

# Complex network approaches to nonlinear time series analysis

Yong Zou<sup>a,\*</sup>, Reik V. Donner<sup>b,\*</sup>, Norbert Marwan<sup>b</sup>, Jonathan F. Donges<sup>b</sup>, Jürgen Kurths<sup>b,c,d,e,f</sup>

<sup>a</sup>*Department of Physics, East China Normal University, Shanghai 200062, China*

<sup>b</sup>*Potsdam Institute for Climate Impact Research, P. O. Box 601203, 14412 Potsdam, Germany*

<sup>c</sup>*Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB243UE, United Kingdom*

<sup>d</sup>*Department of Physics, Humboldt University Berlin, Newtonstraße 15, 12489 Berlin, Germany*

<sup>e</sup>*Department of Control Theory, Nizhny Novgorod State University, Gagarin Avenue 23, 606950 Nizhny Novgorod, Russia*

<sup>f</sup>*Institute of Applied Physics of the Russian Academy of Sciences, 603950 Nizhny Novgorod, Russia*

---

## Abstract

In the last about 10 years, there has been a growing body of literature aiming at the utilization of complex network methods for the characterization of dynamical systems based on time series. While both nonlinear time series analysis and complex network theory are widely considered to be established fields of complex systems sciences with strong links to Nonlinear Dynamics and Statistical Physics, the thorough combination of both approaches has become an active field of research during the last decade, which has allowed addressing fundamental questions regarding the structural organization of nonlinear dynamics as well as the successful treatment of a variety of applications from a broad range of disciplines.

With its three main concepts, phase space / recurrence networks, visibility graphs and transition / Markov chain networks having made their way from abstract concepts to widely used methodologies, the field of time series networks has become mature. These three concepts, as well as several variants thereof, have been studied in great detail regarding their specific properties, potentials and limitations and provided fundamental new insights into the dynamics of complex systems. In addition, these approaches have already found a wide range of applications from such diverse fields as climatology, neurophysiology and economics, to mention only a few examples, demonstrating the great potentials of time series networks to tackling real-world contemporary scientific problems.

To this end, there exists no thorough overview paper covering all existing approaches of time series networks. Consequently, we believe that the time is ripe to deliver such a review covering the methodological foundations, interpretation and (potential) applications of the existing zoo of methods from this field. We are confident that Physics Reports would be an excellent forum for the first review that integrates the state of research on all corresponding concepts that exist so far.

*Keywords:*

---

## Contents

<b>1 Introduction (All)</b>	<b>4</b>
1.1 Nonlinear time series analysis . . . . .	4
1.2 Complex network approaches . . . . .	5
1.3 Outline of the report . . . . .	6

---

\*Corresponding author

Email addresses: [yzou@phy.ecnu.edu.cn](mailto:yzou@phy.ecnu.edu.cn) (Yong Zou), [reik.donner@pik-potsdam.de](mailto:reik.donner@pik-potsdam.de) (Reik V. Donner), [marwan@pik-potsdam.de](mailto:marwan@pik-potsdam.de) (Norbert Marwan), [donges@pik-potsdam.de](mailto:donges@pik-potsdam.de) (Jonathan F. Donges), [kurths@pik-potsdam.de](mailto:kurths@pik-potsdam.de) (Jürgen Kurths)

<b>2 Complex network theory (ALL)</b>	<b>7</b>
2.1 Basic concepts . . . . .	7
2.2 Network characteristics . . . . .	8
2.2.1 Vertex characteristics . . . . .	8
2.2.2 Edge characteristics . . . . .	9
2.2.3 Global network characteristics . . . . .	9
2.3 Stylized facts of complex networks . . . . .	10
2.4 Multiplex and multilayer networks . . . . .	11
2.5 Transformations from time series to network domain . . . . .	11
2.6 Transformations of complex networks to time series . . . . .	12
<b>3 Recurrence networks in phase space (Reik)</b>	<b>13</b>
3.1 Theoretical background . . . . .	13
3.1.1 Phase space and attractor reconstruction . . . . .	13
3.1.2 Recurrences and recurrence plots . . . . .	14
3.2 Types of recurrence networks . . . . .	15
3.2.1 $\varepsilon$ -recurrence networks . . . . .	15
3.2.2 $k$ -nearest neighbor networks . . . . .	16
3.2.3 Adaptive neighbor networks . . . . .	16
3.2.4 Algorithm comparisons and variants . . . . .	17
3.3 Complex network characteristics of RN . . . . .	18
3.3.1 Vertex characteristics . . . . .	19
3.3.2 Edge characteristics . . . . .	20
3.3.3 Global network characteristics . . . . .	21
3.4 Analytical theory of RN . . . . .	22
3.4.1 Preliminaries: random geometric graphs . . . . .	22
3.4.2 Analytical description of $\varepsilon$ -recurrence networks . . . . .	22
3.5 General properties of recurrence networks . . . . .	26
3.5.1 Degree distributions . . . . .	26
3.5.2 Small-world effect . . . . .	27
3.5.3 Assortative vs. disassortative mixing . . . . .	28
3.5.4 Path-based characteristics . . . . .	28
3.5.5 Dimension characteristics by clustering and transitivity . . . . .	29
3.6 Practical considerations . . . . .	31
3.6.1 Dependence on embedding parameters . . . . .	31
3.6.2 Choice of recurrence rate or threshold . . . . .	32
3.6.3 Stability and robustness against noise . . . . .	33
3.6.4 Behavior for larger $\varepsilon$ . . . . .	34
3.7 Multiplex recurrence networks . . . . .	34
3.8 Multilayer networks: coupled networks . . . . .	35
3.8.1 General preliminaries . . . . .	35
3.8.2 Vertex characteristics . . . . .	36
3.8.3 Global characteristics . . . . .	37
3.9 Inter-system recurrence networks . . . . .	38
3.9.1 Cross-recurrence plots . . . . .	38
3.9.2 Coupled networks framework . . . . .	39
3.9.3 Analytical description . . . . .	40
3.9.4 Geometric signatures of coupling . . . . .	42
3.9.5 Examples to characterize flow patterns . . . . .	43
3.10 Joint recurrence networks . . . . .	43
3.10.1 Joint recurrence plots . . . . .	43
3.10.2 Network interpretation . . . . .	43

3.10.3	Network properties and synchronization . . . . .	45
3.11	Other proximity-based time series networks . . . . .	47
3.11.1	Cycle networks . . . . .	47
3.11.2	Correlation networks . . . . .	49
<b>4</b>	<b>Visibility graphs (Jonathan)</b>	<b>50</b>
4.1	Historical roots . . . . .	50
4.2	Algorithmic variants . . . . .	50
4.2.1	Natural visibility graphs . . . . .	50
4.2.2	Horizontal visibility graphs . . . . .	51
4.2.3	Other variants of (H)VG . . . . .	51
4.3	Visibility graph properties . . . . .	52
4.3.1	Degree distributions . . . . .	52
4.3.2	Stochastic vs. deterministic dynamics . . . . .	54
4.3.3	Local network properties . . . . .	55
4.3.4	Global network properties . . . . .	56
4.3.5	Practical considerations . . . . .	57
4.4	Bivariate visibility graph methods . . . . .	58
4.4.1	Visibility graph similarity . . . . .	59
4.4.2	Joint and excess degrees . . . . .	59
4.5	Decomposition of visibility graphs . . . . .	60
4.5.1	Time-directed visibility graphs and characterizations . . . . .	60
4.5.2	Tests for time series irreversibility . . . . .	61
<b>5</b>	<b>Transition networks (Reik)</b>	<b>64</b>
5.1	Markov chains . . . . .	65
5.2	Symbolic encoding of time series . . . . .	65
5.2.1	Threshold-based coarse-graining . . . . .	65
5.2.2	Order pattern-based coarse graining . . . . .	65
5.2.3	Cross and joint ordinal transition networks . . . . .	68
5.2.4	Other approaches . . . . .	69
5.3	Network interpretation of Markov chains . . . . .	70
<b>6</b>	<b>Applications (Reik + Jonathan)</b>	<b>70</b>
6.1	Recurrence networks . . . . .	70
6.1.1	Numerical examples . . . . .	70
6.1.2	Real-world applications . . . . .	74
6.2	Visibility graphs . . . . .	76
6.2.1	Example I: sunspot numbers . . . . .	76
6.2.2	Example II: Asymmetry of sunspots . . . . .	78
6.3	Transition networks . . . . .	80
<b>7</b>	<b>Software implementation – pyunicorn (Jonathan+Norbert)</b>	<b>82</b>
<b>8</b>	<b>Conclusions and future perspectives (All)</b>	<b>84</b>
8.1	Conclusions and discussion . . . . .	84
8.2	Future perspectives . . . . .	84

## 1. Introduction (All)

Artificial Intelligence is driving data in new forms of complexity, leading to the new era of big data [1]. This brings big challenges for researchers from various fields together to extract patterns or new structures from very high volume, high velocity, or high variety. Advanced interdisciplinary data analytics techniques help to capture the hidden structures amidst otherwise chaotic data points, such as machine learning, data mining, statistics, natural language and text processing. In consequence, we transform the messy datasets into something that we can learn fast and make better and faster decisions. There is ample scope for developing new tools for data analysis, based on complex networks and complexity theory.

### 1.1. Nonlinear time series analysis

Time-series analysis is essentially data compression [2]. Given time series, we interpret the underlying dynamical system by a few characteristic numbers that are computed from a large sample of measurements. Therefore, the reduced information as represented by these characteristic numbers must present some specific features of the system. Early approaches of time series analysis heavily relied on linear assumption of the underlying processes, for instance, autoregressive (AR) and moving average (MA) models, which result in almost exponentially decaying auto-correlation function. It is now well accepted that the dynamical laws governing nature or human activities are seldom linear. Nonlinearity is everywhere, for example: (a) phase transition (the melting of ice of a glacier) is an importance signature of nonlinearity in physical systems; (b) animals behave differently (e.g. hunting effort) during time of short food supply versus time of abundant food supply; (c) for most of electronic devices, i.e., transistors, saturation velocity and current are well known nonlinear phenomena. In contrast, the development of nonlinear time series analysis is primarily due to the efforts to overcome the limitations of linear models.

Nonlinear time series analysis is not as well established and is far less well understood than its linear counterpart [3]. The collection of ideas and techniques of nonlinear time series analysis originates from the fast development of dynamical systems theory or so called “chaos theory”, which explores system dynamics by a set of nonlinear difference equations or nonlinear ordinary differential equations. Techniques from chaos theory allow to characterize dynamical systems in which nonlinearities give rise to a complex temporal evolution, for instance, sensitive dependence on initial conditions. Importantly, this concept allows extracting information that cannot be resolved using classical linear techniques such as the power spectrum or spectral coherence.

Since the early stages in 1980s [4], numerous conceptual approaches have been introduced for studying the characteristic features of dynamical systems based on observational time series [3, 5, 6]. The mathematical beauty of this framework analysis is that we characterize the invariant measure in phase space by a number of different ways. Generally speaking, we quantify the system from either geometric or dynamic perspectives. For instance, the correlation dimension (more general for spectrum of  $D_q$  dimension [7]) has been suggested to characterize geometry of phase space; the Lyapunov exponent as a measure for stability of dynamics with respect to infinitesimal perturbations; Kolmogorov-Sinai entropy (or other information theory measures) to quantify uncertainty about the future states of a chaotic trajectory. All these techniques have in common that they quantify certain dynamically invariant phase space properties of the considered system based on temporally discretized realizations of individual trajectories.

One of typical tasks is to perform a precise system characterization from a single time series, which is however not the final goal of most time-series analyses [2]. Here we give a few examples that nonlinear time series analysis can contribute: (1) system characterization from a single time series; (2) discrimination between a signal and some other signals; (3) quantification of various bifurcation transition scenarios to complex dynamics, including period doubling, band merging, intermittency and even more challenging tasks of subtle changes like chaos-to-chaos transitions; (4) detection of change points, regime shifts or tipping points in dynamical properties; (5) test for time series reversibility; (6) noise reduction and filtering; (7) prediction of its future values.

The aforementioned nonlinear time series measures are univariate, i.e., they are applied to single signals measured from individual dynamics. In contrast, bivariate measures are used to analyze pairs of signals

measured simultaneously from two dynamics. There has been considerable interest in the study of the synchronization behavior of coupled chaotic systems, which have been observed in many physical and biological systems [8, 9]. Thus, such bivariate time series analysis measures aim to detect and to distinguish transition forms from non-synchronization states to synchronization ones (for instance, path to phase synchronization, lag synchronization, complete synchronization and generalized synchronization etc). In different synchronization scenarios, it is important to extract not only the coupling strength but also the direction of these couplings, i.e., identifying causal relationships between studied sub-systems. Unraveling the governing functional interactions between sub-systems contained in a large network of complex connectivity topology remains to be a big challenge in modern nonlinear sciences [10, 11]. Various methods have been proposed to extract the statistical associations from data in the literature, for instance, Pearson correlation, mutual information (including time delayed version) [12, 13], Granger causality [14, 15], transfer entropy [16], coupling directions [17–24]. More generally, coupling functions have various forms and we do not expand the discussion here. The readers are referred to several review papers [11, 25, 26].

Nonlinear time series analysis is powerful and useful for many applications, but there are some practical limitations. Some common problems originating from experimental measurements challenge the computations of the nonlinear measures. For instance, most of the existing nonlinear methods hold for low-dimensional dynamics. In reality, very few real-world data sets are measured by perfect sensors operating on low-dimensional dynamics. One has to be aware of non-stationarity, proper choice of embedding parameters, dependence on data length, effects of noise, and irregular samplings etc. [2]. Statistical concerns come also from algorithmic aspects requiring proper choice of parameters. For instance, scaling regimes should be pronounced for implementing linear line fitting to extract dynamic invariant measures of fractal dimensions and Lyapunov exponents, which often influences the results significantly. In addition, computational complexity has to be well evaluated since it varies significantly among these measures. Currently, the choice of algorithmic parameters depend on researchers' experience largely.

In this review report, we will demonstrate that complex network approaches can contribute all aspects that we have discussed above for nonlinear time series analysis. More importantly, complex network approaches can solve partially some fundamental and long standing problems of the existing methods, yielding more robust estimation of dynamical invariants, for instance, using transitivity dimension and local clustering dimension of recurrence networks to approximate fractal dimension of the system [27]; the computation of mean out degree of ordinal transition network and network diameter performs equally well as Lyapunov exponent [28].

### 1.2. Complex network approaches

With the recent increase in available computational capacities and rising data volumes in various fields of science, complex networks have become an interesting and versatile tool for describing structural interdependencies between mutually interacting units [10, 29–31]. Besides “classical” areas of research (such as sociology, transportation systems, computer sciences, or ecology), where these units are clearly (physically) identifiable, the success story of complex network theory has recently lead to a variety of “non-conventional” applications.

One important class of such non-traditional applications of complex network theory are *functional networks*, where the considered connectivity does not necessarily refer to “physical” vertices and edges, but reflects statistical interrelationships between the dynamics exhibited by different parts of the system under study. The term “functional” was originally coined in neuroscientific applications, where contemporaneous neuronal activity in different brain areas is often recorded using a set of standardized EEG channels. These data can be used for studying statistical interrelationships between different brain regions when performing certain tasks, having the idea in mind that the functional connectivity reflected by the strongest statistical dependencies can be taken as a proxy for the large-scale anatomic connectivity of different brain regions [32–34]. Similar approaches have been later utilized for identifying dominant interaction patterns in other multivariate data sets, such as climate data [35–37].

Besides functional networks derived from multivariate time series, there have been numerous efforts for utilizing complex network approaches for quantifying structural properties of individual time series. By means of complex network analysis, the first step is to find a proper network representation for time series,

i.e., an algorithm defining what network vertex and network edge are. Based on the network representations, the rich toolbox of complex network measures [10, 30, 31] provides various quantities that can be used for characterizing the system’s dynamical complexity from networks’ viewpoint and allow discriminating between different types of dynamics (see [38] for a recent mini-review). Several approaches have been proposed for transforming (observational) time series into complex network representations [28, 39–43]. There are three important classes of complex network approaches, which will be the focus of this review report.

The first important class of time series networks make use of similarities or proximity relationships between different parts of a dynamical system’s trajectory [38, 42, 43], including such diverse approaches as cycle networks [39, 44], correlation networks [45], and phase space networks based on a certain definition of nearest neighbors [40]. One especially important example of proximity networks are recurrence networks (RNs) [42, 43], which provide a reinterpretation of recurrence plots in network-theoretic terms and are already widely applied in a variety of fields.

The second class are visibility graphs and related concepts, which characterize some local convexity or record-breaking property within univariate time series data [41, 46, 47]. The standard visibility graph and its various variants have important applications, such as providing new estimates of the Hurst exponent of fractal and multi-fractal stochastic processes [48, 49] or statistical tests for time series irreversibility [50, 51].

The third important class of network approaches are transition networks, which make use of ideas from symbolic dynamics and stochastic processes. Transforming the time series into a transition network is a process of mapping the temporal information into a Markov chain to obtain a compressed or simplified representation of the original dynamics. More specifically, we first discretize the dynamics and then study the transition probabilities between the obtained groups in some Markov chain-like ways [52]. Depending on the particular choice of phase space partitions, we obtain different versions of transition networks. For instance, we construct transition networks by threshold-based coarse graining [38], ordinal pattern-based coarse graining [28, 53].

### 1.3. Outline of the report

The outline of this review report is organized as follows: Section 2, we start with a brief introduction on complex network theory, mainly focusing on the characterization of the structural properties of networks based on the adjacency matrix. All relevant terminologies of network measures will be introduced in this section. We also discuss some important concepts when transforming time series into network representation, namely, the definitions of network vertices and edges.

In Section 3, we focus on recurrence network approaches (RN). We will cover the theoretical background of Poincaré recurrence of dynamical systems and one of most popular visualization techniques of recurrence plots. Furthermore, we summarize the current state of knowledge on the theoretical foundations and potential applications of RN approaches to nonlinear time series. We demonstrate that this type of time series networks naturally arise as random geometric graphs in the phase space of dynamical systems, which determines their structural characteristics and gives rise to a dimensionality interpretation of clustering coefficients and related concepts. Beyond the single-system case, we also provide a corresponding in-depth discussion of cross- and joint recurrence plots from the complex network viewpoint. As a new aspect not previously reported in the literature, we provide a first-time theoretical treatment of a unification of single-system and cross-recurrence plots in a complex network context. Moreover, we discuss some new ideas related to the utilization of multiplex and multilayer multivariate recurrence network-based approaches for studying geometric signatures of coupling and synchronization processes.

In Section 4, both the standard visibility graphs (VG) and horizontal visibility graphs (HVG) will be reviewed. We start with the historical roots of the VG algorithm and summarize the main variants of it. More importantly, we will assess some conjectures of theoretical predictions of (H)VG properties in stochastic and deterministic processes. Some practical considerations will be thoroughly discussed when applying (H)VG analysis to experimental time series. In addition, we will discuss the generalization of (H)VG analysis from a univariate to bivariate time series, for instance, multiplex (H)VGs. We will further show that a decomposition of (H)VGs into time forward (outgoing) and backward (outgoing) directions helps

to test irreversibility of the underlying time series. Some practical issues will be reviewed when performing (H)VG analysis for experimental time series.

In Section 5, we introduce the construction of transition networks by proper coarse graining of phase space and our main focus will be ordinal pattern transition networks. The concept of ordinal pattern transition networks can be traced back to identifying ordinal patterns of time series [54]. We will review the ordinal transition networks of [28] and focus more on the generalizations to multivariate time series [55]. A series of systematic investigations of ordinal methods has been conducted in irregularly sampled time series [56–58], which shows high potential for studies of experimental observation data from climate sciences [59].

In Section 6, we review several applications of network approaches to different situations. In particular, we will review the respective successful applications of recurrence network approaches, visibility graphs, and ordinal pattern transition networks to various nonlinear time series, ranging from numerical example dynamical systems, non-stationary stochastic processes, observational sunspot time series.

In Section 7, we discuss about the software implementations. In particular, we introduce the Python software package `pyunicorn`, which implements methods from both complex network theory and nonlinear time series analysis, and unites these approaches in a performant, modular and flexible way [60].

In the last section 8, we summarize the main topics that we have reviewed in this report. More important, we will pinpoint some important directions for the future research. We wish to emphasize that it is just the beginning of complex network approaches for nonlinear time series analysis and there are lots of interesting topics that need to be explored in the near future.

## 2. Complex network theory (ALL)

We provide a brief introduction of the recent developments in the characterization of the structural properties of a network, focusing on definitions, notations, and basic quantities that are most often used to describe the topologies of networks that are reconstructed from time series. More general descriptions of complex networks can be found in the literature in a number of review articles [10, 29–31] and books [61, 62], which the reader may find useful to consult.

### 2.1. Basic concepts

A complex network is often represented as a graph  $G = (V, E)$  which consists of two sets  $V$  and  $E$ , where  $V$  is the set of vertices (nodes or points) of the graph  $G$ , and  $E$  is the set of edges (links, edges or lines) representing the connection pairs of elements of  $V$  [30]. Each vertex is identified by an integer index  $i = 1, 2, \dots, N$  and each edge is identified by a pair  $e_{ij}$  connecting two vertices  $i$  and  $j$ . A graph  $G$  is called undirected if an edge from vertex  $i$  to  $j$  as denoted by  $e_{ij}$  is equivalent to the edge of  $e_{ji}$  from vertex  $j$  to  $i$ ,  $e_{ij} = e_{ji}$ . On the other hand, in a directed graph,  $e_{ij} \neq e_{ji}$ . A graph may contain loops, i.e. edges from a vertex to itself or multiple edges, i.e. pairs of vertices connected by more than one edges which needs another attribute  $w$  to quantify the weight of  $e_{ij}$ , i.e. denoted by  $w_{ij}$ . A weighted digraph can be completely described by its weight matrix  $W$  so that each entry  $w_{ij}$  expresses the weight of the connection from vertex  $i$  to vertex  $j$ . In this chapter, we first consider undirected and unweighted graphs no need  $w_{ij}$  which will be the case for most of the network approaches for time series. Weighted graphs will receive a special attention and will be discussed in Section 5 for transition networks.

An unweighted graph can be achieved by introducing a proper threshold to the weighted counterpart  $W$  [30], yielding the binary matrix  $A$ . The entry of  $A$  are computed by comparing the corresponding weight matrix  $W$  with a specified threshold  $T$  and we have  $a_{ij} = 1$  if  $w_{ij} > T$ , otherwise,  $a_{ij} = 0$ . The resulting matrix  $A$  is called adjacency matrix of the unweighted graph. Further introduction of symmetry to  $A$  i.e. identifying  $e_{ij} = e_{ji}$  results in an undirected graph. Such an undirected, unweighted graph is also called a simple graph which is the mathematical framework for the structural characterization of complex networks in the following sections.

Depending on the particular mappings transforming a given time series into a complex network, the resulting adjacency matrix  $A$  often depends on some proper chosen algorithmic parameters, for instance, the threshold value  $\varepsilon$  of the recurrence network approach. More importantly, we often have some particular

interpretations for network measures, for instance, in terms of geometry of dynamical system. In the following, we first introduce some general formula in characterizing network structures based on  $A$ . More specific discussions of network measures in terms of the particular network transforming methods will be presented in the later sections.

In addition to the concepts of vertices and edges, the third important concept in complex network theory is the notion of paths. A *path* between two specified vertices  $i$  and  $j$  is an ordered sequence of edges starting at  $i$  and ending at  $j$ , with its *path length*  $d_{ij}$  given by the number of edges in this sequence. There are also various measures characterizing the structural properties of networks based on paths, which will be briefly reviewed here as well.

## 2.2. Network characteristics

### 2.2.1. Vertex characteristics

There are various measures to characterize the structures of a complex network, quantifying the importance of either a vertex or an edge in terms of a particular definition of network property. The conceptually simplest measure characterizing the connectivity properties of a single vertex in a complex network is the *degree* or *degree centrality*

$$k_v = \sum_{i=1}^N A_{iv}, \quad (1)$$

which simply counts the number of edges associated with a given vertex  $v$ . It is also convenient to introduce the averaged degree  $\rho_v = \frac{1}{N-1} k_v$  as the local connectivity density of  $v$ . Furthermore, the topological characterization of the graph  $G$  can be obtained in terms of the degree distribution  $p(k)$ , defined as the probability that a vertex chosen uniformly at random has degree  $k$  or, equivalently, as the fraction of vertices in the graph having degree  $k$ . Note that the variable  $k$  assumes non-negative integer values. Degree distribution  $p(k)$  is often used to classify complex networks, for instance, a scale free network is characterized by  $p(k) \sim k^{-\gamma}$ , which will further discussed in Sec. 2.3.

In order to characterize the density of connections among the neighbors of a given vertex  $v$ , we can utilize the *local clustering coefficient*

$$\mathcal{C}_v = \frac{1}{k_v(k_v - 1)} \sum_{i,j=1}^N A_{vi} A_{ij} A_{jv}, \quad (2)$$

which measures the fraction of pairs of vertices in the neighborhood of  $v$  that are mutually connected.

While degree and local clustering coefficient characterize network structures on the local and meso-scale, there are further vertex characteristics that make explicit use of the concept of shortest paths and, thus, provide measures relying on the connectivity of the entire network. Two specific properties of this kind are the *closeness* or *closeness centrality*

$$c_v = \left( \frac{1}{N-1} \sum_{i=1}^N d_{vi} \right)^{-1}, \quad (3)$$

which gives the inverse arithmetic mean of the shortest path lengths  $d_{vi}$  between vertex  $v$  and all other vertices, and the *local efficiency*

$$e_v = \frac{1}{N-1} \sum_{i=1}^N d_{vi}^{-1}, \quad (4)$$

which gives the inverse harmonic mean of these shortest path lengths. Notably, the latter quantity has the advantage of being well-behaved in the case of disconnected network components, where there are no paths between certain pairs of vertices (i.e.,  $d_{ij} = \infty$ ). In order to circumvent divergences of the closeness due to the existence of disconnected components, it is convenient to always set  $d_{ij}$  to the highest possible value of  $N - 1$  for pairs of vertices that cannot be mutually reached. Both  $c_v$  and  $e_v$  characterize the geometric centrality of vertex  $v$  in the network, i.e., closeness and local efficiency exhibit the highest values for such vertices which are situated in the center of the networks.

Another frequently studied path-based vertex characteristic is the so-called *betweenness* or *betweenness centrality*, which measures the fraction of shortest paths in a network traversing a given vertex  $v$ . Let  $\sigma_{ij}$  denote the total number of shortest paths between two vertices  $i$  and  $j$  and  $\sigma_{ij}(v)$  the multiplicity of these paths that include a given vertex  $v$ , betweenness centrality is defined as

$$b_v = \sum_{i,j=1; i,j \neq v}^M \frac{\sigma_{ij}(v)}{\sigma_{ij}}. \quad (5)$$

Betweenness centrality is commonly used for characterizing the importance of vertices for information propagation in networks.

### 2.2.2. Edge characteristics

In contrast to vertices, whose properties can be characterized by a multitude of graph characteristics, there are fewer measures that explicitly relate to the properties of edges or, more general, pairs of vertices. One such measure is the *matching index*, which quantifies the overlap of the network neighborhoods of two vertices  $v$  and  $w$ :

$$m_{vw} = \frac{\sum_{i=1}^N A_{vi} A_{wi}}{k_v + k_w - \sum_{i=1}^N A_{vi} A_{wi}}. \quad (6)$$

While the concept of matching index does not require the presence of an edge between two vertices  $v$  and  $w$ , there are other characteristics that are explicitly edge-based. To this end, we only mention that the concept of betweenness centrality can also be transferred to edges, leading to the *edge betweenness* measuring the fraction of shortest paths on the graph traversing through a specific edge  $(v, w)$ :

$$b_{vw} = \sum_{i,j=1; i,j \neq v,w}^M \frac{\sigma_{ij}(v, w)}{\sigma_{ij}}, \quad (7)$$

where  $\sigma_{ij}(v, w)$  gives the total number of shortest paths between two vertices  $i$  and  $j$  that include the edge  $(v, w)$ . If there is no edge between two vertices  $v$  and  $w$ , we set  $b_{vw} = 0$  for convenience.

### 2.2.3. Global network characteristics

Some, but not all useful global network characteristics can be derived by averaging certain local-scale (vertex) properties. Prominently, the *edge density*

$$\rho = \frac{1}{N} \sum_{v=1}^N \rho_v = \frac{1}{N(N-1)} \sum_{i,j=1}^N A_{ij} \quad (8)$$

is defined as the arithmetic mean of the degree densities of all vertices and characterizes the fraction of possible edges that are present in the network.

In a similar way, we consider the arithmetic mean of the local clustering coefficients  $\mathcal{C}_v$  of all vertices, resulting in the (Watts-Strogatz) *global clustering coefficient* [63]

$$\mathcal{C} = \frac{1}{N} \sum_{v=1}^N \mathcal{C}_v = \frac{1}{N} \sum_{v=1}^N \frac{\sum_{i,j=1}^N A_{vi} A_{ij} A_{jv}}{k_v(k_v - 1)}. \quad (9)$$

The global clustering coefficient measures the mean fraction of triangles that include the different vertices of the network.

Notably, in the case of a very heterogeneous degree distribution, the global clustering coefficient will be dominated by contributions from the most abundant type of vertices. For example, for a scale-free network with  $p(k) \sim k^{-\gamma}$ , vertices with small degree will contribute predominantly, which can lead to an underestimation of the actual fraction of triangles in the network, since  $\mathcal{C}_v = 0$  if  $k_v < 2$  by definition. In

order to correct for such effects, Barrat and Weigt [64] proposed an alternative definition of the clustering coefficient, which is nowadays frequently referred to as *network transitivity* [10] and is defined as

$$\mathcal{T} = \frac{\sum_{v,i,j=1}^N A_{vi}A_{ij}A_{jv}}{\sum_{v,i,j=1}^N A_{vi}A_{jv}}. \quad (10)$$

Finally, turning to shortest path-based characteristics, we define the *average path length*

$$\mathcal{L} = \frac{1}{N(N-1)} \sum_{i,j=1}^N d_{ij} = \frac{1}{N} \sum_{v=1}^N c_v^{-1} \quad (11)$$

as the arithmetic mean of the shortest path lengths between all pairs of vertices, and the *global efficiency*

$$\mathcal{E} = \left( \frac{1}{N(N-1)} \sum_{i,j=1}^N d_{ij}^{-1} \right)^{-1} = \left( \frac{1}{N} \sum_{v=1}^N e_v \right)^{-1} \quad (12)$$

as the associated harmonic mean. Notably, the average path length can be rewritten as the arithmetic mean of the inverse closeness, and the global efficiency as the inverse arithmetic mean of the local efficiency.

### 2.3. Stylized facts of complex networks

Erdős and Rényi [65] introduced a model to generate random graphs consisting of  $N$  vertices and  $M$  edges. Starting with  $N$  disconnected vertices, the network is constructed by the addition of  $L$  edges at random, avoiding multiple and self connections. Another similar model defines  $N$  vertices and a probability  $p$  of connecting each pair of vertices. The latter model is widely known as Erdős-Rényi (ER) model. For the ER model, in the large network size limit ( $N \rightarrow \infty$ ), the average number of connections of each vertex  $\langle k \rangle$  is given by  $\langle k \rangle = p(N-1)$ , where  $p$  is fixed and often chosen as a function of  $N$  to keep  $\langle k \rangle$  fixed. For this model, degree distribution  $p(k)$  is a Poisson distribution.

In regular hypercubic lattices in  $D$  dimensions, the mean number of vertices one has to pass by in order to reach an arbitrarily chosen vertex, grows with the lattice size as  $N^{1/d}$ . Conversely, in most of the real networks, despite of their often large size  $N$ , there is a relatively short path between any two vertices. This feature is known as the small-world property and is mathematically characterized by an average shortest path length  $\mathcal{L}$  (Eq. (11)) that depends at most logarithmically on the network size  $N$  [10, 29, 31]. As variance with random graphs, the small world property in real networks is often associated with the presence of clustering, denoted by high values of the clustering coefficient. For this reason, Watts and Strogatz [63] have proposed to define small-world networks (WS model) as those networks having both a small value of  $\mathcal{L}$ , like random graphs, and a high clustering coefficient  $\mathcal{C}$ , like regular lattices. This model is situated between an ordered finite lattice and a random graph, presenting the small world property and high clustering coefficient.

Barabási and Albert [66] showed that the degree distribution  $p(k)$  of many real systems is characterized by an uneven distribution. Instead of the vertices of these networks having a random pattern of connections with a characteristic degree, as with the ER and WS models, some vertices are highly connected while others have few connections, with the absence of a characteristic degree. More specifically, the degree distribution has been found to follow a power law for large  $k$ , namely,  $p(k) \sim k^{-\gamma}$ . These networks are called scale free networks (BA model), which is captured by a pronounced linear regime in the double logarithmic plot of  $p(k)$ . The two important ingredients of the BA model is growth and preferential attachment. A proper statistical justification of scale free properties of real networks is a non-trivial task because of effects from finite sizes, intrinsic noise and finite sampling etc [67].

In addition, a large number of real networks are correlated in the sense that the probability that a node of degree  $k$  is connected to another node of degree, say  $k'$ , depends on  $k$ . This problem can be quantified by the average nearest neighbors degree of a vertex  $i$ , or simply the assortativity coefficient  $\mathcal{R}$  correlation coefficient between the degree sequences [68]. In assortative networks, the vertices tend to connect to their connectivity

peers ( $\mathcal{R} > 0$ ), while in disassortative networks vertices with low degrees are more likely connected with highly connected ones ( $\mathcal{R} < 0$ ).

All of these stylized facts of complex networks will be discussed in the respective framework when introducing different network construction algorithms for nonlinear time series.

#### 2.4. Multiplex and multilayer networks

Most complex systems include multiple subsystems and layers of connectivity and they evolve, adapt and transform through internal and external dynamic interactions affecting the subsystems and components at both local and global scale. Predicting their multiscale and multicomponent dynamics is a difficult challenge for the recent research [69]. The issues posed by the multiscale modeling of both natural and artificial complex systems call for a generalization of the “traditional” network theory, by developing a solid foundation and the consequent new associated tools to study multilayer and multicomponent systems in a comprehensive fashion.

We follow with the formal definition of a multilayer network [69] that is a pair  $\mathcal{M} = (\mathcal{G}, \mathcal{C})$  where  $\mathcal{G} = \{G_\alpha; \alpha \in [1, 2, \dots, M]\}$  is a family of graphs  $G_\alpha = (V_\alpha, E_\alpha)$  and

$$\mathcal{C} = \{E_{\alpha\beta} \subseteq V_\alpha \times V_\beta; \alpha, \beta \in [1, 2, \dots, M], \alpha \neq \beta\} \quad (13)$$

is the set of interconnections between nodes of different layers  $G_\alpha$  and  $G_\beta$  with  $\alpha \neq \beta$ . The elements of  $\mathcal{C}$  are called crossed layers and the elements of each  $E_\alpha$  are called intralayer connections of  $\mathcal{M}$  in contrast with elements of each  $E_{\alpha\beta}$  that are called interlayer connections. By using this representation, we simultaneously consider edges that are inside different layers and edges that connect different layers. A multiplex network is a special type of multilayer network that each layer shares the same set of vertices, namely,  $V_1 = V_2 = \dots = V_M$  and the only possible type of interlayer connections are those in which a given node is connected to its counterpart nodes in the rest of layers. In other words, a multiplex network consists of a fixed set of vertices connected by different types of edges [69].

The readers are referred to [69, 70] for a more thorough review on multilayer networks. Furthermore, it is important to remark that concept of multilayer networks has been extended to other relevant notations, for instance, network of networks, interacting or interconnected networks, multidimensional networks, interdependent networks, multilevel networks, and hypernetworks etc.

#### 2.5. Transformations from time series to network domain

The great success of network theory in various fields of research has motivated first attempts to generalize this concept for a direct application to time series [38–43, 45, 71–75]. Then, based on network representation of time series, important complementary features of dynamical systems (i.e., properties that are not captured by existing methods of time series analysis) can be resolved. By means of complex network analysis, the first step is to find a proper network representation for time series, i.e., an algorithm defining what network vertex and network edge are. To this end, several approaches have been proposed. These methods can be roughly distinguished into three classes (see Tab. 1), which are based on

- (i) mutual proximity of different segments of a time series (proximity networks),
- (ii) convexity of successive observations (visibility graphs), and
- (iii) transition probabilities between discrete states (transition networks).

With the exception of visibility graphs, all approaches are related with the concept of recurrence [38]. This is particularly evident for proximity networks, where connectivity is defined in a data-adaptive local way, *i.e.*, by considering distinct regions with a varying center at a given vertex in either the phase space itself or an abstract proximity space. In contrast, for transition networks, the corresponding classes are rigid, *i.e.*, determined by a fixed coarse-graining of the phase space. The distinction between both classes of approaches is conceptually similar to the duality of symbolic time series analysis (*i.e.*, time series analysis

Table 1: Summary of the definitions of vertices and the criteria for the existence of edges in existing complex network approaches.

Method	Vertex	Edge	Directedness
Proximity networks			
<i>Cycle networks</i>	Cycle	Correlation or phase space distance between cycles	undirected
<i>Correlation networks</i>	State vector	Correlation coefficient between state vectors	undirected
<i>Recurrence networks</i>			
<i>k-nearest neighbor networks</i>	State (vector)	Recurrence of states (fixed neighborhood mass)	directed
<i>adaptive nearest neighbor networks</i>	State (vector)	Recurrence of states (fixed number of edges)	undirected
<i><math>\varepsilon</math>-recurrence networks</i>	State (vector)	Recurrence of states (fixed neighborhood volume)	undirected
Visibility graphs	Scalar state	Mutual visibility of states	undirected
Transition networks	Discrete state	Transitions between states	directed

based on a coarse-graining of the dynamics), which may both be used for estimating similar dynamical invariants such as entropies and mutual information.

Among the three classes of methods listed above, the largest group of concepts is given by proximity networks, where the mutual closeness or similarity of different segments of a trajectory can be characterized in different ways. Consequently, there are different types of such proximity networks (see Tab. 1): cycle networks, correlation networks, and recurrence networks. However, all these methods are characterized by two common general properties: Firstly, the resulting networks are invariant under relabeling of their vertices in the adjacency matrix. Hence, the topological characteristics of proximity networks yield nonlinear measures that are invariant against permutation of vertices. In this respect, the network-theoretic approach is distinctively different from traditional methods of time series analysis where the temporal order of observations does explicitly matter. Secondly, we have to point out that particularly proximity networks are spatial networks. In particular, recurrence networks are embedded in the phase space of the considered system, with distances being defined by one of the standard metrics (*e.g.*, Euclidean, Manhattan, etc.). Similar considerations apply to other types of proximity networks as well.

In this report, we will provide an exhaustive review on complex network approaches for nonlinear time series analysis, focusing on some important transformation methods that have been widely applied to various real experimental data analysis, in particular, recurrence networks, visibility graphs, and transition networks and their applications. Certainly, we will also discuss some algorithmic variants of these concepts and the relevant methods wherever it is necessary.

## 2.6. Transformations of complex networks to time series

To investigate how much information is encoded in a network model of a time series, some studies have been undertaken to recover the original time series from the network, to use the network to reconstruct the phase space topology of the original system, or to generate new time series from the networks and compare these with the original [76–79]. This inverse problem of getting back from the network adjacency matrix to time series of the underlying dynamical system remains a big challenge, which certainly has many applications. In general transformations of complex networks to time series are not so easy. In networks, the order of the vertices can be exchanged without affecting the network topology. But for this reconstruction of the trajectory the temporal order of the nodes are required.

Some few algorithms have been proposed so far to reconstruct time series from networks. For instance, under the condition for the reconstructability, Thiel *et al* proposed an algorithm to reconstruct time series from recurrence plot [80]. The reconstructed attractor shows topological equivalence with the original attractor. Furthermore, based on *k*-nearest neighbor networks, the topological properties of the underlying time series have been reconstructed by an inversion algorithm as presented in [77]. Recently, it has been shown that *k*-nearest neighbor and  $\varepsilon$ -recurrence networks can be viewed as identical structures under a change of (equivalent) metrics [81]. Based on this fact, an improved inversion algorithm is proposed in [81], which further validates the use of complex networks as a valid means of studying dynamical systems, whilst also revealing an equivalence between  $\varepsilon$ -recurrence and *k*-nearest neighbor classes of complex networks. Algorithms based on random walks have been proposed in the literature. For instance, a random walk

algorithm has been proposed in [82], which further compares the performance of RNs and adaptive  $k$ -nearest neighbor networks. The performance of these algorithms have been compared in [83]. Recently, a constrained random walk algorithm has been proposed to regenerate time series from ordinal transition networks [79].

### 3. Recurrence networks in phase space (Reik)

In this section, we introduce and discuss RNs as an alternative framework for studying recurrences in phase space from a geometric point of view. We start with the basic setting suitable for single dynamical systems, followed by some detailed considerations on two different multivariate generalizations, taking advantage of corresponding recent extensions [84, 85] of the recurrence plot (RP) concept [86, 87]. Moreover, we provide a short overview on complex network characteristics and their meaning for RNs, highlighting the type of information that can be obtained using this approach – as opposed to other recurrence based techniques like recurrence quantification analysis (RQA) [88, 89], recurrence time statistics, or estimation of dynamical invariants from RPs.

#### 3.1. Theoretical background

##### 3.1.1. Phase space and attractor reconstruction

We start with a (possibly multivariate) time series  $\{x_i\}_{i=1}^N$  with  $x_i = x(t_i)$ , which we interpret as a finite representation of the trajectory of some (deterministic or stochastic) dynamical system. For a discrete system (map), the sampling of the time series is directly given by the map, whereas for a continuous-time system, the time series values correspond to a temporally discretized sampling of a finite part of one trajectory of the system determined by some initial conditions. In the case of observation functions not representing the full variety of dynamically relevant components, we additionally assume that attractor reconstruction has been performed reasonable well (e.g., using time-delay embedding or some related technique) [3, 90–92].

More specifically, given a scalar time series  $\{x_i\}$  ( $i = 1, \dots, N$ ), we first convert the data into state vectors in some appropriately reconstructed phase space. A common method from dynamical systems theory to define such a phase space is time-delay embedding [92]. In fact, the concept of a phase space representation rather than a “simple” time or frequency domain approach is the hallmark of many methods of nonlinear time series analysis, requiring embedding as the first step. Here, we define  $\mathbf{x}_i = (x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau})$  to obtain an  $m$ -dimensional time-delay embedding of  $x_i$  with embedding delay  $\tau$  for obtaining state vectors in phase space [92]. It has been proven that for deterministic dynamical systems, the thus reconstructed phase space is topologically equivalent to the original space if  $m > 2D_F$ , where  $D_F$  is the fractal dimension of the support of the invariant measure generated by the dynamics in the true (but often at most partially observed) state space. Note that  $D_F$  can be much smaller than the dimension of the underlying original (physical) phase space spanned by all relevant system variables.

From a practical perspective, when analyzing a scalar time series of whatever origin, neither embedding dimension  $m$  nor delay  $\tau$  are known a priori. The false nearest-neighbors (FNN) method [91] was introduced to derive a reasonable guess of how to choose  $m$  based on studying whether or not proximity relations between state vectors are lost when the embedding dimension is successively increased. If a reasonable embedding dimension is found, all dynamically relevant coordinates of the system are appropriately represented, so that all proximity relationships are correct and not due to lower-dimensional projection effects.

In a similar spirit, a delay  $\tau$  may be appropriate when the auto-covariance approaches zero, i.e., corresponding to the first root of the auto-correlation function (ACF) of a time series. This minimizes the linear correlation between the components but does not have to mean they are independent. However, the converse is true: if two variables are independent they will be uncorrelated. Therefore, another well established possibility for determining  $\tau$  is to use time-delayed mutual information [90].

The aforementioned approaches to determining  $m$  and  $\tau$  commonly work well for data from deterministic dynamical systems. Let us first assume in the following sections that the proper embeddings have been obtained and then we will discuss the effects of embedding on the reconstructed networks in Sec. 3.6. This is an important issue when dealing with experimental time series and we have to first check if the appropriate

embeddings are applicable. Practically, we need to show the dependence of any analysis on the embeddings explicitly.

### 3.1.2. Recurrences and recurrence plots

Recurrence of states, in the meaning that states are arbitrary close after some time, is a fundamental property of deterministic dynamical systems and is typical for nonlinear or chaotic systems [93, 94]. From the set of (original or reconstructed) state vectors representing a discrete sampling of the underlying system's trajectory (e.g., the chaotic attractor of a dissipative system), recurrences can be visualized by recurrence plots (RP), introduced by Eckmann *et al.* in [86, 87] and the RP is represented by the recurrence matrix  $\mathbf{R}(\varepsilon)$  in the standard way as

$$R_{ij}(\varepsilon) = \Theta(\varepsilon - \|x_i - x_j\|), \quad (14)$$

where  $\|\cdot\|$  can be any norm in phase space (e.g., Manhattan, Euclidean, or maximum norm). For convenience, we will use the maximum norm in all following examples. A RP enables us to investigate the recurrences of the  $m$ -dimensional phase space trajectory through a two-dimensional representation  $R_{ij}$  with dots.

The basic principle is illustrated in Fig. 1 for one realization of the Lorenz system

$$\frac{d}{dt} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sigma(y - x) \\ x(r - y) \\ xy - \beta z \end{pmatrix}. \quad (15)$$

Further definitions of recurrences add dynamical aspects, such as local rank orders or strictly parallel evolution of states (parallel segments of phase-space trajectory considered in iso-directional RPs [95]). For a more detailed overview, we refer to [87].

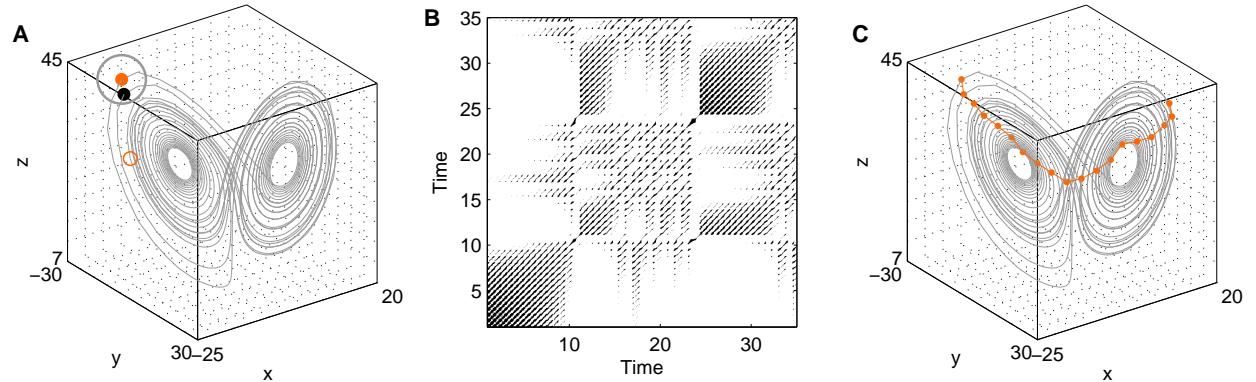


Figure 1: Basic concepts beyond recurrence plots and the resulting recurrence networks, exemplified for one realization of the Lorenz system (Eq. (15)) with the parameters  $r = 28$ ,  $\sigma = 10$  and  $\beta = 8/3$  (sampling time  $\Delta t = 0.02$ , original coordinates, no embedding, recurrences defined based on a fixed threshold  $\varepsilon = 5.0$  using maximum norm). (A) A state at time  $i$  (red dot) is recurrent at another time  $j$  (black dot) when the phase space trajectory visits its close neighborhood (gray circle). This is marked by value 1 in the recurrence matrix at  $(i, j)$ . States outside of this neighborhood (small red circle) are marked with 0 in the recurrence matrix. (B) Graphical representation of the corresponding recurrence matrix (recurrence plot) and adjacency matrix (modulo main diagonal). (C) A particular path in the recurrence network for the same system embedded in the corresponding phase space. Reproduced from [38]. We need to put panel (C) into Fig. 2(b).

RPs of dynamical systems with different types of dynamics exhibit distinct structural properties, which can be characterized in terms of their associated small-scale as well as large-scale features [87]. The study of recurrences by means of RPs has become popular with the introduction of recurrence quantification analysis (RQA) [88, 96]. The initial purpose of this framework has been to introduce measures of complexity which distinguish between different appearances of RPs [97], since they are linked to certain dynamical properties of the studied system. RQA measures use the distribution of small-scale features in the RP, namely individual recurrence points as well as diagonal and vertical line structures. RQA as a whole has been proven to

constitute a very powerful technique for quantifying differences in the dynamics of complex systems and has meanwhile found numerous applications, *e.g.*, in ecology [98], engineering [99], geo- and life sciences [100, 101], or protein research [102, 103]. For a more comprehensive review on the potentials of this method, we refer to [104, 105]. In addition, we would like to remark that even dynamical invariants, like the  $K_2$  entropy and mutual information, or dimensions (information and correlation dimensions  $D_1$ ,  $D_2$ ) can be efficiently estimated from RPs [87, 106]. Moreover, RPs have also been successfully applied to study interrelations, couplings, and phase synchronization between dynamical systems [85, 107–111].

In order to highlight the domains of recurrences in the RPs, some sophisticated algorithms have been proposed recently. For example, Pham *et al* introduced fuzzy recurrence plots, which determines an optimally relation of the phase space states to a number of predefined clusters [112]. This algorithm highlights the recurrence regions giving better visualizations. Recently another algorithm has been proposed to search the recurrence domains in [113]. In particular, intersecting  $\varepsilon$ -balls around sampling points are merged into cells of a phase space partition and a maximum entropy principle defines the optimal size of intersecting balls. This adaptive algorithm in obtaining phase partitions performs better than techniques based on Markov chains which require an ad hoc partition of the system’s phase space. In the same line of research, another algorithm has been proposed in [114] to capture the recurrence density structures in the plot. The computation complexity of RQA measures has been recently evaluated [115].

### 3.2. Types of recurrence networks

#### 3.2.1. $\varepsilon$ -recurrence networks

We can re-interpret the mathematical structure  $\mathbf{R}(\varepsilon)$  as the adjacency matrix of some adjoint complex network embedded in phase space by setting

$$\mathbf{A}(\varepsilon) = \mathbf{R}(\varepsilon) - \mathbf{1}_N, \quad (16)$$

where  $\mathbf{1}_N$  is the  $N$ -dimensional identity matrix. The complex network defined this way is called  *$\varepsilon$ -recurrence network (RN)*, as opposed to other types of proximity-based networks in phase space making use of different definitions of geometric closeness, *e.g.*, considering  $k$ -nearest neighbors [38]. Specifically, the sampled state vectors  $\{x_i\}$  are interpreted as vertices of a complex network, which are connected by undirected edges if they are mutually close in phase space (*i.e.*, describe recurrences). Notably, the binary matrix  $\mathbf{A}(\varepsilon)$  retains the symmetry properties of  $\mathbf{R}(\varepsilon)$ , which implies that the RN is a *simple graph*, *i.e.*, a complex network without multiple edges or self-loops (note that  $A_{ii} = 0$  according to definition (16)). We show an example of an unweighted  $\varepsilon$ -RN network in 2.

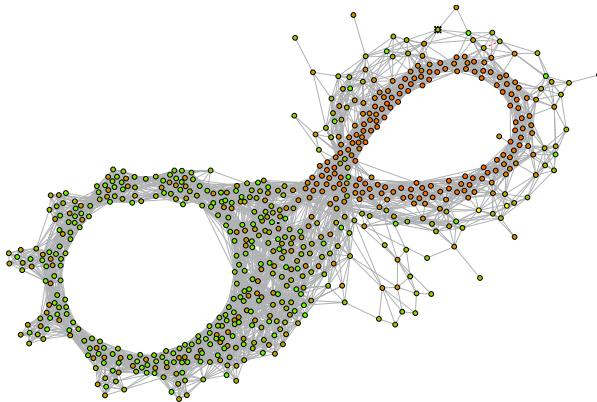


Figure 2: (A) A graphical representation of the Lorenz attractor based on the recurrence matrix represented in Fig. 1. The color of the vertices corresponds to their temporal order (from orange to bright green). (B) A particular path in the RN for the same system embedded in the corresponding phase space. Reproduced from [38]. [We need Fig. 1\(C\) here.](#)

In an  $\varepsilon$ -RN, a *path* between two specified vertices  $i$  and  $j$  is an ordered sequence of edges starting at  $i$  and ending at  $j$ , with its *path length*  $d_{ij}(\varepsilon)$  given by the number of edges in this sequence. An example of a

path is shown in Fig. 2. In the context of RNs, we can thus understand a path as a sequence of mutually overlapping  $\varepsilon$ -balls<sup>1</sup>  $B_\varepsilon(x_i), B_\varepsilon(x_{k_1}), \dots, B_\varepsilon(x_{k_{l_{i,j}-1}}), B_\varepsilon(x_j)$ , where

$$B_\varepsilon(x) = \{y \mid \|x - y\| < \varepsilon\}$$

is an open set describing a volume with maximum distance  $\varepsilon$  (measured in a given norm) from  $x$ , and  $B_\varepsilon(x_i) \cap B_\varepsilon(x_{k_1}) \neq \emptyset, \dots, B_\varepsilon(x_{k_{l_{i,j}-1}}) \cap B_\varepsilon(x_j) \neq \emptyset$ .

Following these considerations, a *shortest path* is a minimum sequence of edges (mutually overlapping  $\varepsilon$ -balls) between two fixed vertices (state vectors)  $i$  and  $j$ . Note that a shortest path does not need to be unique. In turn, due to the discrete character of a network, it is rather typical that there are multiple shortest paths between some specific pair of vertices. In what follows, the shortest path length will be denoted as  $d_{ij}$ , and the multiplicity of such shortest paths as  $\sigma_{ij}$ .

We have to emphasize that the network-theoretic concept of a path on a given graph is distinctively different from the trajectory concept that records the causal dynamic evolution of the system [43]. Furthermore, RNs based on Eq. (16) can be generalized by withdrawing the application of a specific threshold, which leads to weighted networks and unthresholded RPs (distance plots), respectively. For example, the unthresholded RP obtained from one trajectory of a given dynamical system may be re-interpreted as the connectivity matrix of a fully coupled, weighted network.

Most of RN analysis have focused on the network representation using the adjacency matrix and the extraction of new network-theoretic measures, which will be reviewed below. We have to emphasize that the adjacency matrix provides information of vertices and edges, but its graph structural layouts can take variable forms. Network visualization is a non-trivial task and there are many tools in computer science, for example, spring-based layout systems, spectral layout method, and tree layout algorithms etc. These algorithms have been well integrated in many popular network visualization packages, i.e., Mathematica, Gephi, and NetworkX. The network graph shown Fig. 2 has been created using the software package GUESS using a force directed placement algorithm. In [116], Yang *et al* proposed to use the spring-electrical model to explore the self-organized geometry of RNs. In this algorithm, they simulate the recurrence network as a physical system by treating the edges as springs and the nodes as electrically charged particles. Then, force-directed algorithms are developed to automatically organize the network geometry by minimizing the system energy. It has been shown that this self-organized process recovers the attractor of a dynamical system, which provides insights for attractor reconstruction from the adjacency matrix [77, 80].

### 3.2.2. *k*-nearest neighbor networks

Besides the recurrence definition based on a fixed distance threshold  $\varepsilon$  in phase space (i.e., equal neighborhood volumes around all available state vectors), there are alternative ways for defining recurrences and, hence, RPs and RNs. For example, the original definition of a RP by Eckmann *et al.* [86] makes use of *k*-nearest neighbors (i.e., a fixed probability mass of the considered neighborhoods). Re-interpreting the resulting recurrence matrix as the adjacency matrix of a complex network leads to a different type of RN [72], typically referred to as *k*-nearest neighbor network. Since in this definition, the neighborhood relation is not symmetric (i.e.,  $x_j$  being among the  $k$  nearest neighbors of  $x_i$  does not imply  $x_i$  also being among the  $k$  nearest neighbors of  $x_j$ ), the resulting networks are in general directed graphs, and the local density of unidirectional edges (as opposed to bidirectional ones) is related to the gradient of the invariant density.

### 3.2.3. Adaptive neighbor networks

In order to circumvent the directedness of *k*-nearest neighbor networks, Xu *et al.* [40, 74] proposed an algorithm for balancing the neighborhood relationships in such a way that they become symmetric again. The resulting networks embedded in phase space, sometimes also referred to as *adaptive nearest neighbor networks* [38], are conceptually more similar to classical ( $\varepsilon$ )-RNs, but still exhibit somewhat different topological

---

<sup>1</sup>Here,  $\varepsilon$ -balls refers to general (hyper-)volumes according to the specific norm chosen for measuring distances in phase space, e.g., hypercubes of edge length  $2\varepsilon$  in case of the maximum norm, or hyperballs of radius  $\varepsilon$  for the Euclidean norm.

characteristics. In particular, this approach helps to understand the superfamily phenomena of time series, which concern the relative prevalence of motifs of the resulting networks. In particular, the motif distribution of adaptive nearest neighbor networks has been empirically shown to allow a discrimination between different types of dynamics in terms of a different motif ranking [40]. Consequently, this approach has been mainly used for such discriminatory tasks, including applications to turbulence phenomena, instrumental music [38], fractional Brownian motions and multifractal random walks [117].

While these superfamily phenomena have been found in time series from various origins, no concrete theories have been proposed in the literature. Khol *et al* provides a heuristic explanation of superfamily phenomena by examining the dependence of attractor dimension on motif prevalence [81]. Since the reconstructed networks inherently capture the proximity of states, motifs represent specific arrangements of states in space, some of which are more or less likely to occur as dimension changes. Therefore, they found that the relative prevalence of motifs are strongly dependent on the local dimension of the space from which the state vectors are taken. Further evidence is given by identifying comparable superfamily phenomena in networks constructed from states randomly distributed in spaces of varying dimensions [81].

### 3.2.4. Algorithm comparisons and variants

For a detailed discussion of the differences between  $\varepsilon$ -RNs,  $k$ -nearest neighbor and adaptive nearest neighbor networks, we refer to [38]. While these three classes of time series networks exhibit very strong conceptual similarities (the same applies to correlation networks [45] if interpreting the correlation coefficient between two sufficiently high-dimensional state vectors as a generalized distance), the approach proposed by Li *et al.* [118–121] can be understood as being derived from the RN idea. Here, for a set of  $m$ -dimensional embedding vectors, all mutual Euclidean distances are computed. Based on the maximum point-wise Euclidean distance  $d_{max}(m)$ , the threshold distance of an RN is taken as  $\varepsilon(m) = d_{max}(m)/(N-1)$ . This procedure is repeated for different  $m$ , and the critical value of the embedding dimension for which the resulting network gets completely disconnected is interpreted as a complexity index [120]. However, it has not yet been demonstrated that this algorithmic approach has any conceptual benefits in comparison with the classical RN transitivity obtained for a fixed embedding dimension.

Another conceptual approach loosely related to RNs provides the foundation of the frequency-degree mapping algorithm introduced by Li *et al.* [122]. Here, the resulting time series networks contain two types of edges: (i) temporal edges connecting subsequent points in time, and (ii) proximity edges containing observations of similar values, where similarity is defined by an initial grouping of the data into a discrete set of classes, and observed values being connected if and only if they belong to the same class. Here the definition of a class is equivalent to a recurrence interval that is defined by amplitude quantization, for instance, the recurrence interval length  $I = H/Q$  where  $Q$  is the quantization level and  $H$  is the amplitude range of time series. Notably, the latter approach combines the classical recurrence idea and basic concepts of symbolic dynamics [123]. In this spirit, the resulting network's adjacency matrix is given as the recurrence matrix associated with a symbolic recurrence plot [113, 124, 125] plus a “stripe” around the matrix' main diagonal. The frequency-degree mapping algorithm has been successfully applied to characterizing signatures of various types of ventricular arrhythmias in human heart beat time series [122], stock markets [120], and air quality indices [121].

In order to highlight the recurrence domains in the networks, fuzzy recurrence network approach has been proposed in [126], which shares much similarities with fuzzy recurrence plots [112]. Furthermore, a grammatical rewriting algorithm over the recurrence matrix has been proposed to search for recurrence domains in [113], which presents a symbolic description of recurrence properties of time series. It is interesting to see that this algorithm yields an optimal symbolic recurrence representation revealing functional components of brain signals [113]. Note that the computation of recurrence matrix is the first step of this grammatical algorithm.

The computation time of a RN is proportional to  $N^2$  where  $N$  is the number of time points, which calls for more efficient algorithm in constructing RNs for long time series. In the case of long time series, on the other hand, we are more interested in the evolution behavior of RN over time. To this end, sliding window techniques are often suggested but we need to check the dependence on the choice of window sizes [127, 128]. Another idea is to perform coarse graining of the original RNs [114], which stems from original idea of meta

recurrence plots [129]. In [130], the authors proposed to first divide the original long time series into short segments and RNs are reconstructed for each piece. The next step is to build joint recurrence networks for each pair of windowed segments. Then, the global overview of the long term dynamics is characterized by the variations of the network properties that are computed for meta-time series.

### 3.3. Complex network characteristics of RN

Based on the re-interpretation of the recurrence matrix  $\mathbf{R}(\varepsilon)$  as the adjacency matrix of an adjoint RN, we can utilize the large toolbox of complex network measures [10, 29–31] for characterizing the structural organization of a dynamical system in its phase space. Notably, this viewpoint is complementary to other concepts of nonlinear time series analysis making use of RPs. For example, RQA characterizes the statistical properties of line structures in the RP, which implies explicit consideration of dynamical aspects (i.e., sequences of state vectors) [87]. In turn, RNs do not make use of time information, since network properties are generally invariant under vertex relabelling (i.e., permutations of the order of observations) [43]. In this spirit, RN analysis provides geometric instead of dynamical characteristics. This idea of a geometric analysis is similar to some classical concepts of fractal dimensions (e.g., box-counting or correlation dimensions), where certain scaling laws in dependence on the spatial scale of resolution (corresponding here to  $\varepsilon$ ) are exploited. In turn, RN analysis can be performed (as RQA) using only a single fixed scale ( $\varepsilon$ ) instead of explicitly studying scaling properties over a range of threshold values.

The distinction between dynamical and geometric information implies that in case of RN analysis, the typical requirement of a reasonable (typically uniform) temporal sampling of the considered trajectory is replaced by the demand for a suitable spatial coverage of the system’s attractor in phase space. Specifically, under certain conditions the latter could also be obtained by considering an ensemble of short trajectories instead of a single long one. If the trajectory under study is relatively densely sampled, trivial serial correlations can lead to a selection bias in the set of sampled state vectors; the latter could be avoided by reasonable downsampling. In the same context, the possibility of utilizing Theiler windows for removing edges representing short-term auto-correlations (e.g., recurrence points close to the main diagonal in the RP) should be mentioned as another strategy based on a somewhat different rationale [43]. However, from a conceptual perspective, downsampling can provide an unbiased sampling of the attractor as long as the fixed sampling time does not correspond to any integer multiple of some of the system’s natural frequencies. As an alternative, bootstrapping from the set of available state vectors provides another feasible option, which should be preferred if a sufficiently long time series is available. In general, numerical experiments and different applications suggest that stable estimates of RN characteristics can often already be obtained using a sample size of  $N \sim \mathcal{O}(10^2 \dots 10^3)$  data points [127, 128].

In Sec. 2.2, we have provided a general review on various network measures characterizing the structural properties of a complex network as denoted by the adjacency matrix  $A$ . In this section, we further discuss the physical interpretations of these measures in terms of phase space properties as captured by RN representations. In what follows, we will denote all properties computed from a RN consisting of a finite number  $N$  of state vectors as  $\hat{f}$ , pointing to the fact that they are estimated from a given sample of state vectors but shall characterize the entire trajectory of the system under study. Namely, we have specific finite estimations for Eqs. (17)-(11). Furthermore, we will discuss a corresponding continuous framework generalizing all network characteristics described below in Section 3.4. In order to focus the following discussion, we review only the possibly most relevant characteristics associated with RNs. More details including further measures can be found in [43, 131].

When considering quantitative characteristics of complex networks, different classifications of measures are possible. First, we may distinguish measures based on the concept of graph neighborhoods from those making use of shortest path-based characteristics. (This is not an exhaustive classification, since it potentially neglects other important network measures, e.g., such based on diffusion processes or random walks on the network.) Second, network measures can be classified into such making use of local, meso-scale and global information. This scheme is widely equivalent to the first one in that local information refers to properties determined by the graph neighborhood of a given vertex, whereas global information takes contributions due to all vertices of the network into account, which is common for shortest path-based measures. Finally, we can differentiate between measures quantifying properties of single vertices, pairs of vertices, and

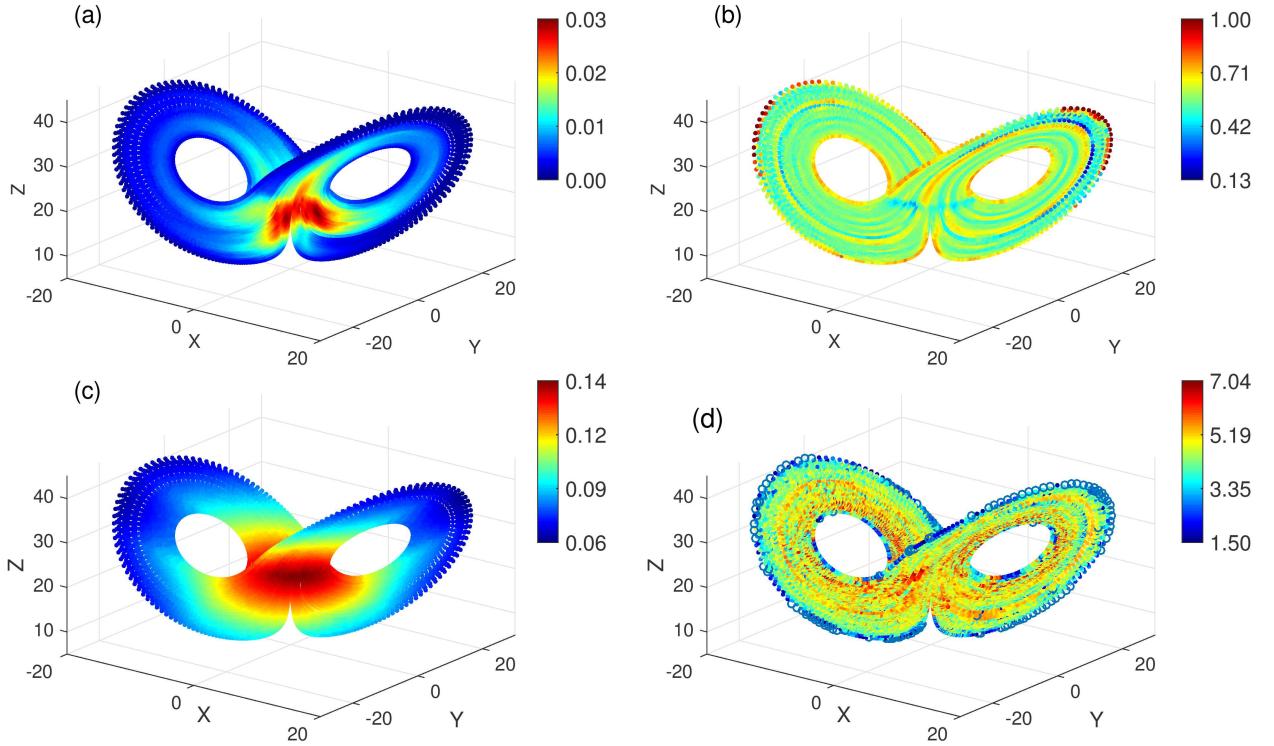


Figure 3: Spatial distributions of vertex characteristics of the  $\varepsilon$ -RN obtained for the Lorenz system (Eq. 15) at the canonical parameters (RPs computed using the maximum norm with  $\hat{\rho} = 0.01$ ,  $N = 20,000$  and sampling time  $\Delta t = 0.1$ ): (a)  $\hat{k}_v(\varepsilon)$ , (b)  $\hat{C}_v(\varepsilon)$ , (c)  $\hat{c}_v(\varepsilon)$ , (d)  $\hat{b}_v(\varepsilon)$ . Reproduced from [43].

the network as a whole. In this chapter, we will utilize the latter way of classification, since it appears most instructive from the applied point of view (i.e., we are commonly interested in either the local or the global geometry of an attractor).

### 3.3.1. Vertex characteristics

The network *degree* or *degree centrality* of a single vertex  $x_v$  (Eq. (17)) has

$$\hat{k}_v(\varepsilon) = \sum_{i=1}^N A_{iv}(\varepsilon). \quad (17)$$

From the perspective of recurrences, it is reasonable to replace the degree by a normalized characteristic, the *degree density*

$$\hat{\rho}_v(\varepsilon) = \frac{\hat{k}_v(\varepsilon)}{N-1} = \frac{1}{N-1} \sum_{i=1}^N A_{iv}(\varepsilon), \quad (18)$$

which corresponds to the definition of the local recurrence rate of the state  $x_v$ .  $\hat{\rho}_v(\varepsilon)$  quantifies the density of states in the  $\varepsilon$ -ball around  $x_v$ , i.e., the probability that a randomly chosen member of the available sample of state vectors is  $\varepsilon$ -close to  $x_v$ . An illustration of this fact for the Lorenz system is presented in Fig. 3(a); here, phase space regions with a high density of points (i.e., a high residence probability of the sampled trajectory) are characterized by a high degree density.

The *local clustering coefficient* (Eq. (2)) reads

$$\hat{C}_v(\varepsilon) = \frac{1}{\hat{k}_v(\varepsilon)(\hat{k}_v(\varepsilon) - 1)} \sum_{i,j=1}^N A_{vi}(\varepsilon) A_{ij}(\varepsilon) A_{jv}(\varepsilon), \quad (19)$$

which measures the fraction of pairs of vertices in the  $\varepsilon$ -ball around  $x_v$  that are mutually  $\varepsilon$ -close. For vertices with  $\hat{k}_v(\varepsilon) < 2$ , we define  $\hat{C}_v(\varepsilon) = 0$ . It has been shown that the local clustering coefficient in a RN is associated with the geometric alignment of state vectors. Specifically, close to dynamically invariant objects such as unstable periodic orbits (UPOs) of low period, the dynamics of the system is effectively lower-dimensional, which results in a locally enhanced fraction of closed paths of length 3 (“triangles”) and, thus, a higher local clustering coefficient. The latter behavior is exemplified in Fig. 3(b) for the Lorenz system, where we recognize certain bands with higher values of  $\hat{C}_v$  corresponding to the positions of known low-periodic UPOs [43].

The *closeness* or *closeness centrality* (Eq. (3)) reads

$$\hat{c}_v(\varepsilon) = \left( \frac{1}{N-1} \sum_{i=1}^N \hat{d}_{vi}(\varepsilon) \right)^{-1}, \quad (20)$$

and the *local efficiency* (Eq. (4)) is

$$\hat{e}_v(\varepsilon) = \frac{1}{N-1} \sum_{i=1}^N \hat{d}_{vi}(\varepsilon)^{-1}. \quad (21)$$

Here again, we set  $\hat{d}_{ij}$  to the highest possible value of  $N - 1$  for pairs of vertices that cannot be mutually reached. Both  $\hat{c}_v(\varepsilon)$  and  $\hat{e}_v(\varepsilon)$  exhibit the highest values for vertices which are situated in the center of the RN (see Fig. 3(c) for an illustration for the Lorenz system).

The *betweenness* or *betweenness centrality* (Eq. (5)) reads

$$\hat{b}_v(\varepsilon) = \sum_{i,j=1; i,j \neq v}^M \frac{\hat{\sigma}_{ij}(v)}{\hat{\sigma}_{ij}}. \quad (22)$$

In the RN context, it can be interpreted as indicating the local degree of fragmentation of the underlying attractor [43]. To see this, consider two densely populated regions of phase space that are separated by a poorly populated one. Vertices in the latter will “bundle” the shortest paths between vertices in the former ones, thus forming geometric bottlenecks in the RN. In this spirit, we may understand the spatial distribution of betweenness centrality for the Lorenz system (Fig. 3(d)) which includes certain bands with higher and lower residence probability reflected in lower and higher betweenness values.

### 3.3.2. Edge characteristics

The *matching index* (Eq. (6)) quantifying the overlap of the network neighborhoods of two vertices  $v$  and  $w$  reads:

$$\hat{m}_{vw}(\varepsilon) = \frac{\sum_{i=1}^N A_{vi}(\varepsilon) A_{wi}(\varepsilon)}{\hat{k}_v(\varepsilon) + \hat{k}_w(\varepsilon) - \sum_{i=1}^N A_{vi}(\varepsilon) A_{wi}(\varepsilon)}. \quad (23)$$

From the above definition, it follows that  $\hat{m}_{vw}(\varepsilon) = 0$  if  $\|x_v - x_w\| \geq 2\varepsilon$ . In turn, there can be mutually unconnected vertices  $v$  and  $w$  ( $A_{vw} = 0$ ) with  $\varepsilon \leq \|x_v - x_w\| < 2\varepsilon$  that have some common neighbors and, thus, non-zero matching index. In the context of recurrences in phase space,  $\hat{m}_{vw}(\varepsilon) = 1$  implies that the states  $x_v$  and  $x_w$  are twins, i.e., share the same neighborhood in phase space. In this spirit, we interpret  $\hat{m}_{vw}(\varepsilon)$  as the degree of twinness of two state vectors. Note that twins have important applications in creating surrogates for testing for the presence of complex synchronization [132, 133].

The *edge betweenness* (Eq. (7)) reads

$$\hat{b}_{vw}(\varepsilon) = \sum_{i,j=1; i,j \neq v,w}^M \frac{\hat{\sigma}_{ij}(v,w)}{\hat{\sigma}_{ij}}, \quad (24)$$

As the (vertex-based) betweenness centrality, in a RN edge betweenness characterizes the local fragmentation of the studied dynamical system in its phase space.

For the specific case of  $\varepsilon$ -recurrence networks, we emphasize that there is no simple correspondence between matching index and edge betweenness, since both quantify distinctively different aspects of phase space geometry. Specifically, there are more pairs of vertices with non-zero matching index than edges, even though there are also pairs of vertices with  $\hat{b}_{vw}(\varepsilon) > 0$  but  $\hat{m}_{vw}(\varepsilon) = 0$  (i.e., there is an edge between  $v$  and  $w$ , but both have no common neighbors). However, for those pairs of vertices for which both characteristics are non-zero, we find a clear anti-correlation. One interpretation of this finding is that a large matching index typically corresponds to very close vertices in phase space; such pairs of vertices can in turn be easily exchanged as members of shortest paths on the network, which implies lower edge betweenness values. A similar argument may explain the coincidence of high edge betweenness and low non-zero matching index values.

### 3.3.3. Global network characteristics

The *edge density* (Eq. (8)) reads

$$\hat{\rho}(\varepsilon) = \frac{1}{N} \sum_{v=1}^N \hat{\rho}_v(\varepsilon) = \frac{1}{N(N-1)} \sum_{i,j=1}^N A_{ij}(\varepsilon) \quad (25)$$

Notably, for a RN the edge density equals the recurrence rate  $RR(\varepsilon)$  of the underlying RP. Strictly speaking, this is only true if the recurrence rate is defined such as the main diagonal in the RP is excluded in the same way as potential self-loops from the RN's adjacency matrix. It is trivial to show that  $\hat{\rho}(\varepsilon)$  is a monotonically increasing function of the recurrence threshold  $\varepsilon$ : the larger the threshold, the more neighbors can be found, and the higher the edge density.

The arithmetic mean of the local clustering coefficients  $\hat{C}_v(\varepsilon)$  of all vertices (Eq. (9)) reads

$$\hat{C}(\varepsilon) = \frac{1}{N} \sum_{v=1}^N \hat{C}_v(\varepsilon) = \frac{1}{N} \sum_{v=1}^N \frac{\sum_{i,j=1}^N A_{vi}(\varepsilon) A_{ij}(\varepsilon) A_{jv}(\varepsilon)}{\hat{k}_v(\varepsilon)(\hat{k}_v(\varepsilon) - 1)}. \quad (26)$$

Given our interpretation of the local clustering coefficient in a RN,  $\hat{C}(\varepsilon)$  can be interpreted as a proxy for the average local dimensionality of the dynamical system in phase space.

The *network transitivity* (Eq. (10)) is defined as

$$\hat{T}(\varepsilon) = \frac{\sum_{v,i,j=1}^N A_{vi}(\varepsilon) A_{ij}(\varepsilon) A_{jv}(\varepsilon)}{\sum_{v,i,j=1}^N A_{vi}(\varepsilon) A_{jv}(\varepsilon)}. \quad (27)$$

When interpreting  $\hat{C}(\varepsilon)$  as a proxy for the average local dimensionality,  $\hat{T}(\varepsilon)$  characterizes the effective global dimensionality of the system.

The *average path length* (Eq. (11)) is

$$\hat{L}(\varepsilon) = \frac{1}{N(N-1)} \sum_{i,j=1}^N \hat{d}_{ij}(\varepsilon) = \frac{1}{N} \sum_{v=1}^N \hat{c}_v(\varepsilon)^{-1} \quad (28)$$

and the *global efficiency* (Eq. (12))

$$\hat{E}(\varepsilon) = \left( \frac{1}{N(N-1)} \sum_{i,j=1}^N \hat{d}_{ij}(\varepsilon)^{-1} \right)^{-1} = \left( \frac{1}{N} \sum_{v=1}^N \hat{e}_v(\varepsilon) \right)^{-1} \quad (29)$$

We can easily convince ourselves that the average path length must exhibit an inverse relationship with the recurrence threshold, since it approximates (constant) distances in phase space in units of  $\varepsilon$  [43].

### 3.4. Analytical theory of RN

As we will demonstrate in the following, the properties of RNs can be described analytically supporting their better understanding and, hence, applicability. For this purpose, we can exploit the formal equivalence of RNs and random geometric graphs (RGGs), a well-studied concept in graph theory and computational geometry. In this section, we motivate this equivalence and demonstrate how the variety of RN characteristics can be reformulated in the continuum limit  $N \rightarrow \infty$  for any finite  $\varepsilon$  [131]. This framework allows gaining deep insights into the geometric organization of chaotic attractors by exploring the multitude of characteristics provided by complex network theory. Moreover, this analytical considerations will be extended to inter-system recurrence networks in Section 3.9.

#### 3.4.1. Preliminaries: random geometric graphs

Random geometric graphs [134] are based on a (finite) set of vertices randomly positioned in a  $d$ -dimensional ( $d \in \mathbb{N}^+$ ) metric space according to some probability density function  $p(x)$ . In general, the connectivity among this set of vertices is taken to be distance-dependent, i.e., for two vertices  $i$  and  $j$ , the probability of being connected in the RGG has the form  $P(A_{ij} = 1) = f(\|x_i - x_j\|)$  with some predefined function  $f$ , which is monotonically decreasing. As a consequence, spatially close vertices are more likely to be connected than distant ones. A particularly well studied special case is  $f(\delta) = \Theta(\varepsilon - \delta)$  ( $\delta$  denoting here the distance between any two points in the considered metric space), often referred to as RGG (in the strict sense). Notably, the latter definition has fundamental real-world importance (e.g., in terms of ad-hoc communication networks or, more general, contact networks) and matches that of the adjacency matrix of a RN (Eq. 16) if we identify  $p(x)$  with the invariant density of the dynamically invariant object under study (e.g., some attractor in case of a dissipative system), and take the space in which the RGG is embedded as that spanned by the respective dynamical variables of the system. In this respect, for all following considerations, it is sufficient restricting attention to the support of  $p(x)$  (respectively its closure), which is described by some manifold  $S = \text{supp}(p)$  embedded in the considered metric space (e.g., the attractor manifold).

From a practical perspective, the spatial coverage of  $p(x)$  by the RGG's vertices can be strongly affected by the sampling, leading to a spatial clustering of vertices if the sampling frequency is close to an integer multiple of the chaotic attractor's characteristic frequency. In such a situation, it is advantageous to follow alternative sampling strategies for  $p(x)$ . Note that for ergodic systems, sampling from one long trajectory, ensembles of short independent realizations of the same system, or directly from the invariant density should lead to networks with the same properties at sufficiently large  $N$ . As already mentioned above, generating the RGG/RN representation based on bootstrapping from the ensemble of available state vectors is superior to a regular sampling of a given trajectory.

As outlined above, the importance of RGGs for the considerations on RNs is that some of their properties (like the degree distribution [135] or transitivity [136]) have been intensively studied in the past for the generic case of a hard distance threshold in  $f$  and arbitrary probability density functions  $p(x)$  for metric spaces of various integer dimensions. For example, Hermann *et al.* [135] give a closed-form expression of the degree distribution for arbitrary  $p(x)$ , whereas Dall and Christensen [136] provide a deep discussion of the transitivity properties of RGGs. Notably, the latter aspect has become particularly relevant in the interpretation of RN properties as well as their multivariate generalizations, as will be further discussed in Section 3.9.

#### 3.4.2. Analytical description of $\varepsilon$ -recurrence networks

By making use of the fact that RNs are a specific type of RGGs, all relevant graph-theoretical measures for recurrence networks can be seen as discrete approximations of more general and continuous geometrical properties of a dynamical system's underlying attractor characterized by a set  $S$  together with an associated invariant density  $p(x)$ ,  $x \in S$ . This point of view allows obtaining deeper insights into the geometrical meaning of the network quantifiers introduced in Section 3.3 and enables us to establish surprising connections to other fields, *e.g.*, the close relationship of transitivity measures like the local clustering coefficient and transitivity to the local and global fractal dimension of the dynamical system's attractor, respectively [27]

(see Section 3.5.5). In the following, we review a corresponding analytical framework for general spatially embedded networks which is specifically tailored for defining continuous variants of the common discrete complex network characteristics [131].

*General setting.* Let  $S$  be a compact and smooth manifold with a non-vanishing continuous probability density function  $p : S \rightarrow (0, \infty)$  with  $\int_S dx p(x) = 1$ . For the purpose of the present work, we identify  $S$  with the set of points defining the attractor of a (dissipative) dynamical system. In case of chaotic attractors in time-continuous systems, we obtain a closure of the open attractive set by considering its union with the set of (infinitely many) unstable periodic orbits embedded in the attractor.

Continuous analogs of the discrete complex network characteristics introduced in Section 3.3 should be approximated by taking the limit  $N \rightarrow \infty$  and  $\varepsilon \rightarrow 0$  (note that the latter limit may not be assessible in the case of fractal sets  $S$ , which we will not further consider in the following). Here, “continuous” refers to a network with uncountably many vertices and edges, which is determined by the *adjacency function*

$$A(x, y) = \Theta(\varepsilon - \|x - y\|) - \delta(x - y) \quad (30)$$

for all  $x, y \in S$ , which is a continuous analog of the adjacency matrix. In the latter expression,  $\delta(x - y) = 1$  if  $x = y$ , and 0 otherwise.

*Shortest paths and geodesics.* A large variety of complex network characteristics introduced in Section 3.3 relies on the concept of shortest paths. Examples include closeness and betweenness centrality, local and global efficiency, and average path length. In the continuum limit, we consider a path in  $S$  as a closed curve described by a properly parametrized function  $f : [0, 1] \rightarrow S$ , and define the associated path length

$$l(f) = \sup_{n > 0; \{t_i\}_{i=1}^n} \left\{ \sum_{i=1}^n d(f(t_{i-1}), f(t_i)) \middle| 0 = t_0 \leq t_1 \leq \dots \leq t_n = 1 \right\} \in [0, \infty] \quad (31)$$

where  $d(\cdot)$  denotes some metric used for defining distances on  $S$ . The *geodesic distance* between two points  $x, y \in S$ , which serves as the analog of the shortest path length on a network, is then defined as (cf. Fig. ??).

$$g(x, y) = \inf_f \{l(f) \mid f : [0, 1] \rightarrow S, f(0) = x, f(1) = y\}. \quad (32)$$

A path of length  $g(x, y)$  is called a *global geodesic* on  $S$ . Depending on the specific geometry of the considered set  $S$ , the multiplicity of global geodesics connecting  $x$  and  $y$  may differ, including no, one, or even infinitely many distinct global geodesics.

Regarding a continuum limit for RNs, we note that shortest paths in such networks approximate global geodesics on the underlying invariant set  $S$  in the limit of  $\varepsilon \rightarrow 0$  and  $N \rightarrow \infty$ . Specifically, in the latter limit the shortest path length  $l_{ij}(\varepsilon)$  between two points  $x(t_i), x(t_j) \in S$  behaves as

$$\varepsilon l_{ij}(\varepsilon) \rightarrow g(x(t_i), x(t_j)) \quad (33)$$

independently of the chosen metric [131].

For defining betweenness centrality, we do not only require information on the lengths of global geodesics, but also their total multiplicity  $\sigma(y, z; \varepsilon)$  as well as their multiplicity conditional on a third point  $x \in S$  being part of the curve, denoted as  $\sigma(y, z|x; \varepsilon)$  in the following. The definition of the latter quantity is, however, not unique for a given finite  $\varepsilon$ . Two possible, yet generally not equivalent expressions read [131]

$$\sigma_1(y, z|x; \varepsilon) = \sum_{k=1}^{\sigma(y, z; \varepsilon)} \int_0^1 dt \delta(f_k(t) - x) \quad (34)$$

$$\sigma_2(y, z|x; \varepsilon) = \sum_{k=1}^{\sigma(y, z; \varepsilon)} \int_0^1 dt \Theta(\varepsilon - \|f_k(t) - x\|), \quad (35)$$

where  $f_k(t)$  denotes the family of global geodesics between  $y$  and  $z$ . Note that this family can have uncountably many members (to see this, consider, for example, the set of geodesics between the two poles on a sphere). In this case, the sum in Eqs. (34) and (35) should be replaced by an integral. Furthermore, we emphasize that the  $\varepsilon$ -dependence in the multiplicities of shortest paths is implicit rather than explicit, since the chosen discretization level  $\varepsilon$  can affect the effective “shape” of  $S$  and, hence, the positions of possible edges in the considered space.

*Local (vertex-based) measures.* The *continuous  $\varepsilon$ -degree density*

$$\rho(x; \varepsilon) = \int_{B_\varepsilon(x)} d\mu(y) \quad (36)$$

gives the probability that a point  $y \in S$  randomly drawn according to  $p$  falls into an  $\varepsilon$ -neighborhood  $B_\varepsilon(x) = \{y \in S \mid \|x - y\| < \varepsilon\}$  around  $x$ . Its discrete estimator is given by the classical degree density  $\hat{\rho}_v(\varepsilon)$  (Eq. 18).

In order to quantify the density of closed paths of length 3 in the network, we can consider the *continuous local  $\varepsilon$ -clustering coefficient*

$$\mathcal{C}(x; \varepsilon) = \frac{\int \int_{B_\varepsilon(x)} d\mu(y) d\mu(z) \Theta(\varepsilon - \|y - z\|)}{\rho(x; \varepsilon)^2}. \quad (37)$$

This measure characterizes the probability that two points  $y$  and  $z$  randomly drawn according to  $p$  from the  $\varepsilon$ -neighborhood of  $x \in S$  are mutually closer than  $\varepsilon$ . Its discrete approximation is provided by the classical local clustering coefficient  $\hat{\mathcal{C}}_v(\varepsilon)$  (Eq. 19).

Let  $y \in S$  be drawn randomly according to  $p$ . For a fixed  $x \in S$ , the *continuous  $\varepsilon$ -closeness centrality*

$$c(x; \varepsilon) = \left( \int_S d\mu(y) \frac{g(x, y)}{\varepsilon} \right)^{-1} \quad (38)$$

and the *continuous local  $\varepsilon$ -efficiency*

$$e(x; \varepsilon) = \int_S d\mu(y) \left( \frac{g(x, y)}{\varepsilon} \right)^{-1} \quad (39)$$

give the inverse expected geodesic distance and the expected inverse geodesic distance of  $y$  to  $x$ , respectively. Hence, both measures quantify the geometric closeness of  $x$  to any other point in  $S$  according to the probability density function  $p$ . By making use of RNs, they can be approximated by the classical closeness centrality  $\hat{c}_v(\varepsilon)$  (Eq. 20) and local efficiency  $\hat{e}_v(\varepsilon)$  (Eq. 21).

Finally, the probability that a point  $x$  lies on a randomly chosen global geodesic connecting two points  $y, z \in S$  according to  $p$  is measured by the *continuous  $\varepsilon$ -betweenness centrality*

$$b(x; \varepsilon) = \int \int_S d\mu(y) d\mu(z) \frac{\sigma(y, z|x; \varepsilon)}{\sigma(y, z; \varepsilon)}. \quad (40)$$

Its discrete estimator is given by the standard RN betweenness centrality  $\hat{b}_v(\varepsilon)$  (Eq. 22) with the different possible expressions for  $\sigma(y, z|x; \varepsilon)$  (Eqs. 34,35) [131].

*Pairwise vertex and edge measures.* The *continuous  $\varepsilon$ -matching index*

$$m(x, y; \varepsilon) = \frac{\int_{B_\varepsilon(x) \cap B_\varepsilon(y)} d\mu(z)}{\int_{B_\varepsilon(x) \cup B_\varepsilon(y)} d\mu(z)} \quad (41)$$

quantifies the mutual overlap between the neighborhoods of two vertices  $x, y \in S$ . In other words,  $m(x, y; \varepsilon)$  is the probability that a point  $z \in S$  randomly chosen from  $B_\varepsilon(x)$  according to  $p$  is also contained in  $B_\varepsilon(y)$ .

and vice versa. For  $x \rightarrow y$ , we have  $B_\varepsilon(x) \rightarrow B_\varepsilon(y)$  and, consequently,  $m(x, y; \varepsilon) \rightarrow 1$ , whereas  $m(x, y; \varepsilon) = 0$  if  $\|x - y\| > 2\varepsilon$ . As in the case of the other measures described above,  $m(x, y; \varepsilon)$  can be approximated by the discrete RN matching index  $\hat{m}_{vw}(\varepsilon)$  (Eq. 23).

Note that  $m(x, y; \varepsilon)$  does not require mutual closeness between  $x$  and  $y$  (i.e.,  $\|x - y\| \in (\varepsilon, 2\varepsilon)$  is possible). In contrast, the *continuous  $\varepsilon$ -edge betweenness*

$$b(x, y; \varepsilon) = \int \int_S d\mu(z) d\mu(z') \frac{\sigma(z, z'|x, y; \varepsilon)}{\sigma(z, z'; \varepsilon)} \quad (42)$$

(with  $\sigma(z, z'|x, y; \varepsilon)$  denoting the number of global geodesics between  $z$  and  $z'$  containing both  $x$  and  $y$  under the condition  $\|x - y\| \leq \varepsilon$ , and  $\sigma(z, z'; \varepsilon)$  the total number of global geodesics between  $z$  and  $z'$ ) is a measure whose discrete estimator  $\hat{b}_{vw}(\varepsilon)$  (Eq. 24) is related to the presence of an edge between  $x(t_v)$  and  $x(t_w)$ , i.e.,  $\|x(t_v) - x(t_w)\| < \varepsilon$ . However, although this property has been originally introduced as an explicit edge property, it can be understood in a more general way as a two-vertex property such that  $b(x, y; \varepsilon)$  measures the probability that two specific (not necessarily  $\varepsilon$ -close) points  $x$  and  $y$  both lie on a  $p$ -randomly drawn global geodesic connecting two points  $z, z' \in S$  and are mutually closer than  $\varepsilon$ . Further generalizations towards  $n$ -point relationships are possible, but not instructive within the scope of this work.

*Global network measures.* The *continuous  $\varepsilon$ -edge density*

$$\rho(\varepsilon) = \int_S d\mu(x) \rho(x; \varepsilon) \quad (43)$$

is the  $p$ -expectation value of the continuous  $\varepsilon$ -degree density and approximated by the discrete edge density  $\hat{\rho}(\varepsilon)$  of a RN (Eq. 25).

In the same spirit, the *continuous global  $\varepsilon$ -clustering coefficient*

$$\mathcal{C}(\varepsilon) = \int_S d\mu(x) \mathcal{C}(x; \varepsilon) \quad (44)$$

is the  $p$ -expectation value of the continuous local  $\varepsilon$ -clustering coefficient. Its associated discrete estimator is the classical global (Watts-Strogatz) clustering coefficient  $\hat{\mathcal{C}}(\varepsilon)$  (Eq. 26). As an alternative measure characterizing geometric transitivity, we can define the *continuous  $\varepsilon$ -transitivity*

$$\mathcal{T}(\varepsilon) = \frac{\int \int \int_S d\mu(x) d\mu(y) d\mu(z) \Theta(\varepsilon - \|x - y\|) \Theta(\varepsilon - \|y - z\|) \Theta(\varepsilon - \|z - x\|)}{\int \int \int_S d\mu(x) d\mu(y) d\mu(z) \Theta(\varepsilon - \|x - y\|) \Theta(\varepsilon - \|z - x\|)}, \quad (45)$$

which gives the probability that among three points  $x, y, z \in S$  randomly drawn according to  $p$ ,  $y$  and  $z$  are mutually closer than  $\varepsilon$  given they are both closer than  $\varepsilon$  to  $x$ . The corresponding discrete estimator is the RN transitivity  $\hat{\mathcal{T}}(\varepsilon)$  (Eq. 27).

As examples of shortest path-based characteristics, we can define the *continuous  $\varepsilon$ -average path length*

$$\mathcal{L}(\varepsilon) = \int \int_S d\mu(x) d\mu(y) \frac{g(x, y)}{\varepsilon} \quad (46)$$

and the *continuous global  $\varepsilon$ -efficiency*

$$\mathcal{E}(\varepsilon) = \left( \int \int_S d\mu(x) d\mu(y) \left( \frac{g(x, y)}{\varepsilon} \right)^{-1} \right)^{-1}, \quad (47)$$

which measure the expected geodesic distance and the inverse of the expected inverse geodesic distance, respectively, both measured in units of  $\varepsilon$  between two points  $x, y \in S$  drawn randomly according to  $p$ . Their discrete estimators are given by the classical RN average path length  $\hat{\mathcal{L}}(\varepsilon)$  (Eq. 28) and global efficiency  $\hat{\mathcal{E}}(\varepsilon)$

(Eq. 29), respectively. Notably, we can reformulate  $\mathcal{L}(\varepsilon)$  as the  $p$ -expectation value of the inverse continuous  $\varepsilon$ -closeness centrality,

$$\mathcal{L}(\varepsilon) = \int_S d\mu(x) c(x; \varepsilon)^{-1}, \quad (48)$$

and  $\mathcal{E}(\varepsilon)$  the inverse  $p$ -expectation value of the continuous local  $\varepsilon$ -efficiency

$$\mathcal{E}(\varepsilon) = \left( \int_S d\mu(x) e(x; \varepsilon) \right)^{-1}. \quad (49)$$

*Further characteristics.* The selection of measures discussed above is far from being complete. Continuous versions of further complex network characteristics, such as assortativity, network diameter and radius, and network motifs are discussed in [131], where also some outlook on corresponding generalizations of other measures like eigenvector centrality or random walk betweenness has been given. To this end, we restrict ourselves to the measures discussed above, since they have been most commonly used in recent applications of the RN framework.

### 3.5. General properties of recurrence networks

With the general RN framework (Section 3.3) and the associated analytical treatment of RNs (Section 3.4) in mind, it is possible to study the properties of RNs as well as their multivariate generalizations from a solid theoretical basis. In the following, we will first discuss some general aspects of complex networks often found in real-world systems, such as small-world effects, the emergence of scale-free degree distributions, or assortative mixing (i.e., the tendency of vertices to connect with other vertices that exhibit a similar degree), regarding their presence or absence in RNs. Subsequently, we will turn to the transitivity characteristics of RNs, motivating their particular usefulness for detecting geometric signatures of qualitative changes in the dynamics of a single system.

#### 3.5.1. Degree distributions

A general analytical expression for the degree distribution  $P(k)$  of a RGG and, hence, a RN has been given by Herrmann *et al.* [135]. For this purpose, let us make the following assumptions: (i) The system under study is ergodic. (ii) The sampled trajectory is sufficiently close to its attractor, i.e., we exclude the presence of transient behavior. (iii) The sampling interval is co-prime to any possible periods of the system. If these three conditions are met, the vertices of the RN can be considered as being randomly sampled from the probability density function  $p(x)$  associated with the invariant measure  $\mu$  of the attractor [137].

For a RGG with arbitrary  $p(x)$ , the degree distribution  $P(k)$  can be derived from  $p(x)$  in the limit of large sample size  $N$  as

$$P(k) = \int dx p(x) e^{-\alpha p(x)} (\alpha p(x))^k / k! \quad (50)$$

(representing an  $n$ -dimensional integral in case of an  $n$ -dimensional system) with  $\alpha = \langle k \rangle / \int dx p(x)^2$  [135]. In order to understand this relationship, note that for each  $x$ , the probability that a sampled point falls into the  $\varepsilon$ -ball centered at  $x$  is approximately proportional to  $p(x)$ . Hence, the degree of a node at  $x$  has a binomial distribution. For sufficiently large  $N$ , the latter can be approximated by a Poissonian distribution with parameter  $\alpha p(x)$ , leading to Eq. (50).

For the specific case of one-dimensional maps (i.e., the Logistic map), Eq. (50) can be explicitly evaluated, leading to a general characterization of the conditions under which scale-free distributions can emerge in RNs. When projecting higher-dimensional time-continuous systems to such one-dimensional maps by making use of appropriate (Poincaré) return maps, the corresponding considerations can be generalized to such systems, given the specific Poincaré surface is “representative” for the system’s geometric structure. A detailed discussion has been presented in [138]. To this end, we only recall the main result that when the system’s invariant density  $p(x)$  exhibits a singularity with power-law shape, Eq. (50) implies that the resulting RN’s degree distribution must also display a power-law in the limit  $N \rightarrow \infty$  for sufficiently small  $\varepsilon$ . In turn, if  $\varepsilon$  is chosen too large, the scale-free behavior cannot be detected anymore, since it is masked by too large

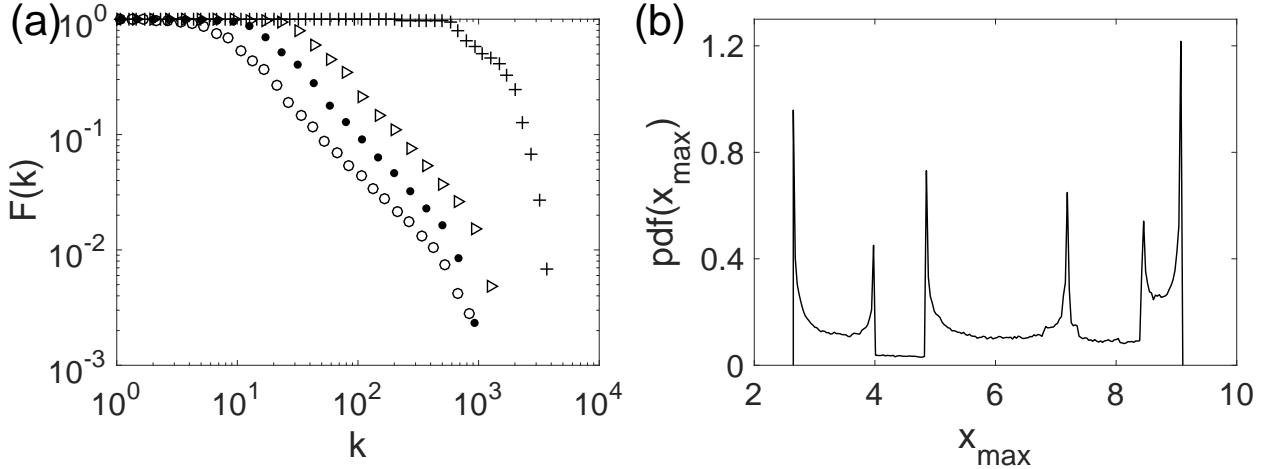


Figure 4: (a) Complementary cumulative distribution function  $F(k) = \sum_{k'=k}^{\infty} p(k')$  for RNs obtained from the  $x$ -component of the first return map of the Rössler system (with  $a = b = 0.2$ ,  $c = 5.7$ ) through the  $y = 0$  plane, using edge densities  $\hat{\rho}_1 = 0.02\%$  ( $\circ$ ),  $\hat{\rho}_2 = 0.03\%$  ( $\bullet$ ),  $\hat{\rho}_3 = 0.05\%$  ( $\triangleright$ ), and  $\hat{\rho}_4 = 3\%$  ( $+$ ). All curves have been obtained as mean values taken from 5 independent realizations of the system with length  $N = 2 \times 10^5$  and using the Euclidean norm. For  $\hat{\rho}_1$  to  $\hat{\rho}_3$ , we find power-law behavior with a characteristic exponent of  $\gamma = 2.16 \pm 0.03$ , whereas no clear scaling region is found in the denser RN with edge density  $\hat{\rho}_4$ . (b) PDF of the  $x$  values, where power-law shaped singularities are observed. Redrawn after [138].

neighborhoods of the points close to the singularity. Figure 4 demonstrates the latter effect for the specific case of the Rössler system (Eq. ??).

Notably, it is not trivial to provide an exhaustive characterization of the conditions under which scale-free distributions can emerge for higher-dimensional systems. As a consequence, generally applicable necessary and sufficient conditions for the presence of power-laws in the degree distributions of RNs have not been established so far. Based on the degree distribution  $P(k)$ , some higher order statistics have been proposed in [139] quantifying the heterogeneity properties of the connectivity.

We note that in general complex systems, the emergence of power-laws is often associated with a hierarchical organization related to certain fractal properties. In contrast, for RNs it has been shown that the presence of power-laws is not directly related to some (global) fractal structure of the system, but rather the local shape of its invariant density. Consequently, although there are examples of dynamical systems where the scaling exponent of the degree distribution coincides well with the associated fractal dimension, there is no such relationship in general. It will be subject of future studies under which conditions regarding the structural organization of the attractor, fractal structure and power-law singularities are sufficiently closely related so that the RN's degree distribution allows quantifying the system's fractal properties.

### 3.5.2. Small-world effect

A first generic property shared by many real-world networks is the so-called small-world effect, first described as the outcome of studies on social interrelationships, predominantly Milgram's famous chain-letter experiment in the 1960s [140]. In the spirit of the latter studies, the term “small-world effect” originally denoted the fact that average shortest path lengths in social networks, but also other real-world networks, are much shorter than we would expect from random connectivity configurations. Given the importance of redundancy in such networks, Watts and Strogatz [63] suggested including the presence of a high clustering coefficient (i.e., higher than in random graphs) as a second criterion for identifying the small-world effect in real-world networks.

From the latter considerations, it is clear that RNs cannot obey small-world effects: although they may exhibit a high degree of transitivity (typically depending on the specific system under study), for any fixed value of  $\varepsilon$  their average path lengths can only take specific values, which become independent of the network size  $N$  in case of sufficiently large samples. On the one hand, for any chosen pair of vertices  $i$  and  $j$  at positions  $x_i$  and  $x_j$ , the shortest path length is bounded from below as  $\hat{l}_{ij} \geq \lceil \|x_i - x_j\|/\varepsilon \rceil$  (respectively,

the geodesic distance on the attractor  $S$  divided by the recurrence threshold  $\varepsilon$ ). Specifically, each shortest path length will converge to a finite value for  $N \rightarrow \infty$ . On the other hand, due to the finite diameter of chaotic attractors, the average path length  $\hat{\mathcal{L}}(\varepsilon)$  cannot exceed a maximum value of  $\lceil \max_{i,j} \{ \|x_i - x_j\| \} / \varepsilon \rceil$  independent of  $N$ . Hence, the average path length is bounded from above by a value independent of  $N$ , which is distinct from the common behavior of small-world networks ( $\hat{\mathcal{L}} \sim \log N$ ) [63]. Moreover, as another immediate consequence of the latter considerations, we observe that  $\hat{\mathcal{L}} \sim \varepsilon^{-1}$  [43]. This implies that by tuning  $\varepsilon$ , it is possible to achieve any desired average shortest path length  $\hat{\mathcal{L}}$ ; this fact notably reduces the explanatory power of this global network characteristic.

add results from Jacob et al. here; check structure of this section...

### 3.5.3. Assortative vs. disassortative mixing

Unlike small-world effects and scale-free degree distributions, there are hardly any available results regarding the mixing properties of RNs. In general, RNs obey a tendency towards showing assortative mixing (i.e., vertices tend to link to other vertices with similar degree), which is reasonable in situations where the invariant density  $p(x)$  is continuous or even differentiable, which is supported by recent numerical results [43, 128].

### 3.5.4. Path-based characteristics

One main field of application of RQA as well as other quantitative approaches to characterizing the distribution of recurrences in phase space (e.g., recurrence time statistics) is identifying and quantifying different degrees of dynamical complexity among realizations of the same system under different conditions (e.g., different values of the control parameter(s)), or even within a single time series given the system is non-stationary. While the line-based characteristics of RQA are founded on heuristic considerations (e.g., the higher the predictability of the observed dynamics, the longer the diagonal line structures off the main diagonal should be), we have argued in Section 3.4 that RNs have an analytical foundation in RGGs. Notably, the corresponding characteristics are based on the same binary structure (the recurrence matrix) as the RQA measures and. Hence, both concepts allow deriving a similar kind of information, with the important difference being that RQA quantifies dynamical properties whereas RNs encode topological/geometric characteristics. However, since both aspects are ultimately linked in the case of chaotic attractors, this general observation suggests that RN analysis is in principle suitable for characterizing dynamical complexity in the same way as other established concepts. Therefore, one natural question arises: How do RN measures perform in this task, and which of the multiple possible network measures are particularly suited for this purpose?

The latter questions have been the main motivation behind much of the early work on RNs focussing on numerical studies of various paradigmatic model systems for low-dimensional chaos [38, 42, 43, 141–144]. The latter studies suggest that for characterizing dynamical complexity, global network characteristics are conceptually easier to use and could provide potentially more stable and distinctive results than certain statistics over local network properties such as the distributions of vertex degrees [145] or local clustering coefficients [144]. Among the set of possible global RN measures, two properties have been found particularly useful: network transitivity  $\hat{T}$  and average path length  $\hat{\mathcal{L}}$ .

Regarding the average path length, the discriminatory skills regarding different degrees of dynamical complexity can be understood by the fact that for time-continuous systems, chaotic systems can display different degrees of spatial filling of the “populated” area in phase space, i.e., a high (fractal) dimension of a chaotic attractor close to the (integer) dimension of the corresponding phase space gives rise to a more homogeneous filling than lower ones, which has a natural geometric consequence for the possible path lengths between pairs of sampled state vectors on the attractor. However, it needs to be noted that quantifying dynamical complexity by means of  $\hat{\mathcal{L}}$  suffers from two important drawbacks:

On the one hand, the measure is not normalized and depends crucially on the choice of  $\varepsilon$ . Hence, working in different methodological settings (e.g., using fixed recurrence thresholds  $\varepsilon$  vs. fixed recurrence rates  $RR = \hat{\rho}$ ) can provide potentially ambiguous results, since numerical values of  $\hat{\mathcal{L}}$  cannot necessarily be directly compared with each other.

On the other hand, even the qualitative behavior of  $\hat{\mathcal{L}}$  in dependence on the system's dynamical complexity depends on whether the system is a discrete map or time-continuous. In the latter case, a periodic orbit would result in a higher average path length than a chaotic one, since a chaotic attractor is a “spatially extended” object in phase space on which there are “shortcuts” between any two state vectors connecting points corresponding to different parts of the trajectory [43]. In turn, for discrete maps, a periodic orbit contains only a finite set of  $p$  mutually different state vectors, so that for sufficiently low  $\varepsilon$  and large  $N$ , the RN is decomposed into  $p$  disjoint, fully connected components. In such a situation with not just single isolated vertices, but a completely decomposed network, a reasonable redefinition of  $\hat{\mathcal{L}}$  would be summing up only over pairs of mutually reachable vertices in Eq. (28). Consequently, we approach the minimum possible value of  $\hat{\mathcal{L}} = 1$  [42], whereas chaotic orbits typically lead to larger  $\hat{\mathcal{L}}$ .

According to the above observations, there is no fully developed theoretical understanding and description of the influence of attractor dimensionality on the resulting average path length beyond the general considerations presented in Section 3.4. Corresponding further investigations might be an interesting subject for future studies.

### 3.5.5. Dimension characteristics by clustering and transitivity

As mentioned in Section 3.5.1, the scaling exponent of a possible power-law degree distribution has no direct relationship to the fractal dimension of the system. In turn, such a relationship naturally exists when studying the corresponding integrated measure (i.e., the edge density  $\hat{\rho}(\varepsilon)$ ) in terms of its scaling properties as the recurrence threshold is systematically varied. The latter approach has been extensively discussed in the literature in connection with the estimation of dynamical invariants from RPs [106, 146] and gives rise to estimates of the correlation dimension  $D_2$ . Notably, one of the classical approaches to estimating  $D_2$  from time series data, the Grassberger-Procaccia algorithm [7, 147], makes use of the correlation sum, which can be easily formulated in terms of the recurrence rate or RN edge density.

The relatively high computational complexity of the latter approaches to estimating the correlation dimension from a RP stems from the fact that a sequence of RPs for different values of  $\varepsilon$  needs to be studied for obtaining a proper scaling relationship. In turn, as shown by us in previous studies [27], network transitivity provides an alternative approach to defining and estimating a different notion of fractal dimension. For this purpose, note that for a classical RGG embedded in some integer-dimensional metric space, the expected network transitivity (which is numerically estimated as the ensemble mean over sufficiently many realizations of the stochastic generation of the RGG) is an analytical function of the dimension  $m$ , which decays (exactly when using the maximum norm, otherwise approximately) exponentially with  $m$  [136]. This analytical relationship can be generalized to attractor manifolds with non-integer fractal dimensions, which can in turn be estimated from the RN transitivity by inverting this function.

*Transitivity dimensions.* For the general case, the latter idea leads to a pair of quantities referred to as upper and lower transitivity dimensions [27],

$$D_{\mathcal{T}}^u = \limsup_{\varepsilon} \frac{\log(\mathcal{T}(\varepsilon))}{\log(3/4)}, \quad (51)$$

$$D_{\mathcal{T}}^l = \liminf_{\varepsilon} \frac{\log(\mathcal{T}(\varepsilon))}{\log(3/4)}, \quad (52)$$

where the two definitions originate from the fact that certain systems (in particular, chaotic maps whose attractors form Cantor sets in at least one direction in phase space [27]) can exhibit an oscillatory behavior between some upper and lower accumulation point of  $\mathcal{T}(\varepsilon)$  as the recurrence threshold  $\varepsilon$  is varied. For systems without such fragmented structure, upper and lower transitivity dimension seem to coincide, which allows estimating them from the sample RN transitivity with reasonable accuracy using only a single network instance with one suitably chosen value of  $\varepsilon$ . A detailed analytical investigation of the qualitatively different behavior of the RN transitivity for chaotic attractors with continuous and fragmented invariant densities in dependence on  $\varepsilon$  will be subject of future work. Note that in the above definition, we do not explicitly consider a scaling behavior for  $\varepsilon \rightarrow 0$ , since the definition does not explicitly contain  $\varepsilon$  (as it is the case for

other classical notions of fractal dimensions), but makes use of normalized characteristics with a probabilistic interpretation (cf. Section 3.4). In this spirit, the fraction on the right-hand side of the former equations is a well-defined object for each value of  $\varepsilon$  (i.e., the specific scale under which the system is viewed) individually.

More numerical results have been presented in [131]. We clearly recognize that  $\hat{D}_T(\varepsilon)$  assumes approximately stable (i.e.,  $N$ - and  $\varepsilon$ -independent) values if the recurrence threshold is chosen sufficiently large. In general, there exist two limits that need to be taken into account: For too large recurrence rates, the RN characteristics lose their discriminatory skills, since too many edges are present masking subtle small-scale properties of the attractor [131, 142]. In turn, if  $\varepsilon$  is too low (e.g., if  $\hat{\rho}$  is below the RN's percolation threshold) [131], the network decomposes into mutually disjoint components, and the resulting network characteristics can become ambiguous.

Notably, the analytical relationship (Eqs. 51,52) between the effective (geometric) dimension of chaotic attractors and RN transitivity provides the theoretical justification and foundation for applying  $\hat{T}$  as a characteristic discriminating between high and low dynamical complexity of chaotic attractors. Unlike for  $\hat{\mathcal{L}}$ , the transitivity shows qualitatively the same behavior for discrete and time-continuous systems and is normalized, so that its values can be directly used as a quantitative measure of dynamical complexity associated with the effective geometric dimensionality and, hence, structural complexity of the attractor in phase space.

relationship with classical Rényi dimensions (discussions with Grassberger)

add brief information on behavior of other network characteristics?

*Clustering dimension.* With the same rationale as for the global network transitivity, we can make use of the local clustering properties of RNs for defining local measures of attractor dimensionality, referred to as upper and lower clustering dimensions [27]:

$$D_C^u(x) = \limsup_{\varepsilon} \frac{\log(\mathcal{C}(x; \varepsilon))}{\log(3/4)}, \quad (53)$$

$$D_C^l(x) = \liminf_{\varepsilon} \frac{\log(\mathcal{C}(x; \varepsilon))}{\log(3/4)}. \quad (54)$$

Following the same argument as for the (global) transitivity dimensions, we do not need to consider the limit  $\varepsilon \rightarrow 0$  here.

With similar considerations regarding the possible existence of two distinct accumulation points of  $\mathcal{C}(x)$  as  $\varepsilon$  varies, we may utilize this framework for characterizing the point-wise dimension of chaotic attractors in a unique way without making explicit use of scaling characteristics as in the common point-wise dimensions [27]. However, we need to keep in mind that the considered concept of (geometric) dimensionality is largely affected by the profile of the invariant density, e.g., the existence of sharp attractor boundaries or supertrack functions [27, 43, 141]. For example, if the attractor has distinct tips (e.g., in the case of the Hénon system [27, 43]), the geometric dimension at these points is effectively reduced to zero, which is reflected by  $\hat{\mathcal{C}}_v = 1$  for vertices  $v$  sufficiently close to the tips. A similar behavior can be observed for the logistic map at the attractor boundaries and the supertrack functions [27, 43, 141].

The latter observations point to a prospective application of the local clustering properties of RNs. In case of chaotic attractors of time-continuous dynamical systems, it is known that an infinite number of unstable periodic orbits (UPOs) provide the skeleton of the chaotic dynamics and are densely embedded in the attractor. The localization of such UPOs is, however, known to be a challenging task. Since UPOs are relatively weakly repulsive (from a practical perspective, those UPOs with low periods are typically least unstable), a trajectory getting close to the vicinity of an UPO will stay close to this orbit for some finite amount of time [148]. As a result, the dynamics close to UPOs is quasi one-dimensional, and state vectors sampled from the trajectories approximate some lower-dimensional (in the limiting case one-dimensional) subset of the attractor manifold. In such case, the above theoretical considerations suggest that the local clustering coefficient  $\hat{\mathcal{C}}_v$  of vertices  $v$  close to low-periodic UPOs should be higher than the values typical for other parts of the chaotic attractor. This conceptual idea is supported by numerical results from our

previous work [43, 142] (cf. also the band structures with increased  $\hat{\mathcal{C}}_v$  in Fig. 3(b)), but has not yet been systematically applied to the problem of UPO localization. Notably, the detection limit of UPOs should be ultimately determined by the recurrence threshold  $\varepsilon$  in conjunction with the RN size  $N$ . Specifically, for every finite  $\varepsilon > 0$ , there are infinitely many UPOs intersecting with the  $\varepsilon$ -neighborhood of some point  $x_v$  in phase space, whereas we will (for a finite sample of state vectors) only resolve the signatures of the least unstable orbits.

### 3.6. Practical considerations

#### 3.6.1. Dependence on embedding parameters

We focus on two important algorithmic parameters of the RN approach, embedding dimension and delay. The impact of other parameters such as recurrence threshold  $\varepsilon$ , sampling rate, or even the selection of variables in multi-dimensional systems has been extensively discussed elsewhere [142, 149] for deterministic systems, but not yet for stochastic ones. For the sake of brevity, we present only a brief corresponding discussion here. Specifically, since we consider discrete-time univariate stochastic processes, only  $\varepsilon$  is relevant, but can be treated mostly alongside the theoretical considerations presented in [131].

Given a scalar time series  $\{x_i\}$  ( $i = 1, \dots, N$ ), in order to apply RN analysis we first have to convert the data into state vectors in some appropriately reconstructed phase space. A common method from dynamical systems theory to define such a phase space is time-delay embedding [92]. In fact, the concept of a phase space representation rather than a “simple” time or frequency domain approach is the hallmark of many methods of nonlinear time series analysis, requiring embedding as the first step. Here, we define  $\mathbf{x}_i = (x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau})$  to obtain an  $m$ -dimensional time-delay embedding of  $x_i$  with embedding delay  $\tau$  for obtaining state vectors in phase space [92]. It has been proven that for deterministic dynamical systems, the thus reconstructed phase space is topologically equivalent to the original space if  $m > 2D_F$ , where  $D_F$  is the fractal dimension of the support of the invariant measure generated by the dynamics in the true (but often at most partially observed) state space. Note that  $D_F$  can be much smaller than the dimension of the underlying original (physical) phase space spanned by all relevant system variables.

From a practical perspective, when analyzing a scalar time series of whatever origin, neither embedding dimension  $m$  nor delay  $\tau$  are known a priori. The false nearest-neighbors (FNN) method [91] was introduced to derive a reasonable guess of how to choose  $m$  based on studying whether or not proximity relations between state vectors are lost when the embedding dimension is successively increased. If a reasonable embedding dimension is found, all dynamically relevant coordinates of the system are appropriately represented, so that all proximity relationships are correct and not due to lower-dimensional projection effects. In a similar spirit, the first root of the auto-correlation function (ACF) of a time series often yields a good estimate for  $\tau$ . A more refined method is to use time-delayed mutual information [90].

The aforementioned approaches to determining  $m$  and  $\tau$  commonly work well for data from deterministic dynamical systems. Embeddings are not applicable for non-stationary processes though they are more ubiquitous in real time series analysis, for instance, fractional Brownian motions (fBm) and related processes arising from an integration of stationary processes (e.g., fractional Lévy motion, (F)ARIMA models, etc.). More specifically, we have to keep in mind some severe conceptual problems may appear when applying them to non-stationary processes: First, finite estimates of  $D_F$  are spurious due to the finite amount of data used. The latter result is reasonable since an infinite amount of data (i.e., the innovations at each time step) are necessary to fully describe the evolution of a stochastic process. Thus, from a conceptual perspective, the embedding dimension should be chosen infinitely large. In turn, finite  $m$  will necessarily cause spurious results since the full complexity of the system’s (discrete) trajectory is not captured.

On the other hand, the embedding delay  $\tau$  is not considered in the mathematical embedding theorems for deterministic dynamical systems. Embeddings with the same embedding dimension  $m$  but different  $\tau$  are topologically equivalent in the mathematical sense [3], but in reality a good choice of  $\tau$  facilitates further analysis. If  $\tau$  is small compared to the relevant internal time-scales of the system, successive elements of the delay vectors are strongly correlated. This leads to the practical requirement that the embedding delay should cover a much longer time interval than the largest characteristic time-scale that is relevant for the dynamics of the system. However, in fBm arbitrarily long time-scales are relevant due to the self-similar

nature of the process [150]. This makes finding a feasible value of  $\tau$  a challenging (and, regarding formal optimality criteria, even theoretically impossible) task.

We emphasize that in the case of non-stationary fBm, the fundamental concepts of phase space reconstruction and low-dimensional dynamics do not apply (not even approximately) anymore. Therefore, any attempt to applying RN analysis to fBm directly necessarily yields results that hold only for the particular embedding parameters chosen and the specific length of the given time series [151]. In [150], we have demonstrated that RN analysis can indeed provide meaningful results for stationary stochastic processes, given a proper selection of its intrinsic methodological parameters, whereas it is prone to fail to uniquely retrieve RN properties for non-stationary stochastic processes like fBm. In cases of non-stationarity, a proper transformation is required to remove the particular type of non-stationarity from the data. This can be achieved by additive detrending, phase adjustment (de-seasonalization), difference filtering (incrementation) or other techniques, with the one mentioned last being the proper tool for the particular case of fBm transforming the original process into stationary fractional Gaussian noise.

### 3.6.2. Choice of recurrence rate or threshold

The crucial algorithmic parameter of recurrence-based time series analysis is  $\varepsilon$ , which has been discussed extensively in the literature [87, 142]. The more empirical choice of  $\varepsilon$  often depends on time series embedded in phase space. Too small  $\varepsilon$  causes very sparsely connected RN with many isolated components; too large  $\varepsilon$  results in an almost completely connected network. Several invariants of a dynamical system e.g., the second-order Rényi entropy  $K_2$  can be estimated by taking its recurrence properties for  $\varepsilon \rightarrow 0$  [7, 147], which suggests that for a feasible analysis of recurrence networks, a low  $\varepsilon$  is preferable as well. This is supported by the analogy to complex networks based on spatially extended systems, where attention is usually restricted to the strongest links between individual vertices i.e., observations from different spatial coordinates for retrieving meaningful information about relevant aspects of the systems' dynamics. In contrast, a high edge density, does not yield feasible information about the actually relevant structures, because these are hidden in a large set of mainly less important edges [142].

As a consequence, only those states should be connected in a recurrence network that are closely neighbored in phase space, leading to rather sparse networks. Following a corresponding rule of thumb confirmed for recurrence quantification analysis [152], we suggest choosing  $\varepsilon$  as corresponding to an edge density  $\rho \lesssim 0.05$  [42, 43], which yields neighborhoods covering appropriately small regions of phase space. Note that since many topological features of recurrence networks are closely related to the local phase space properties of the underlying attractor [43], the corresponding information is best preserved for such low  $\varepsilon$  unless the presence of noise requires higher  $\varepsilon$  [152].

The heuristic criterion proposed by Gao *et al.* [75], which selects  $\varepsilon$  as the (supposedly unique) turning point of the plot of  $\rho$  vs.  $\varepsilon$ , is *not* generally applicable as we have discussed in [142]. In particular, this heuristic criterion can *not* attribute certain network features to specific *small-scale* attractor properties in phase space [142]. Moreover, besides our general considerations supporting low  $\varepsilon$ , application of the turning point criterion leads to serious pitfalls. One has to emphasize that various typical examples for both discrete and continuous dynamical systems are characterized by *several* turning points. Depending on the particular types of signals from real measurements from civil engineering structures, surrogate-assisted method for choosing optimal threshold by searching for a turning point of proper defined quality loss function might be a good solution [153].

One of the problems preventing a uniform choice of  $\varepsilon$  between different time series is that the size of the attractor after embedding is arbitrary. To overcome this, Jacob [154] proposed first to transform the time series into a uniform deviate so that the size of the attractor gets rescaled into the unit interval  $[0, 1]$ . Then, their choice of  $\varepsilon$  is based on empirical results from numerical computations such that the following two criteria are fulfilled [154]: a) the resulting RN has to remain mostly as “one single cluster” and b) the measures derived from the RN should uniquely represent the underlying attractor. Note that the first condition fixes the lower bound for  $\varepsilon$  which ensures the network becomes fully connected. The second one fixes the upper bound. Furthermore, they show that the above two conditions together provide an identical optimum  $\varepsilon$  range for time series from all chaotic attractors. It should be noted that the above choice of

the critical range of  $\varepsilon$  is, in fact, analogous to the selection of a scaling region in the conventional nonlinear time-series analysis for computing dynamical invariants like correlation dimension  $D_2$  [147].

The above strategies in choosing  $\varepsilon$  help us to overcome the problem of sliding-window-based analyses of systems with varying amplitude fluctuations (as coming from different dynamical regimes or non-stationarities), for instance, those based on normalizing time series or fixing recurrence density. However, in real-world applications, time series are not usually smooth all the time. When considering the time series by a RN representation, extreme points (very high jumps or falls in the fluctuation of time series) in the time series could break the connected components in the network since the distance between an extreme point and other points would be larger than the threshold value [155]. These unconnected components would cause problems for some complex network measures, since some of them need a connected network to be computed for the entire network. For example, even if we have just one node that is not connected to the network, the average path length will always be infinite for the entire network. The normalization method would then result in non-optimal recurrence thresholds biasing the recurrence analysis. An even more important motivation for avoiding isolated components in the RN is that the RN provides a large amount of information about the dynamics of the underlying system, although it contains only binary information. To find a sufficiently small threshold  $\varepsilon$  that fulfills the desired condition of connected neighborhoods, Eroglu *et al.* proposed to use the connectivity properties of the network. In particular, they choose the value for  $\varepsilon$  that is the smallest one for the RN to be connected. The connectivity of a RN is measured by the second-smallest eigenvalue  $\lambda_2$  of the Laplacian matrix associated to the recurrence matrix [155]. Note that this criterion in choosing an adaptive  $\varepsilon$  shares the same idea as used [154, 156]. Note that in the case that phase space is consisted by several disjoint partitions, the method guaranteeing the connectivity might not be feasible, for instance, phase space is not continuous when the control parameter is before the band merging of the Logistic map. Another counter examples has been presented in the standard map where there are several components in phase space [142, 157]).

As mentioned before, small variations of  $\varepsilon$  may lead to very different modular structure in the associated recurrence network, which prompts a big challenge for identifying metastable states in real-world time series. Note that modular regions of RNs may correspond to trajectory bundle, which are associated to the existence of metastable states. This way, selecting an inadequate recurrence threshold can hide important geometric information of the state space of a system. In [158], the authors suggested that an adequate recurrence threshold should lie in a range of values producing recurrence networks with similar modular structures. This means that there is a range of recurrence thresholds for which the associated recurrence networks describe reconstructed state spaces with equivalent topology. However, this region of values depends on the distribution of the particular time series data, which might not be uniform. To this end, they define a filtration procedure to self-adaptively set an adequate  $\varepsilon$  from RNs that are associated to a set of recurrence thresholds. The adequate  $\varepsilon$  belongs to the subset of values in the filtration for which the modular structures of their associated RNs are the least dissimilar. Furthermore, in searching for metastable states [158], the authors suggested to compute modular similarity measures like the Adjusted Rank Index, which then further help to compute the final adequate recurrence threshold  $\varepsilon$ .

### 3.6.3. Stability and robustness against noise

The results of the previous subsections show that RN approaches are able to clearly distinguish between periodic and chaotic dynamics under noise free condition. In experimental time series, one is always confronted with measurement errors. Hence, it is necessary to analyze the influence of noise on the reconstructed RNs. In the framework of recurrence plot, a larger  $\varepsilon$  has been suggested to overcome the noise effect [159], for instance, a threshold  $\epsilon$  that is at least 5 times the standard deviation of the observational Gaussian noise  $\sigma$  can yield reliable statistics. This criterion is based on the analytical computation of the probability of a recurrence point in the RP to be correctly recognized in the presence of observational noise. We suggest to use this criterion if weak observational noise is present as it has been found that the choice  $\epsilon \sim 5\sigma$  is optimal for a wide class of processes [159].

Furthermore, in [160] the authors show that the influence of noise on the network measure  $\mathcal{C}$  can be minimized by an appropriate choice of  $\rho(\varepsilon)$  (e.g., by setting  $\rho(\varepsilon) > 0.02$ ), while the influence on  $\mathcal{L}$  is independent of  $\rho(\varepsilon)$ . However, for noise levels greater than 40% in case of  $\mathcal{C}$  and 20% in case of  $\mathcal{L}$ , the recurrence

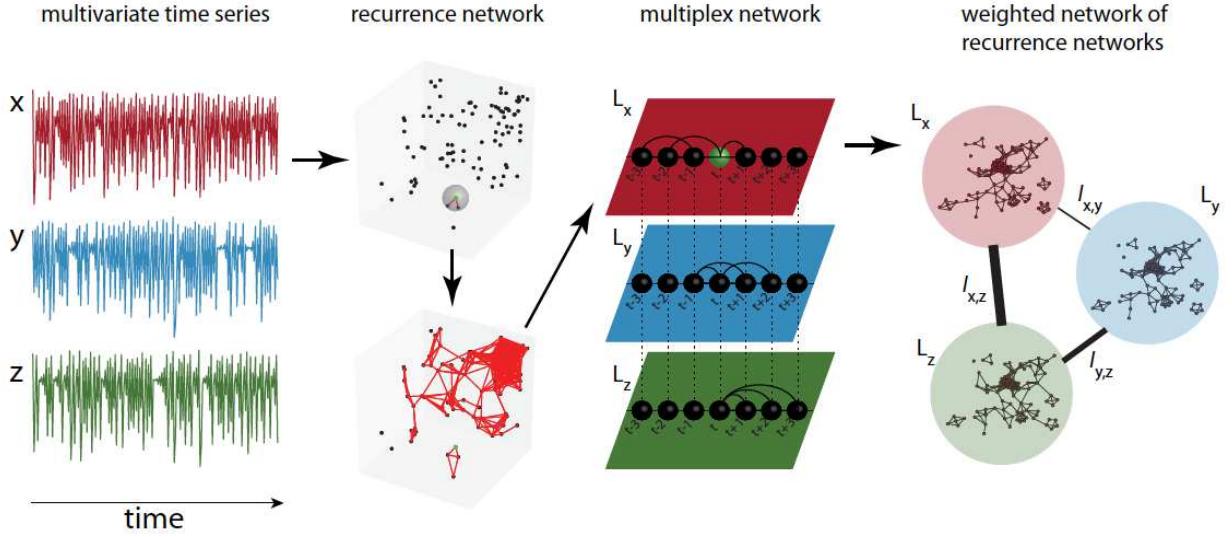


Figure 5: Construction of multiplex recurrence networks for multivariate time series. Courtesy of [162].

network measures fail in distinguishing between noisy periodic dynamics and noisy chaotic dynamics.

It is also possible to test the efficiency of RN measures as discriminating statistics by means of hypothesis testing using surrogate technique [161]. For instance, network measures of RNs are effective if the data involves both white noise and colored noise, which have been further illustrated by experimental time series that has been simulated by a black hole light curves model.

### 3.6.4. Behavior for larger $\varepsilon$

new papers by Jacob et al., cross-over to small-world (spatial network like) behavior

## 3.7. Multiplex recurrence networks

So far, RN approaches has been discussed in the framework of a single system. In the next two subsections, we focus on several different generalizations to multivariate analysis, for instance, multiplex recurrence networks (MRN) and two different versions of multilayer recurrence networks: inter-system recurrence networks (ISRN) and joint recurrence networks (JRN) which are based on cross recurrence plots and joint recurrence plots respectively.

Let us start with the construction of multiplex recurrence networks, as schematically illustrated in Fig. 5. If in a multilayer network of  $M$  layers, each layer has the same set of vertices and the connections between layers are only between a node and its counterpart in the other layers, then we call such networks “multiplex”. In [162], Eroglu *et al* proposed to construct multiplex recurrence networks, which shares much similarities in constructing multiplex visibility graphs [163]. Consider an  $M$ -dimensional multivariate time series  $\{\mathbf{x}(t)\}_{t=1}^n$ , with  $\mathbf{x}(t) = (x^{[1]}(t), x^{[2]}(t), \dots, x^{[M]}(t)) \in \mathbb{R}^M$  for any value of  $t$ . Then, the RN of the  $\alpha$ th component of  $\mathbf{x}(t)$  is created and located into the associated layer  $\alpha$  of the multiplex network. For the  $M$ -dimensional multivariate time series, we can create  $M$  different RNs which have the same number of nodes and each node is labeled by its associated time. These networks will form the different layers of a multilayer network. The layers are connected each other with the same time labeled nodes. Furthermore this procedure requires that the time sampling is the same for all of the used time series. Networks, transformed from multivariate time series, are compatible with the definition of multiplex networks, because each node is uniquely assigned to a certain time point of the multivariate time series, i.e., we find the equally time-labeled nodes in all layers.

can I get the source figure from Deniz? Figure 1 to explain the construction steps, highly acknowledged.

More specifically, we denote the adjacency matrix of the  $\alpha$ th layer as  $A^{[\alpha]} = a_{ij}^{[\alpha]}$  and  $a_{ij}^{[\alpha]} = 1$  if nodes  $i$  and  $j$  are connected in layer  $\alpha$ ,  $a_{ij}^{[\alpha]} = 0$  otherwise. Then the entire multiplex network  $\mathcal{A}$  is represented by the vector of adjacency matrices of its layers  $\mathcal{A} = \{A^{[1]}, A^{[2]}, \dots, A^{[M]}\}$ , which read

$$\mathcal{A} = \begin{bmatrix} \mathbf{A}^{[1]} & \mathbf{I}_{[N]} & \dots & \mathbf{I}_N \\ \mathbf{I}_{[N]} & \mathbf{A}^{[2]} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{I}_N \\ \mathbf{I}_{[N]} & \dots & \mathbf{I}_N & \mathbf{A}^{[m]} \end{bmatrix}_{Nm \times Nm}, \quad (55)$$

where  $\mathbf{I}_N$  is the identity matrix of size  $N$ .

Two different measures have been proposed to quantify the similarity between layer  $\alpha$  and  $\beta$  of the giant multiplex network [162, 163]. The first is to compute the interlayer mutual information  $I_{\alpha,\beta}$ :

$$I_{\alpha,\beta} = \sum_{k^{[\alpha]}} \sum_{k^{[\beta]}} P(k^{[\alpha]}, k^{[\beta]}) \log \frac{P(k^{[\alpha]}, k^{[\beta]})}{P(k^{[\alpha]})P(k^{[\beta]})}, \quad (56)$$

where  $P(k^{[\alpha]}, k^{[\beta]})$  is the joint probability of the existence of nodes which has  $k^{[\alpha]}$  degree at layer  $\alpha$  and  $k^{[\beta]}$  at layer  $\beta$ ,  $P(k^{[\alpha]})$  and  $P(k^{[\beta]})$  are the degree distributions of RNs at layer  $\alpha$  and  $\beta$  respectively. Since  $I_{\alpha,\beta}$  is computed based on the degree sequences, instead of the original time series, the mutual information  $I_{\alpha,\beta}$  considers the topological recurrence structures in phase space.

Another measure to quantify the coherence of the original multivariate system is the average edge overlap [162, 163]:

$$\omega = \frac{\sum_i \sum_{j>i} \sum_{\alpha} a_{ij}^{[\alpha]}}{m \sum_i \sum_{j>i} (1 - \delta_{0, \sum_{\alpha} a_{ij}^{[\alpha]}})}, \quad (57)$$

where  $\delta_{ij}$  is the Kronecker delta symbol. This measure represents the average number of identical edges over all layers of the multiplex network [163]. Like the interlayer mutual information Eq. (56),  $\omega$  estimates the similarity and coherence with averaged existence of overlapped links from nodes  $i$  to  $j$  between all layers  $\alpha$  and  $\beta$ .

Note that the interlayer similarity measures (Eqs. (56), (57)) are computed for each pair layers. The giant adjacency matrix  $\mathcal{A}$  of the multiplex network can be projected onto one weighted network representation encompassing the interlayer information only. In other words, we consider each single layer of the multiplex as a node and weighted edges between nodes  $\alpha$  and  $\beta$  are determined by the quantity  $I_{\alpha,\beta}$ , which yields that the size of the projection weighted network  $M \times M$ .

### 3.8. Multilayer networks: coupled networks

#### 3.8.1. General preliminaries

When describing multilayer and multiplex networks, we introduced the traditional notations in Sec. 2.4. However, in the particular case of networks reconstructed from two or more interacting time series, we prefer to utilize a recently introduced general framework [164, 165]. Let us consider an arbitrary undirected and unweighted simple graph  $G = (V, E)$  with adjacency matrix  $\mathbf{A} = \{A_{ij}\}_{i,j=1}^N$ . Furthermore, let us assume that there is a given partition of  $G$  with the following properties:

1. The vertex set  $V$  is decomposed into  $K$  disjunct subsets  $V_k \subseteq V$  such that  $\bigcup_{k=1}^K V_k = V$  and  $V_k \cap V_l = \emptyset$  for all  $k \neq l$ . The cardinality of  $V_k$  will be denoted as  $N_k$ .
2. The edge set  $E$  consists of mutually disjoint sets  $E_{kl} \subseteq E$  with  $\bigcup_{k,l=1}^K E_{kl} = E$  and  $E_{kl} \cap E_{mn} = \emptyset$  for all  $(k, l) \neq (m, n)$ .
3. Let  $E_{kl} \subseteq V_k \times V_l$ . Specifically, for all  $k = 1, \dots, K$ ,  $G_k = (V_k, E_{kk})$  is the induced subgraph of the vertex set  $V_k$  with respect to the full graph  $G$ .

Under these conditions,  $E_{kk}$  comprises the (internal) edges within  $G_k$ , whereas  $E_{kl}$  contains all (cross-) edges connecting  $G_k$  and  $G_l$ . Specifically, for the “natural” partition of an IRN, the  $G_k$  correspond to the single-system RNs constructed from the systems  $X_k$ , whereas the cross-recurrence structure is encoded in  $E_{kl}$  for  $k \neq l$ .

We are now in a position to study the interconnectivity structure between two subnetworks  $G_k, G_l$  on several topological scales drawing on the lineup of local and global graph-theoretical measures generalizing those used for single network characterization (Section 3.3). In this context, local measures  $\hat{f}_v^{kl}$  characterize a property of vertex  $v \in V_k$  with respect to subnetwork  $G_l$ , while global measures  $\hat{f}_{kl}$  assign a single real number to a pair of subnetworks  $G_k, G_l$  to quantify a certain aspect of their mutual interconnectivity structure. Most interconnectivity characteristics discussed below have been originally introduced in [164] which see for more detailed discussions.

### 3.8.2. Vertex characteristics

The *cross-degree* (or *cross-degree centrality*)

$$\hat{k}_v^{kl} = \sum_{i \in V_l} A_{vi} \quad (58)$$

counts the number of neighbors of  $v$  within  $G_l$ , i.e., direct connections between  $G_k$  and  $G_l$  (Fig. ??A). Thus, this measure provides information on the relevance of  $v$  for the network “coupling” between  $G_k$  and  $G_l$ <sup>2</sup>. For the purpose of the present work, it is useful studying a normalized version of this measure, the *cross-degree density*

$$\hat{\rho}_v^{kl} = \frac{1}{N_l} \sum_{i \in V_l} A_{vi} = \frac{1}{N_l} \hat{k}_v^{kl}. \quad (59)$$

For an IRN,  $\hat{\rho}_v^{kl}(\varepsilon_{kl})$  equals the (cross-) recurrence rate  $RR_{kl}(\varepsilon_{kl})$  (for  $k = l$ , it gives the corresponding single-system recurrence rate  $RR_k(\varepsilon_k)$ ).

As for the single network case, important information is governed by the presence of triangles in the network. Given two subnetworks, the *local cross-clustering coefficient*

$$\hat{C}_v^{kl} = \frac{1}{\hat{k}_v^{kl}(\hat{k}_v^{kl} - 1)} \sum_{i,j \in V_l} A_{vi} A_{ij} A_{jv}, \quad (60)$$

measures the relative frequency of two randomly drawn neighbors  $i, j \in V_l$  of  $v \in V_k$  are mutually connected (Fig. ??B). For  $\hat{k}_v^{kl} < 2$ , we define  $\hat{C}_v^{kl} = 0$ . In general,  $\hat{C}_v^{kl}$  characterizes the tendency of vertices in  $G_k$  to connect to clusters of vertices in  $G_l$ .

The *cross-closeness centrality*

$$\hat{c}_v^{kl} = \left( \frac{\sum_{i \in V_l} d_{vi}}{N_l} \right)^{-1} \quad (61)$$

(where  $d_{vi}$  is the graph-theoretical shortest-path length between  $v$  and  $i$ ) characterizes the topological closeness of  $v \in G_k$  to  $G_l$ , i.e., the inverse arithmetic mean of the shortest path lengths between  $v$  and all vertices  $i \in V_l$ . If there exist no such paths,  $d_{vi}$  is commonly set to the maximum possible value  $N - 1$  given the size of  $G$ . As in the single network case, replacing the arithmetic by the harmonic mean yields the *local cross-efficiency*

$$\hat{e}_v^{kl} = \frac{\sum_{i \in V_l} d_{vi}^{-1}}{N_l}, \quad (62)$$

which can be interpreted in close analogy to  $\hat{c}_v^{kl}$ . Note that in the case of IRNs, topological closeness directly implies geometric closeness.

---

<sup>2</sup>In the specific case of an IRN, we interpret this as geometric signatures of the coupling between the underlying dynamical systems  $X_k$  and  $X_l$  [166, 167].

As a final vertex characteristic, we may generalize the betweenness concept to the case of coupled subnetworks, which results in the *cross-betweenness centrality*

$$\hat{b}_v^{kl} = \sum_{i \in V_k, j \in V_l; i, j \neq v} \frac{\hat{\sigma}_{ij}(v)}{\hat{\sigma}_{ij}}. \quad (63)$$

Here,  $\hat{\sigma}_{ij}(v)$  and  $\hat{\sigma}_{ij}$  are defined as in the case of a single network. Note that unlike the other vertex characteristics discussed above, in the case of cross-betweenness centrality, we do not require  $v$  belonging to  $G_k$  or  $G_l$  (Fig. ??C). The reason for this is that vertices belonging to any subnetwork may have a non-zero betweenness regarding two given subgraphs  $G_k$  and  $G_l$ , in the sense that shortest paths between  $i \in V_k$  and  $j \in V_l$  can also include vertices in other subnetworks.

### 3.8.3. Global characteristics

The density of connections between two subnetworks can be quantified by taking the arithmetic mean of the local cross-degree density (Eq. 59), yielding the *cross-edge density*

$$\hat{\rho}^{kl} = \frac{1}{N_k N_l} \sum_{i \in V_k, j \in V_l} A_{ij} = \hat{\rho}^{lk}. \quad (64)$$

Notably,  $\hat{\rho}^{kl}$  corresponds to the definition of the *cross-recurrence rate*  $RR_{kl}$  (Eq. 70). Since we consider here only undirected networks (i.e., bidirectional edges), the cross-edge density is invariant under mutual exchange of the two considered subnetworks.

The *global cross-clustering coefficient*

$$\hat{\mathcal{C}}^{kl} = \left\langle \hat{\mathcal{C}}_v^{kl} \right\rangle_{v \in V_k} = \frac{1}{N_k} \sum_{v \in V_k, \hat{k}_v^{kl} > 1} \frac{\sum_{i, j \in V_l} A_{vi} A_{ij} A_{jv}}{\sum_{i \neq j \in V_l} A_{vi} A_{vj}} \quad (65)$$

estimates the probability of vertices in  $G_k$  to have mutually connected neighbors in  $G_l$ . Unlike the cross-edge density, the corresponding “cross-transitivity” structure is typically asymmetric, i.e.,  $\hat{\mathcal{C}}^{kl} \neq \hat{\mathcal{C}}^{lk}$ . As in the single network case, we need to distinguish  $\hat{\mathcal{C}}^{kl}$  from the *cross-transitivity*

$$\hat{\mathcal{T}}^{kl} = \frac{\sum_{v \in V_k; i, j \in V_l} A_{vi} A_{ij} A_{jv}}{\sum_{v \in V_k; i \neq j \in V_l} A_{vi} A_{vj}}, \quad (66)$$

for which we generally have  $\hat{\mathcal{T}}^{kl}(\varepsilon) \neq \hat{\mathcal{T}}^{lk}(\varepsilon)$  as well. Again, we have to underline that cross-transitivity and global cross-clustering coefficient are based on a similar concept, but capture distinctively different network properties.

Regarding the quantification of shortest path-based characteristics, we define the *cross-average path length*

$$\hat{\mathcal{L}}^{kl} = \frac{1}{N_k N_l} \sum_{i \in V_k, j \in V_l} d_{ij} \quad (67)$$

and the *global cross-efficiency*

$$\hat{\mathcal{E}}^{kl} = \left( \frac{1}{N_k N_l} \sum_{i \in V_k, j \in V_l} d_{ij}^{-1} \right)^{-1} \quad (68)$$

Unlike  $\hat{\mathcal{C}}^{kl}$  and  $\hat{\mathcal{T}}^{kl}$ ,  $\hat{\mathcal{L}}^{kl}$  and  $\hat{\mathcal{E}}^{kl}$  are (as shortest path-based measures) symmetric by definition, i.e.,  $\hat{\mathcal{L}}^{kl}(\varepsilon) = \hat{\mathcal{L}}^{lk}(\varepsilon)$  and  $\hat{\mathcal{E}}^{kl}(\varepsilon) = \hat{\mathcal{E}}^{lk}(\varepsilon)$ . In the case of disconnected network components, the shortest path length  $d_{ij}$  is defined as discussed for the corresponding local measures.

In the same spirit as shown above, other single network characteristics [10, 30] can be adopted as well for defining further interdependent network measures. This includes measures characterizing edges or, more

generally, pairs of vertices like edge betweenness or matching index, further global network characteristics (assortativity, network diameter and radius), mesoscopic structures (motifs), or even characteristics associated with diffusion processes on the network instead of shortest paths (e.g., eigenvector centrality or random walk betweenness). The selection of measures introduced above reflects those characteristics which have the most direct interpretation in the context of IRNs and have also been utilized in studying the interdependence structure between complex networks in other contexts [164, 165].

### 3.9. Inter-system recurrence networks

In the last decade, two different widely applicable bi- and multivariate extensions of RPs and RQA have been proposed [87]: cross-recurrence plots [84, 96] and joint recurrence plots [85]. In the following, we discuss some possibilities for utilizing these approaches in a complex network framework, following previous considerations in [166–168]. For this purpose, let us consider  $K$  (possibly multivariate) time series  $\{x_i^k\}_{i=1}^{N_k}$  with  $x_i^k = x^k(t_i^k)$  sampled at times  $\{t_i^k\}$  from dynamical systems  $\{X_k\}$  with  $k = 1, \dots, K$ .

#### 3.9.1. Cross-recurrence plots

One way of extending recurrence analysis to the study of multiple dynamical systems is looking at *cross-recurrences*<sup>3</sup>, i.e., encounters of the trajectories of two systems  $X_k$  and  $X_l$  sharing the same phase space, where  $x_i^k \approx x_j^l$  [84, 96] (see Fig. ?? for some illustration). Unlike the traditional recurrence matrix  $\mathbf{R}$  of a single system, the elements of the cross-recurrence matrix  $\mathbf{CR}^{kl}$  are defined as

$$CR_{ij}^{kl}(\varepsilon_{kl}) = \Theta(\varepsilon_{kl} - \|x_i^k - x_j^l\|), \quad (69)$$

where  $i = 1, \dots, N_k$ ,  $j = 1, \dots, N_l$ , and  $\varepsilon_{kl}$  is a prescribed threshold distance in the joint phase space of both systems. As in the single-system case,  $\varepsilon_{kl}$  determines the number of mutual neighbors in phase space, quantified by the *cross-recurrence rate*

$$RR_{kl}(\varepsilon_{kl}) = \frac{1}{N_k N_l} \sum_{i=1}^{N_k} \sum_{j=1}^{N_l} CR_{ij}^{kl}(\varepsilon_{kl}), \quad (70)$$

which is a monotonically increasing function of  $\varepsilon_{kl}$  (i.e., the larger the distance threshold in phase space, the more neighbors are found). Note that unlike  $\mathbf{R}^k$  and  $\mathbf{R}^l$ , the cross-recurrence matrix  $\mathbf{CR}^{kl}$  is asymmetric, since we typically have  $\|x_i^k - x_j^l\| \neq \|x_j^l - x_i^k\|$ . Even more, it can be non-square if time series of different lengths ( $N_k \neq N_l$ ) are considered.

Due to the aforementioned characteristics,  $\mathbf{CR}^{kl}$  cannot be directly interpreted as the adjacency matrix of a network with similar properties as single-system RNs. This is because the indices  $i$  and  $j$  label two distinct sets of state vectors belonging to systems  $X_k$  ( $i$ ) and  $X_l$  ( $j$ ), respectively. In turn, we can interpret the state vectors  $\{x_i^k\}$  and  $\{x_j^l\}$  as two distinct groups of vertices, and  $\mathbf{CR}^{kl}$  as being an adjacency matrix of a *cross-recurrence network (CRN)* providing a binary encoding of the presence of edges between vertices belonging to different groups. This is the defining property of bipartite graphs [31].

Bipartite networks are found in a wide range of fields [169, 170] and can be understood as a generic way for describing arbitrary complex networks [171, 172]. The large variety of applications of bipartite graphs has triggered great interest in models describing their properties in an appropriate way. Particular attention has been spent on the problem of community detection [173], involving new definitions for the modularity function [169, 174–176] and the development of proper algorithms for community detection [174, 177–179], partially relating to the spectral properties of the networks. However, their specific structure renders some traditional definitions of network-theoretic measures non-applicable, calling for generalizations or even redefinitions of quantities such as the clustering coefficient [180, 181]. This is why we do not further consider explicit quantification of the properties of the bipartite CRN, but follow a different approach detailed below.

---

<sup>3</sup>It is important to realize that cross-recurrences are not to be understood in the classical sense of Poincaré's considerations, since they do not indicate the return of an isolated dynamical system to some previously assumed state. In contrast, they imply an arbitrarily delayed close encounter of the trajectories of two *distinct* systems and, therefore, should be better named *cross encounters* instead. Following the same reasoning, terms such as *cross-recurrence plot* or *cross-recurrence rate* are suggestive, but potentially misleading. However, to comply with the existing literature on cross-recurrence plots, we will adopt the established terms even despite their conceptual ambiguities.

### 3.9.2. Coupled networks framework

As mentioned in Section ??, there is a lack of appropriate measures for characterizing explicit bipartite network structures as compared with the rich toolbox of general-purpose complex network characteristics [10, 30]. Therefore, instead of explicitly investigating the bipartite structure of the CRN, it is more useful to combine the information contained in the single-system recurrence matrices  $\mathbf{R}^k(\varepsilon_k)$  and the cross-recurrence matrices  $\mathbf{CR}^{kl}(\varepsilon_{kl})$  to construct an inter-system recurrence matrix [167]

$$\mathbf{IR}(\varepsilon) = \begin{pmatrix} \mathbf{R}^1(\varepsilon_{11}) & \mathbf{CR}^{12}(\varepsilon_{12}) & \dots & \mathbf{CR}^{1K}(\varepsilon_{1K}) \\ \mathbf{CR}^{21}(\varepsilon_{21}) & \mathbf{R}^2(\varepsilon_{22}) & \dots & \mathbf{CR}^{2K}(\varepsilon_{2K}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CR}^{K1}(\varepsilon_{K1}) & \mathbf{CR}^{K2}(\varepsilon_{K2}) & \dots & \mathbf{R}^K(\varepsilon_{KK}) \end{pmatrix}. \quad (71)$$

Here,  $\varepsilon = (\varepsilon_{kl})_{kl}$  is a  $K \times K$  matrix containing the single-system recurrence thresholds  $\varepsilon_{kk} = \varepsilon_k$  and (cross-recurrence) distance thresholds  $\varepsilon_{kl}$ . The corresponding *inter-system recurrence network (IRN)* [167] (see Fig. ?? for an example) is fully described by its adjacency matrix

$$\mathbf{A}(\varepsilon) = \mathbf{IR}(\varepsilon) - \mathbf{1}_N, \quad (72)$$

where  $N = \sum_{k=1}^K N_k$  is the number of vertices and  $\mathbf{1}_N$  the  $N$ -dimensional identity matrix. As in the case of single-system RNs, the IRN is an undirected and unweighted simple graph, which additionally obeys a natural partition of its vertex and edge set (see Section ??). Vertices represent state vectors in the phase space common to all systems  $X_k$  and edges indicate pairs of state vectors from *either* the same *or* two different systems that are mutually close, whereby the definition of closeness can vary between different pairs of systems. To this end, we briefly mention two specific choices that may be convenient:

- Since we assume the considered systems to share the same phase space, it can be reasonable to measure distances in a way disregarding the specific membership of vertices to the different systems under study. This would imply choosing  $\varepsilon_{kl} = \varepsilon$  as equal values for all  $k, l = 1, \dots, K$ . In such a case, we can reinterpret the IRN as the RN constructed from the concatenated time series

$$\{y_i\}_{i=1}^N = (x_1^1, \dots, x_{N_1}^1, x_1^2, \dots, x_{N_2}^2, \dots, x_1^K, \dots, x_{N_K}^K).$$

In this situation, we can reconsider the general framework of single-system RN analysis as discussed above for studying the geometric properties of the combined system as reflected in a RN. Note, however, that in this case it is hardly possible to explicitly exploit the given natural partitioning of the concatenated data<sup>4</sup>. In contrast, all state vectors are treated in exactly the same way.

- An alternative choice of recurrence and distance thresholds is based on considering the individual single-system RNs are quantitatively comparable. Since some of the network measures discussed in Section 3.3 explicitly depend on the number of existing edges in the network, this requirement calls for networks with the same edge density  $\rho_k = \rho$  for all  $k = 1, \dots, K$ . In other words, the recurrence thresholds  $\varepsilon_{kk}$  ( $k = 1, \dots, K$ ) should be chosen such that the (single-system) recurrence rates are equal ( $RR_1 = \dots = RR_K = RR$ ). Given the natural partitioning of the IRN vertex set, such network can be viewed and statistically analyzed as a network of networks (see Section ??). In this case, in order to highlight the interconnectivity structure of the individual RNs, it is beneficial to choose the distance thresholds  $\varepsilon_{kl}$  for  $k \neq l$  such that the resulting cross-recurrence rates  $RR_{kl}$  yield  $RR_{kl} < RR_k = RR_l = RR$  and possibly also take the same values  $RR_{kl} = CRR < RR$  for all  $k \neq l$ .

---

<sup>4</sup>One corresponding strategy could be utilizing methods for community detection in networks [173], such as consideration of modularity [182]. Notably, such idea has not yet been explored in the context of RN analysis, and it is unclear to what extent the inferred possible community structure of an IRN could exhibit relevant information for studying any geometric signatures associated with the mutual interdependences between different dynamical systems. To this end, we leave this problem for future research.

Table 2: Comparison of two multivariate generalizations of RN analysis regarding the principal requirements on the time series to be analyzed. *Identical* means that a specific property must be the same for all involved time series, while *arbitrary* implies that this does not need to be the case.

	IRN	JRN
Length	arbitrary	identical
Sampling	arbitrary	identical
Physical units	identical	arbitrary
Phase space dimension	identical	arbitrary

As already stated above, the meaningful construction and analysis of IRNs requires time series  $\{x_i^k\}_{i=1}^{N_k}$  that share the same phase space and, hence, describe the same observables with identical physical units (Table 2). However, the time series under study can in principle be sampled at arbitrary times  $\{t_i^k\}_{i=1}^{N_k}$  and have different lengths  $N_k$ , because the method discards all information on time and focuses exclusively on neighborhood relationships in phase space. This type of geometric information is what can be exploited for studying coupling structures between different dynamical systems as reflected by the spatial arrangement of state vectors in the joint phase space (see Section ??).

### 3.9.3. Analytical description

In the same spirit as for the single-system RNs (Section ??), we can consider the graph-theoretical measures for studying the interconnections between subnetworks within IRNs (Section ??) as discrete approximations of more general geometric properties [183]. Let  $S_k \subset Y$  be a subset of an  $m$ -dimensional compact smooth manifold  $Y$  and  $p_k(x)$  represent its invariant density for all  $k = 1, \dots, K$ , where  $x \in S_k$ . In the following, the  $S_k$  and  $p_k$  are assumed to fulfill the same requirements that are stated for  $S$  and  $p$  in Section ???. Notably, the  $S_k$  are assumed to have a considerable non-empty pairwise intersections. We will use the abbreviation  $\int d\mu_k(x) = \int_{S_k} d^m x p_k(x)$ , where  $\mu_k$  is a probability measure on  $S_k$ . For simplicity, only a single recurrence threshold  $\varepsilon = \varepsilon_{kl}$  for all  $k, l$  will be used in the following. The generalization to different values of  $\varepsilon_{kl}$  is straightforward.

*Local measures.* The *continuous  $\varepsilon$ -cross-degree density*

$$\rho^{kl}(x; \varepsilon) = \int_{B_\varepsilon(x) \cap S_l} d\mu_l(y) = \int d\mu_l(y) \Theta(\varepsilon - \|x - y\|) \quad (73)$$

measures the probability that a randomly chosen point in  $S_l$  is found in the neighborhood  $B_\varepsilon(x)$  of  $x \in S_k$ . Its discrete version is the cross-degree density  $\hat{\rho}_v^{kl}(\varepsilon)$  (Eq. 59).

The *continuous local  $\varepsilon$ -cross-clustering coefficient*

$$\mathcal{C}^{kl}(x; \varepsilon) = \frac{\iint_{B_\varepsilon(x) \cap S_l} d\mu_l(y) d\mu_l(z) \Theta(\varepsilon - \|y - z\|)}{\rho^{kl}(x; \varepsilon)^2} \quad (74)$$

gives the probability that two randomly chosen points  $y, z \in S_l$  are  $\varepsilon$ -close to each other ( $\|y - z\| < \varepsilon$ ) if they both lie in the neighborhood of  $x \in S_k$ .  $\mathcal{C}^{kl}(x; \varepsilon)$  is approximated by the discrete local cross-clustering coefficient  $\hat{\mathcal{C}}_v^{kl}(\varepsilon)$  (Eq. 60).

Considering the mutual global geometry of the sets  $S_k, S_l$ , we furthermore introduce *continuous  $\varepsilon$ -cross-closeness centrality*

$$c^{kl}(x; \varepsilon) = \left( \int d\mu_l(y) \frac{g(x, y)}{\varepsilon} \right)^{-1} \quad (75)$$

quantifying the closeness of  $x \in S_k$  to all points of the set  $S_l$  along geodesics together with the related harmonic *continuous local  $\varepsilon$ -cross-efficiency*

$$e^{kl}(x; \varepsilon) = \int d\mu_l(y) \left( \frac{g(x, y)}{\varepsilon} \right)^{-1}. \quad (76)$$

Here, geodesics are defined with respect to the union of all involved systems' attractors  $S = \bigcup_{k=1}^K S_k$  and  $g(x, y)$  is a suitable distance metric on such geodesics (Section ??). The proposed local path-based measures for interdependent networks are approximated by the discrete cross-closeness centrality  $\hat{c}_v^{kl}(\varepsilon)$  (Eq. 61) and local cross-efficiency  $\hat{e}_v^{kl}(\varepsilon)$  (Eq. 62).

Finally, we define the *continuous  $\varepsilon$ -cross-betweenness centrality*

$$b^{kl}(x; \varepsilon) = \int \int d\mu_k(y) d\mu_l(z) \frac{\sigma(y, z|x; \varepsilon)}{\sigma(y, z; \varepsilon)}. \quad (77)$$

As in the single network case,  $\sigma(y, z|x; \varepsilon)$  denotes the number of times  $x \in S$  (i.e., from any arbitrary subnetwork) lies on a geodesic between  $y \in S_k$  and  $z \in S_l$ , and  $\sigma(y, z; \varepsilon)$  denotes the total number of such geodesics. Regarding the appropriate parametrization of  $\sigma(y, z|x; \varepsilon)$ , we refer to our discussion for the single network case in Section ???. The discrete estimator  $\hat{b}_v^{kl}(\varepsilon)$  of  $b^{kl}(x; \varepsilon)$  is given in Eq. (63).

*Global measures.* The simplest continuous global property describing the geometric overlap between the sets  $S_k$  and  $S_l$  is the *continuous  $\varepsilon$ -cross-edge density*

$$\rho^{kl}(\varepsilon) = \int \int d\mu_k(x) d\mu_l(y) \Theta(\varepsilon - \|x - y\|) = \rho^{lk}(\varepsilon) \quad (78)$$

that is empirically estimated by the discrete cross-edge density  $\hat{\rho}^{kl}(\varepsilon)$  (Eq. 64).

The expectation value of the continuous local  $\varepsilon$ -cross-clustering coefficient  $\mathcal{C}^{kl}(x; \varepsilon)$  is referred to as the *continuous global  $\varepsilon$ -cross-clustering coefficient*

$$\mathcal{C}^{kl}(\varepsilon) = \int d\mu_k(x) \mathcal{C}^{kl}(x; \varepsilon), \quad (79)$$

which is approximated by the discrete global cross-clustering coefficient  $\hat{\mathcal{C}}^{kl}(\varepsilon)$  (Eq. 65). Moreover, designed for quantifying transitivity in the cross-recurrence structure, the *continuous  $\varepsilon$ -cross-transitivity*

$$\mathcal{T}^{kl}(\varepsilon) = \frac{\iiint d\mu_k(x) d\mu_l(y) d\mu_l(z) \Theta(\varepsilon - \|x - y\|) \Theta(\varepsilon - \|y - z\|) \Theta(\varepsilon - \|z - x\|)}{\iiint d\mu_k(x) d\mu_l(y) d\mu_l(z) \Theta(\varepsilon - \|x - y\|) \Theta(\varepsilon - \|x - z\|)} \quad (80)$$

gives the probability that two randomly chosen points  $y, z \in S_l$  which are  $\varepsilon$ -close to a randomly chosen point  $x \in S_k$  are also  $\varepsilon$ -close with respect to each other.  $\mathcal{T}^{kl}(\varepsilon)$  is approximated by the discrete cross-transitivity  $\hat{\mathcal{T}}^{kl}(\varepsilon)$  (Eq. 66). As in the case of the discrete estimators, the two latter quantities are in general not symmetric, i.e.,  $\mathcal{C}^{kl}(\varepsilon) \neq \mathcal{C}^{lk}(\varepsilon)$  and  $\mathcal{T}^{kl}(\varepsilon) \neq \mathcal{T}^{lk}(\varepsilon)$ .

While the two former measures depend only on the local overlap structure between  $S_k$  and  $S_l$  together with the invariant densities  $p_k(x)$  and  $p_l(x)$ , path-based measures contain information on the global geometry of both sets. The *continuous  $\varepsilon$ -cross-average path length*

$$\mathcal{L}^{kl}(\varepsilon) = \int \int d\mu_k(x) d\mu_l(y) \frac{g(x, y)}{\varepsilon} = \mathcal{L}^{lk}(\varepsilon) \quad (81)$$

gives the average length of geodesic paths starting in  $S_k$  and ending in  $S_l$  or vice versa. Similarly, we define the *continuous global  $\varepsilon$ -cross-efficiency*

$$\mathcal{E}^{kl}(\varepsilon) = \left( \int \int d\mu_k(x) d\mu_l(y) \left( \frac{g(x, y)}{\varepsilon} \right)^{-1} \right)^{-1} = \mathcal{E}^{lk}(\varepsilon) \quad (82)$$

which is the harmonic mean geodesic distance between  $S_k$  and  $S_l$ . Discrete approximations of these global path-based quantifiers are provided by the cross-average path length  $\hat{\mathcal{L}}^{kl}(\varepsilon)$  (Eq. 67) and global cross-efficiency  $\hat{\mathcal{E}}^{kl}(\varepsilon)$  (Eq. 68), respectively. As for their discrete estimators, the path-based characteristics  $\mathcal{L}^{kl}(\varepsilon)$  and  $\mathcal{E}^{kl}(\varepsilon)$  are invariant under an exchange of  $S_k$  and  $S_l$ .

### 3.9.4. Geometric signatures of coupling

The new class of statistical network measures designed for investigating the topology of networks of networks discussed in Section ?? is readily applicable for analyzing the interdependency structure of multiple complex dynamical systems. For the special case of two coupled systems  $X$  and  $Y$ , we have demonstrated numerically that in an IRN, the asymmetry intrinsic to the global measures cross-transitivity  $\hat{\mathcal{T}}^{XY}$  and global cross-clustering coefficient  $\hat{\mathcal{C}}^{XY}$  can be exploited to reliably detect the direction of coupling between chaotic oscillators over a wide range of coupling strengths, requiring only a relatively small number of samples  $N_{X,Y} \sim \mathcal{O}(10^2 \dots 10^3)$  [167]. For this purpose, we make again use of the fact that transitivity-based characteristics quantify subtle geometric properties that can be easily evaluated both analytically and numerically.

In order to see how cross-transitivities and global cross-clustering coefficients capture dynamical signatures of asymmetric vs. symmetric coupling configurations, let us assume a diffusive coupling with positive sign (i.e., an attractive interaction) as in Eq. (??). In the uncoupled case, cross-triangles arise randomly according to the sampling from the systems' respective invariant densities. In this case, eventual asymmetries between  $\hat{\mathcal{T}}^{XY}$  and  $\hat{\mathcal{T}}^{YX}$  (or, equivalently,  $\hat{\mathcal{C}}^{XY}$  and  $\hat{\mathcal{C}}^{YX}$ ) originate from the geometry of the respective sets and the associated  $p(x)$ , which should already be reflected in the single-system RN transitivities and global clustering coefficients. In turn, if both systems are represented by the same set of state variables (a prerequisite for the application of IRNs) and obey similar values of  $\hat{\mathcal{T}}^X$  and  $\hat{\mathcal{T}}^Y$  ( $\hat{\mathcal{C}}^X$  and  $\hat{\mathcal{C}}^Y$ ), it is likely that also  $\hat{\mathcal{T}}^{XY}$  and  $\hat{\mathcal{T}}^{YX}$  ( $\hat{\mathcal{C}}^{XY}$  and  $\hat{\mathcal{C}}^{YX}$ ) take similar values. Note that minor asymmetries in the interdependent network characteristics can already occur if both systems are only weakly non-identical, e.g., when considering uncoupled identical Rössler systems with just a small detuning of their natural frequencies [167].

Let us suppose now that there is a unidirectional coupling  $X \rightarrow Y$ . In this case, the trajectory of the driven system  $Y$  is attracted by that of the driver  $X$  due to the considered form of coupling. As a result, it is likely to find more states in  $Y$  that are close to mutually connected pairs of states in  $X$  than in the uncoupled case. This implies that  $\hat{\mathcal{T}}^{YX}$  ( $\hat{\mathcal{C}}^{YX}$ ) increases since  $X$  is “pulling” the trajectory of  $Y$  and, hence, the number of triangles having their baseline in system  $X$  increases relatively to those having their baseline in  $Y$ . Consequently, we expect to have  $\hat{\mathcal{T}}^{YX} > \hat{\mathcal{T}}^{XY}$  and  $\hat{\mathcal{C}}^{YX} > \hat{\mathcal{C}}^{XY}$ , which is confirmed by numerical studies [167]. An alternative way for understanding the observed asymmetry of the interdependent network characteristics is illustrated in Fig. ??: moderate unidirectional coupling (below the onset of synchronization) increases the driven system's dimension [109, 184] (we will numerically demonstrate this behavior in Section ??), so that former neighbors of pairs of recurrent states in  $X$  are not mutually close in  $Y$  anymore. In this case, the number of “cross-triangles” with baseline in  $Y$  decreases in comparison with those having their baseline in  $X$ . In fact, a corresponding decrease in  $\hat{\mathcal{T}}^{XY}$  ( $\hat{\mathcal{C}}^{XY}$ ) and an increase in  $\hat{\mathcal{T}}^{YX}$  ( $\hat{\mathcal{C}}^{YX}$ ) can often be observed in parallel (see Fig. ??).

Figure ?? shows the illustrative example of global cross-clustering coefficients for two unidirectionally coupled Rössler systems in the funnel regime with the same parameters  $a$ ,  $b$  and  $c$ , but a weak detuning of  $\nu = 0.02$ , following the setting of [167]. The obtained results are consistent with our above heuristic explanation for the emergence of asymmetries between the interdependent network characteristics in the presence of unidirectional coupling. Specifically, for a wide range of moderate coupling strengths, the difference between the two global cross-clustering coefficients allows to correctly identify the direction of the imposed coupling. At large coupling strengths (i.e., close to and beyond the onset of generalized synchronization, which is indicated by the second largest Lyapunov exponent of the system approaching zero as shown in Fig. ??), both global cross-clustering coefficients become statistically indistinguishable, which is consistent with the fact that the behavior of the driven system is completely locked to the dynamics of the driver (cf. Section ??). In turn, the indistinguishability of both coupling directions at very low coupling strengths is most likely due to the fact that the geometric deformations of the driven system's attractor are too small to be detected by the given finite values of  $\varepsilon_X$ ,  $\varepsilon_Y$  and  $\varepsilon_{XY}$  and the chosen network size. We expect that for larger IRNs and smaller distance thresholds, the lower boundary of the interval of coupling strengths for which the two global cross-clustering coefficients differ statistically significantly from each other will shift towards zero.

We emphasize that the same results can be obtained using the cross-transitivity replacing the global cross-clustering coefficient. Moreover, it is notable that the reported distinction can already be obtained at

comparably small network sizes of some hundred vertices [167].

### 3.9.5. Examples to characterize flow patterns

As one of successful applications of inter-system recurrence network approach as proposed in [166, 167], Gao *et al* characterize different oil-water flow patterns by reconstructing networks from multi-channel measurements [185–187]. In this series of works, Gao *et al* constructed multivariate RNs based on cross recurrence plots. In this example, we obtain experimental time series by four-sector conductance sensor, measuring the local flow behavior in the top, right, bottom, and left part of the horizontal pipe, respectively. Let us consider four dimensional multi-channel time series (after embedding in the same appropriate phase space) as  $\vec{M}_A, \vec{M}_B, \vec{M}_C$  and  $\vec{M}_D$ . Then the inter-system recurrence matrix of Eq. (71) is rewritten as

$$IR(\varepsilon) = \begin{pmatrix} \mathbf{R}^A(\varepsilon_{AA}) & \mathbf{CR}^{AB}(\varepsilon_{AB}) & \mathbf{CR}^{AC}(\varepsilon_{AC}) & \mathbf{CR}^{AD}(\varepsilon_{AD}) \\ \mathbf{CR}^{BA}(\varepsilon_{BA}) & \mathbf{R}^B(\varepsilon_B) & \mathbf{CR}^{BC}(\varepsilon_{BC}) & \mathbf{CR}^{BD}(\varepsilon_{BD}) \\ \mathbf{CR}^{CA}(\varepsilon_{CA}) & \mathbf{CR}^{BC}(\varepsilon_{BC}) & \mathbf{R}^C(\varepsilon_C) & \mathbf{CR}^{CD}(\varepsilon_{CD}) \\ \mathbf{CR}^{DA}(\varepsilon_{DA}) & \mathbf{CR}^{DB}(\varepsilon_{DB}) & \mathbf{CR}^{DC}(\varepsilon_{DC}) & \mathbf{R}^D(\varepsilon_D) \end{pmatrix}, \quad (83)$$

Note that both threshold values  $\varepsilon_A$  and  $\varepsilon_{AB}$  are chosen according to the rules of thumb as we discussed in Sec. 3.6.2. In order to consider  $IR$  as a network of networks, Gao *et al* [185] proposed a subjective criterion in choosing different threshold values to obtain the adjacency matrix of the multivariate networks. More specifically, two different thresholds are chosen such that the cross recurrence rate is significantly lower than that of the auto-recurrence. The so-obtained multivariate RNs are largely influenced by threshold values. One solution is that we consider each entry of  $IR$  as the connection weight of each link, which then quantifies the structural properties of the resulting weighted multivariate RNs [186].

## 3.10. Joint recurrence networks

### 3.10.1. Joint recurrence plots

Besides cross-recurrences, another possible multivariate generalization of RPs is studying joint recurrences of different systems in their individual (possibly different) phase spaces. Here, the basic idea is that the simultaneous occurrence of recurrences in two or more systems  $\{X_k\}$  (see Fig. ??) contains information on possible interrelationships between their respective dynamics, for example, the emergence of generalized synchronization [85, 108]. Consequently, based on time series  $\{x_i^k\}$ , the joint recurrence matrix  $\mathbf{JR}$  with elements

$$JR_{ij}(\varepsilon_1, \dots, \varepsilon_K) = \prod_{k=1}^K R_{ij}^k(\varepsilon_k) \quad (84)$$

is defined as the element-wise product of the single-system recurrence matrices  $\mathbf{R}^k$  with elements

$$R_{ij}^k(\varepsilon_k) = \Theta(\varepsilon_k - \|x_i^k - x_j^k\|), \quad (85)$$

where  $(\varepsilon_1, \dots, \varepsilon_K)$  is the vector of recurrence thresholds that can be selected for each time series individually, typically such as to yield the same global recurrence rates  $RR_k = RR$  for all  $k = 1, \dots, K$ .

### 3.10.2. Network interpretation

Analogously to single-system recurrence network analysis, we can take a graph-theoretical perspective by defining a *joint recurrence network* (*JRN*) by its adjacency matrix

$$\mathbf{A}(\varepsilon_1, \dots, \varepsilon_K) = \mathbf{JR}(\varepsilon_1, \dots, \varepsilon_K) - \mathbf{1}_N, \quad (86)$$

where  $\mathbf{1}_N$  again denotes the  $N$ -dimensional identity matrix. Hence, edges  $(i, j)$  of a JRN indicate joint recurrences occurring simultaneously in *all*  $K$  time series under study. Alternatively,  $\mathbf{A}(\varepsilon_1, \dots, \varepsilon_K)$  may be viewed as the element-wise product of the single-system recurrence networks' adjacency matrices  $\mathbf{A}^k(\varepsilon_k)$ .

As single-system RN and IRN, the JRN describes an undirected and unweighted simple graph. However, due to the temporal simultaneity condition of the joint recurrence concept, vertices  $i$  are explicitly associated

with points in time  $t_i^k = t_i^l$  common to the  $K$  considered time series (cf. Tab. 2). This is conceptually different from RNs and IRNs where time information is not taken into account so that network characteristics are invariant under permutations of the state vectors (i.e., the – possibly embedded – observations). More specifically, it is not possible to relabel the observations in the underlying time series prior to the computation of the JRN, whereas the JRN vertices can be shuffled again without altering the resulting network properties.

By construction, the time series  $\{x_i^k\}$  used for constructing a JRN need to be sampled at identical times  $\{t_i^k\}$  and have to have the same length, *i.e.*,  $N_1 = N_2 = \dots = N_K = N$ . However, since recurrences are compared instead of state vectors, the  $\{x_i^k\}$  neither have to represent the same physical quantity measured in identical units, nor need they reside in the same phase space (Tab. 2).

From a conceptual perspective, a JRN can be regarded as a simple RN for the combined system  $(X_1 \otimes \dots \otimes X_K)$  in its higher-dimensional phase space spanned by all state variables. However, recurrences are defined here in some non-standard way taking distances in the subspaces associated with the individual systems separately into account. This implies that the properties of JRNs can be studied in essentially the same way as those of single-system RNs (but with possibly more subtle geometric interpretations of the respective network characteristics). In turn, comparing the same properties for JRN and single-system RNs provides important information about the similarity of neighborhood relationships in the combined phase space and projections on the individual systems' subspaces. Specifically, we can gain insights about the effective degrees of freedom of the combined system, which may be reduced in comparison with the sum of the degrees of freedom of the uncoupled systems due to dynamical interdependences between its components. We will further detail this idea in Section ??.

*$\alpha$ -joint recurrence networks.* Equivalently to their interpretation outlined in Section ??, we can also consider JRNs as the reduction of a generalized graph, where the vertices correspond to time points  $t_i$ , which can be connected by at most  $K$  different types of (labelled) edges representing the mutual closeness of states in the  $K$  different systems. In this viewpoint, the reduction towards the JRN follows from the requirement that for a given pair of vertices, in the generalised graph *all*  $K$  possible labelled edges must be present. With other words, in terms of Boolean logics the entries of the binary recurrence matrices  $\mathbf{R}^k$  are connected by a logical AND for defining the elements of **JRN**.

Notably, the presence of a joint recurrence becomes increasingly unlikely as the number of possibly interacting systems  $K$  increases. Even in the case of very strong interdependences, there may be stochastic fluctuations in the individual systems (*e.g.*, observational noise) that mask recurrences in individual systems and, thus, subsequently reduce the *joint recurrence rate*

$$JRR(\varepsilon_1, \dots, \varepsilon_K) = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N JR_{ij}(\varepsilon_1, \dots, \varepsilon_K) \quad (87)$$

aka JRN edge density  $\rho_J$ .

One possibility to circumvent the problem sketched above is relaxing the requirement of having simultaneous recurrences in all systems (*i.e.*, the logical AND operation connecting the recurrence matrices of the individual systems in a component-wise way), but considering the case where at least a fraction  $\alpha \in (0, 1]$  of all systems exhibit recurrences (the standard JRN follows for  $\alpha = 1$ ). This point of view allows defining a hierarchy of networks, which we call  *$\alpha$ -joint recurrence networks* ( $\alpha$ -JRN). Starting from the union of the single-system RNs providing a network with  $K$  different edge types corresponding to recurrences of the individual systems, we require that there exist at least  $\lceil \alpha K \rceil$  edges between two specified vertices (*i.e.*, time points). In the specific case of  $K = 2$  systems and  $\alpha \in (0, 0.5]$  (*or*, more generally, for  $\alpha \in (0, 1/K]$ ), we can rewrite this requirement with a simple logical (Boolean) operation connecting the single-system recurrence matrices in a component-wise way as  $JR_{ij}^\alpha(\varepsilon_1, \varepsilon_2) = R_{ij}^{X_1}(\varepsilon_1) \text{ OR } R_{ij}^{X_2}(\varepsilon_2)$ .

For the more general case, in order to mathematically formulate the requirement of  $\lceil \alpha K \rceil$  simultaneous recurrences, it is convenient to start from a practically equivalent re-definition of the joint recurrence matrix,

$$JR_{ij}^*(\varepsilon_1, \dots, \varepsilon_K) = \Theta \left( \sum_{k=1}^K R_{ij}^k(\varepsilon_k) - K - \delta \right), \quad (88)$$

with the usual Heaviside function  $\Theta(\cdot)$  and  $\delta \rightarrow 0^+$  being infinitesimally small (to assure  $JR_{ij}^* = 1$  if  $\sum_{k=1}^K R_{ij}^k = K$ ), and set

$$JR_{ij}^\alpha(\varepsilon_1, \dots, \varepsilon_K) = \Theta\left(\sum_{k=1}^K R_{ij}^k(\varepsilon_k) - \alpha K - \delta\right), \quad (89)$$

to be the  $\alpha$ -joint recurrence matrix. We can use the latter definition to define  $\alpha$ -joint recurrence plots as well as  $\alpha$ -JRN in full analogy to the classical case  $\alpha = 1$ .

Trivially, the number of edges in an  $\alpha$ -JRN decreases monotonically for increasing  $\alpha$  if all single-system recurrence thresholds  $\varepsilon_k$  are kept fixed. We note that a similar relaxation of the strict requirement of a conjecture (AND relation) between the (Boolean) entries of different recurrence matrices has been recently discussed in the framework of symbolic recurrence plots [124]. Moreover, it might be interesting (but has not yet been explored) to use concepts from fuzzy logic as the basis for somewhat weaker requirements than in the rather restrictive definition of the original JRN.

The conceptual idea of  $\alpha$ -JRN has not yet been further developed and studied elsewhere. One possible field of application could be finding proper values of  $\alpha$  (for example, in dependence on the magnitude of some observational noise) for which results commonly obtained using “normal” JRN become stable in the case of real-world time series. To this end, we only emphasize the possibility of defining  $\alpha$ -JRN and studying the properties of these entities (e.g., the scaling of network characteristics as a function of  $\alpha$ ), but leave a corresponding investigation as a subject for future research.

### 3.10.3. Network properties and synchronization

The concept of joint recurrence plots (JRP) has been found very useful for studying the otherwise hard to detect emergence of generalized synchronization (GS) between two coupled chaotic systems  $X$  and  $Y$  [108]. GS describes the presence of a general functional relationship between the trajectories of both systems,  $y(t) = f(x(t))$ , which can arise at sufficiently large coupling strengths in both uni- and bidirectional coupling configurations. Most available methods for identifying GS from time series data have been developed for driver-response relationships, and only few approaches are also suitable for studying GS in the presence of symmetric couplings [168]. Among the latter, JRP have recently attracted specific interest.

Romano *et al.* [108] argued that in case of GS, recurrences in the two coupled systems need to occur simultaneously (or with a given fixed time lag in the special case of lag synchronization,  $y(t) = f(x(t - \tau))$ ). Hence, comparing the joint recurrence rate  $JRR$  with the recurrence rates of the individual single-system RPs (taken to be the same for both systems) should show convergence of both values. The latter fact is quantified in terms of the *joint probability of recurrence (JPR) index*

$$JPR = \max_\tau \frac{S(\tau) - RR}{1 - RR} \quad (90)$$

with the lagged joint recurrence rate ratio

$$S(\tau) = \frac{1}{N^2 RR} \sum_{i,j=1}^N \Theta(\varepsilon_X - \|x_i - x_j\|) \Theta(\varepsilon_Y - \|y_{i+\tau} - y_{j+\tau}\|) \quad (91)$$

and  $RR$  being the recurrence rate taken equal for both considered systems. Since for GS, we can expect that  $S(\tau) \rightarrow 1$  for some  $\tau$ ,  $JPR \rightarrow 1$ . However, the latter measure has some disadvantages. On the one hand, testing for significance of a specific value of  $JPR$  usually requires complex surrogate data approaches for properly approximating the distribution of the underlying null hypothesis (no synchronization) adapted to the specific time series under study [132]. On the other hand, comparing the single-system and joint recurrence rates may be insufficient since due to the complexity of fluctuations or the presence of stochastic components (observational noise), we hardly ever capture all single-system recurrence in the JRP. Consequently, a solely RR-based characterization does not necessarily lead to the expected “optimum” value of the synchronization index ( $JPR = 1$ ) in case of fully developed GS.

As an alternative, we have suggested that looking at higher-order characteristics (specifically, three-point instead of two-point relationships) may improve the results [168], especially when relying on probabilistic arguments. One convenient way is utilizing again the concept of transitivities from RN and JRN. The exploitation of alternative higher-order characteristics might be possible, but has not yet been explored. Joint recurrence networks can be analyzed by standard statistical measures from complex network theory [31, 142], which, however, need to be reinterpreted in terms of the underlying systems' joint recurrence structure [47, 166, 167]. Indeed, the transitivity properties of joint recurrence networks have been shown to reveal complex synchronization scenarios, notably including the detection of the onset of generalized synchronization, in coupled chaotic oscillators such as Rössler systems [167]. Notably, the specific requirements on the time series data render JRNs a promising approach for detecting intricate interconnections between qualitatively distinct observables in observational or experimental real-world data.

As a heuristic indicator for the presence of GS, we have proposed using the *transitivity ratio* [168]

$$\hat{Q}_{\mathcal{T}} = \frac{\hat{\mathcal{T}}^J}{(\hat{\mathcal{T}}^X + \hat{\mathcal{T}}^Y)/2}, \quad (92)$$

i.e., the ratio between the JRN transitivity and the arithmetic mean of the single-system RN transitivities. The rationale behind this definition is that for systems exhibiting GS, all degrees of freedom are completely locked, implying that both approach the same effective (fractal) dimension and should thus have the same RN transitivities, which approximately equal the JRN transitivity. Alternatively, we could also use other means of  $\hat{\mathcal{T}}^{X,Y}$ , such as the geometric or harmonic means, for obtaining a meaningful ratio. However, numerical experiments show that using the arithmetic mean provides values of  $\hat{Q}_{\mathcal{T}}$  that are mostly confined to the interval  $[0, 1]$  with only minor exceedances in the fully developed GS regime [168]. Since the arithmetic mean is always larger than the geometric one, normalizing with respect to the geometric mean  $\sqrt{\hat{\mathcal{T}}^X \hat{\mathcal{T}}^Y}$  would lead to higher values of  $\hat{Q}_{\mathcal{T}}$  and, hence, an even stronger violation of the desired normalization of the transitivity ratio. However, even when considering the normalization by the arithmetic mean of single-system RN transitivities, the thus defined transitivity ratio has two major drawbacks:

On the one hand, if the single-system RN transitivities are essentially different (a case that has not been studied in [168]), the contribution of the lower-dimensional system (higher transitivity) dominates the arithmetic mean in the denominator of Eq. (92) and, hence, the transitivity ratio itself irrespective of a possible well-defined driver-response relationship.

On the other hand, there is no rigorous theoretical justification for  $\hat{Q}_{\mathcal{T}}$  being a good indicator of GS. Notably, the definition of the transitivity ratio is based on the idea that the transitivities are related with the effective dimensions of the individual systems. In the uncoupled case, the degrees of freedom of both systems are independent; hence, the effective dimension of the composed system  $X \otimes Y$  just reads  $D^{X \otimes Y} = D^X + D^Y$  (notably, due to the logarithmic transform between RN transitivity and transitivity dimension, this additivity does *not* apply to the RN transitivities). In turn, in case of GS, the degrees of freedom of both systems become mutually locked, leading to  $D^{X \otimes Y} = D^X = D^Y$  (i.e., one system can be viewed as a – possibly nonlinear – projection of the other), with  $D^X$  and  $D^Y$  eventually differing from their values in the uncoupled case depending on the specific coupling configuration (e.g., uni- versus bidirectional coupling). Taking the estimated transitivity dimensions  $\hat{D}_{\mathcal{T}^{X,Y}}$  as proxies for  $D^{X,Y}$  and the *pseudo-dimension*  $\hat{D}_{\mathcal{T}^J} = \log(\hat{\mathcal{T}}^J) / \log(3/4)$  as an approximation of the true dimension  $D^{X \otimes Y}$  of the composed system  $X \otimes Y$ <sup>5</sup>, the latter case would translate into  $\hat{Q}_{\mathcal{T}} = 1$ , which is approximately attained in numerical studies for coupled Rössler systems in different dynamical regimes [168].

In order to circumvent both problems, we suggest here utilizing an alternative indicator, which is directly based on the concept of effective dimensions (degrees of freedom) of the individual systems. In analogy with the mutual information (sometimes also called redundancy [188, 189]) frequently used in nonlinear time

---

<sup>5</sup>In fact, we should take here the transitivity dimension of the RN obtained for  $X \otimes Y$ , i.e.,  $\hat{D}_{\mathcal{T}^{X \otimes Y}} = \log(\hat{\mathcal{T}}^{X \otimes Y}) / \log(3/4)$ , which is in general not identical to the pseudo-dimension  $\hat{D}_{\mathcal{T}^J}$  due to the different metrics used for the definition of recurrences of  $X \otimes Y$  and joint recurrences of  $X$  and  $Y$ .

series analysis, we define the *transitivity dimension redundancies*

$$\hat{D}_{\mathcal{T}^R} = \hat{D}_{\mathcal{T}^X} + \hat{D}_{\mathcal{T}^Y} - \hat{D}_{\mathcal{T}^J}, \quad (93)$$

$$\hat{D}_{\mathcal{T}^R} = \hat{D}_{\mathcal{T}^X} + \hat{D}_{\mathcal{T}^Y} - \hat{D}_{\mathcal{T}^{X \otimes Y}}, \quad (94)$$

which should assume zero values in the uncoupled case and exhibit  $\hat{D}_{\mathcal{T}^X} = \hat{D}_{\mathcal{T}^Y} = \hat{D}_{\mathcal{T}^{X \otimes Y}} = \hat{D}_{\mathcal{T}^J}$  in case of GS. In order to obtain a normalized measure for the presence of GS, we define the *dimensional locking index (DLI)*

$$\widetilde{\widetilde{DLI}} = \frac{\hat{D}_{\mathcal{T}^R}}{\hat{D}_{\mathcal{T}^J}}, \quad (95)$$

$$\widehat{DLI} = \frac{\hat{D}_{\mathcal{T}^R}}{\hat{D}_{\mathcal{T}^{X \otimes Y}}}. \quad (96)$$

Notably, this index is tailored to the dimensionality interpretation of RN transitivity. In a strict sense, this argument only applies if using the single-system RN transitivity (dimension) of the composed system  $X \otimes Y$  instead of the JRN transitivity (dimension)  $\hat{D}_{\mathcal{T}^J}$ . However, at this point, we suggest using the latter as an approximation. A detailed comparison between the two definitions will be subject to future research.

In order to further illustrate the behavior of the (J)RN-based characteristics for detecting the emergence of GS, we reconsider the example of two unidirectionally coupled identical but slightly detuned Rössler systems from Section ???. In contrast to [168], who studied different settings for uni- and bidirectional configurations with single realizations of the same system, we present here results obtained from ensembles of realizations. The results shown in Fig. ?? demonstrate that the estimated values of  $\mathcal{T}^J$  and  $\widetilde{\widetilde{DLI}}$  exhibit a marked increase at the onset of GS. Specifically, the new *DLI* index approaches one (with little overshooting) in the synchronized regime as expected, but takes values of only about 0.2 or lower in the non-synchronous case (in comparison with values of about 0.7 exhibited by  $Q_{\mathcal{T}}$ , cf. Fig. 2B in [168]).

As a second important observation, we find a systematic and significant decrease in the RN transitivity of the driven system at moderate coupling strengths before the onset of GS, which corresponds to an increase of the associated transitivity dimension. This behavior is precisely what was claimed in the context of coupling analysis in Section ?? for providing an explanation of the numerically observed asymmetry between the transitivity-based interdependent network characteristics. These results underline that some integrated utilization of single-system, inter-system and joint recurrence networks can eventually provide deep insights into the coupling regime and strength from bivariate observations.

### 3.11. Other proximity-based time series networks

#### 3.11.1. Cycle networks

Zhang *et al* [39] first suggested to study the topological features of pseudo-periodic time series by means of complex networks. Suppose that a dynamical system possesses pronounced oscillations (examples are the well-known Lorenz and Rössler systems). In this case, we identify the individual cycles contained in a time series of this system with the vertices of an undirected network. Edges between pairs of vertices are established if the corresponding segments of the trajectory behave very similarly. For quantifying the proximity of cycles in phase space, different measures have been proposed. In [190], Zhang *et al* introduced a generalization of the correlation coefficient applicable to cycles of possibly different lengths. Specifically, this correlation index is defined as the maximum of the cross correlation between the two signals when the shorter of both is slid relative to the longer one. That is, if the two cycles being compared are  $C_1 = \{x_1, x_2, \dots, x_\alpha\}$  and  $C_2 = \{y_1, y_2, \dots, y_\beta\}$  with (without loss of generality)  $\alpha \leq \beta$ , then we compute

$$\rho(C_1, C_2) = \max_{i=0, \dots, (\beta-\alpha)} \langle (x_1, x_2, \dots, x_\alpha), (y_{1+i}, y_{2+i}, \dots, y_{\alpha+i}) \rangle, \quad (97)$$

where  $\langle \cdot, \cdot \rangle$  denotes the standard correlation coefficient of two  $\alpha$ -dimensional vectors, and set

$$A_{i,j} = \Theta(\rho(C_i, C_j) - \rho_{max}) - \delta_{i,j}. \quad (98)$$

where  $\rho_{max}$  is a properly chosen threshold value and  $\delta_{i,j}$  is the Kronecker delta necessary in order to obtain a network without self-loops. As an alternative, the phase space distance [190]

$$D(C_1, C_2) = \min_{i=0, \dots, (\beta-\alpha)} \frac{1}{\alpha} \sum_{j=1}^{\alpha} \|x_j - y_{j+i}\| \quad (99)$$

has been suggested, leading to the following definition:

$$A_{i,j} = \Theta(D_{max} - D(C_i, C_j)) - \delta_{i,j}. \quad (100)$$

Of course, there are other calculations one could perform as well.

The advantage of cycle networks is that explicit time delay embedding is avoided. In addition, the method is more robust against additive noise, given a small enough noise magnitude to allow a clear identification of the individual cycles from the time series. Moreover, cycle networks are invariant under reordering of the cycles (this is precisely the same property that was also exploited for cycle-shuffled surrogate methods [191] but not the pseudo-periodic surrogate method [192]). However, for chaotic and nonlinear systems in a near-periodic regime, we typically observe significant orderly variation in the appearance of individual cycles. For systems that are linear or noise driven, that orderly variation will be less pronounced. As a consequence, the networks constructed with these methods will have characteristic and distinct properties: linear and periodic systems have cycle networks that appear randomly, while chaotic and nonlinear systems generate highly structured networks [39, 44]. Therefore, the vertex and edge properties of the resultant networks can be used to distinguish between distinct classes of dynamical systems. Moreover, in [190], authors used meso-scale properties of the networks — and in particular the clustering of vertices — to locate unstable periodic orbits (UPOs) within the system. This approach is feasible, since a chaotic system will exhibit a dense hierarchy of unstable periodic orbits, and these orbits act as accumulation points in the Poincaré section. Hence, the corresponding vertices form clusters in the cycle network.

For an implementation of the cycle network approach, the time series must be divided into distinct cycles. In [39, 71] the preferred method for defining cycles is splitting the trajectory at peaks (or equally troughs). In order to quantify the mutual proximity of different cycles, different measures can be applied depending on the specific application. On the one hand, the cycle correlation index  $\rho_{i,j}$  (Eq. (97)) can be properly estimated without additional phase space reconstruction (embedding), which has advantages when analyzing noisy and non-stationary time series, *e.g.*, experimental data [39]. Moreover, this choice effectively smoothes the effect of an additive independent and identically distributed noise source [190]. On the other hand, the phase space distance  $D_{i,j}$  (Eq. (99)) is physically more meaningful [71]. For the example systems as well as some real-world clinical electrocardiogram recordings studied in [39, 71], both methods have been found to perform reasonably well. However, whether the previously considered approaches also lead to feasible results for other cases has to be further investigated in future research.

In general, the construction and quantitative analysis of cycle networks requires a sufficiently high sampling rate, *i.e.*, we require that both cycle lengths  $\alpha$  and  $\beta$  in Eqs. (97) and (99) are reasonably large. The main reason for this requirement is that even two cycles that are fully identical but sampled in a different way may have rather different cycle correlation indices (and phase space distances) depending on the exact values of the observed quantity. Hence, for a very coarse sampling, it is possible that two cycles that are actually close in phase space may not be connected in the cycle network. However, for large sampling rates, the variance of this measure decreases, resulting in a more reliable network reconstruction.

Instead of computing correlation coefficients to quantifying linear correlation between two cycles, the mutual information is proposed to capture nonlinear effects in a time series and, hence, provides more accurate estimates of the similarity of nonlinear time series intervals. Also, this builds directly on results obtained from other fields in which estimates of mutual information values have been used to infer causal gene networks. Second, in [193] the authors defined a node in the constructed network as an episode. An episode is a temporal interval of the time series that may consist of  $n_e \geq 1$  consecutive cycles. That means an episode is  $n_e$  times longer than a cycle. The extended length of an episode, compared to a cycle, has the advantage of increasing the accuracy of the statistical estimates of the mutual information value. The

reason for this is that a cycle does not need to have a certain minimal length to qualify as a cycle. However, it is clear that very short cycles convey less information about the time series than long cycles. Due to the fact that the notion of a “cycle” is parameter free, one cannot adjust for this shortcoming. For this reason we extend the principle idea behind the usage of a cycle in the construction of a network [39] by means of an episode. Third, our network construction model is a parametric method because an episode is a function of  $n_e$ , the number of consecutive cycles. This gives us a parameter that can be optimized to result in the “best” network for a given time series. The so-obtained optimal value of  $n_e$  provides a procedure to estimate the best network representation.

The choice of the threshold  $\rho_{max}$  influences the link density of the resulting network, which could be discussed in a similar framework as constructing recurrence networks. One solution is to show the dependence of network characteristics on  $\rho_{max}$  explicitly [39]. Furthermore, Zhang *et al* constructed cycle networks for sinus rhythm electrocardiogram recordings of the coronary care unit patients and healthy volunteers. application of cycle networks. It has been demonstrated that the degree distributions of the resulted networks show more prominent fluctuations in comparison to that of the healthy volunteers which vary rather smoothly. Other network measures including clustering coefficients and the average path length also show significant difference between the healthy and the coronary care patients. Furthermore, cycle network has been applied to characterizing electrical signals of acupuncture [194], showing different network topologies when the control parameter is in different regime, for instance, either twisting or lifting and thrusting conditions.

Most of the vertex and edge properties of cycle networks have been explained by unstable periodic orbits of the underlying chaotic systems [39]. Chaotic dynamical systems possess infinitely many unstable periodic orbits (UPOs) embedded in a chaotic set, which play an important role in characterizing properties of chaos, such as, Lyapunov exponents, fractal dimensions, and bifurcation structures. In [195], Kobayashi *et al* performed a network analysis of UPOs. By means of the Poincaré map, they first numerically extract a large number of UPOs, which are considered as vertices of the network. Note that most of the existing algorithms can only detect UPOs of lower orders, which have been sufficient for characterizing the properties of the underlying chaotic system [196, 197]. The edges between two UPOs are established by a transition process of a typical chaotic orbit. More specifically, when a typical orbit  $\{x_n\}$  travels close to UPO<sub>i</sub> at time  $n$  and later shifts to the neighborhood of UPO<sub>j</sub> at time  $n + 1$ , we build a connection between these two UPOs. Since the chaotic nature of the typical orbit, the transitions between different UPOs are irregular. The resulting network presents small world and scale free features, which confirm those results that have been reported in [39].

### 3.11.2. Correlation networks

By embedding an arbitrary time series, individual state vectors  $\vec{x}_i$  in the  $m$ -dimensional phase space of the embedded variables can be considered as vertices of an undirected complex network. Specifically, if the Pearson correlation coefficient

$$r_{i,j} = \langle \vec{x}_i, \vec{x}_j \rangle = \frac{\sum_{\alpha=1}^m \left( x_i^{(\alpha)} - \frac{1}{m} \sum_{\beta=1}^m x_i^{(\beta)} \right) \left( x_j^{(\alpha)} - \frac{1}{m} \sum_{\beta=1}^m x_j^{(\beta)} \right)}{\sqrt{\sum_{\alpha=1}^m \left( x_i^{(\alpha)} - \frac{1}{m} \sum_{\beta=1}^m x_i^{(\beta)} \right)^2 \sum_{\alpha=1}^m \left( x_j^{(\alpha)} - \frac{1}{m} \sum_{\beta=1}^m x_j^{(\beta)} \right)^2}} \quad (101)$$

is larger than a given threshold  $r$ , the vertices  $i$  and  $j$  are considered to be connected [45, 75]:

$$A_{i,j} = \Theta(r - r_{i,j}) - \delta_{i,j}. \quad (102)$$

Interpreting  $1 - r_{i,j}$  as a proximity measure, the condition  $r_{i,j} \geq r$  corresponds to the definition (Eq. (16)) of a recurrence with  $\varepsilon = 1 - r$ . The consideration of correlation coefficients between two phase space vectors usually requires a sufficiently large embedding dimension  $m$  for a proper estimation of  $r_{i,j}$ . This high embedding  $m$  often includes several oscillation periods as compared to cycle network. Hence, information about the short-term dynamics might get lost. Moreover, since embedding is known to induce spurious correlations [198], the results of the correlation method of network construction may suffer from related effects. The

correlation network method has been applied to stock price series, showing Gaussian distributions for the degree sequences that are reconstructed from return and amplitude series [45]. Furthermore, different two phase gas-liquid flow patterns have been well characterized by correlation network approaches [75].

The statistical concerns of the Pearson correlation coefficients arise when smaller values of  $m$  are used, say  $m = 10$ , which requires more statistical robust measures. In [199], Hou *et al* proposed to use the inner composition alignment (IOTA) to quantify the connectivity strength between two embedded vectors, which is a permutation based measure that was originally introduced to identify couplings from rather short gene-expression data [200]. Comparing to the standard symmetric undirected correlation network, a directed correlation network is obtained which has been further applied to characterize the pathological changes in the cardiovascular system from short-term heartbeat time series [199].

## 4. Visibility graphs (Jonathan)

### 4.1. Historical roots

Among other approaches, the so-called visibility graph (VG) has attracted considerable interest. Originally, this concept has been introduced for the analysis of mutual visibility relationships between points and obstacles in two-dimensional landscapes in the framework of computational geometry, with applications ranging from robot motion planning to architectural design and topographic descriptions of geo-graphical space [201–204]. Lacasa *et al.*[41] adopted the VG approach to the analysis of structures in scalar, univariate time series. It has been shown that certain statistical features of the resulting complex networks are closely related with fractal and multi-fractal properties of the underlying time series [48, 49]. A simplifying methodological variant called horizontal visibility graphs has been proposed (HVGs, see [46, 205–207]).

A mini review of (H)VGs has been presented in [208, 209], in particular, the application of this approach to geophysical processes has been reported in [47], which link the complete variety of different network properties describing the structure of VGs with specific structural features of geophysical processes in some more detail. Here we summarize the recent developments of the method and discuss some practical issues which pose considerable challenges to VG analysis of geophysical data, such as missing data, homo- and heteroscedastic uncertainty of observations, and time-scale uncertainty. We also discuss some successful applications to testing time irreversibility of nonlinear time series.

### 4.2. Algorithmic variants

In (H)VG analysis, individual observations are considered as vertices. For instance, given a univariate time series  $x(t_i)_{i=1,\dots,N}$ , the binary adjacency matrix  $A$  has dimension  $N \times N$ . Depending on the particular visibility conditions in defining the edges of the resulted graph, we have several different versions of VGs.

#### 4.2.1. Natural visibility graphs

First, in the framework of the standard visibility graph (VG), the non-zero entries of  $A_{i,j}$  correspond to two time points  $t_i$  and  $t_j$  are mutually connected vertices if the following criterion

$$\frac{x(t_i) - x(t_k)}{t_k - t_i} > \frac{x(t_i) - x(t_j)}{t_j - t_i} \quad (103)$$

is fulfilled for all time points  $t_k$  with  $t_i < t_k < t_j$  [41]. Therefore, the edges of the network take into account the temporal information explicitly. In Fig. 6(a), we show the algorithm of constructing VGs for almost periodic sunspot series. More detailed discussions on VGs analysis for sunspot series will be reviewed in Sec. 6.2. By default, two consecutive observations are connected and the graph forms a completely connected component without disjoint subgraphs. There are pronounced boundary effects, for instance, the first time point can only be visible to points that are in the future time axis (Fig. 6(a)). Furthermore, the VG is not affected by choice of algorithmic parameters – in contrast to most other methods of constructing complex networks from time series data are dependent on the choice of some parameters (*e.g.* the threshold  $\varepsilon$  of recurrence networks, see more details in [43]).

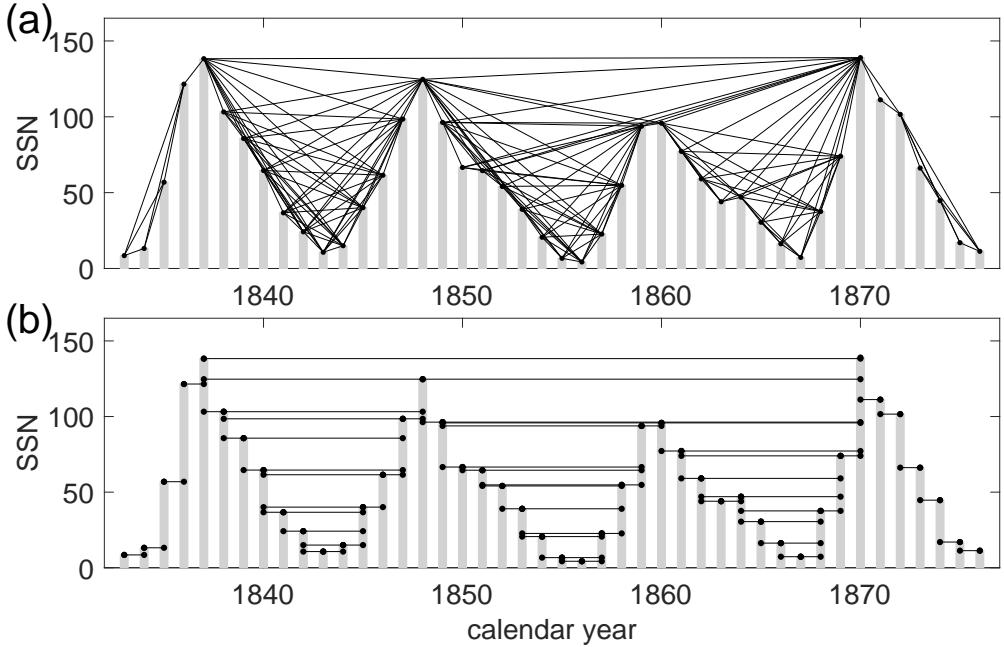


Figure 6: Algorithm of constructing natural visibility graphs (a), and (b) horizontal VG for time series of sunspot number series. Reproduced from [210].

#### 4.2.2. Horizontal visibility graphs

As a notable modification of the standard VG algorithm, Luque *et al.* [46] proposed utilizing a simplified criterion of horizontal visibility for transforming a time series into a complex network. Specifically, they considered two observations made at times  $t_i$  and  $t_j$  to be connected in a horizontal visibility graph (HVG) if and only if

$$x(t_k) < \min(x(t_i), x(t_j)) \quad (104)$$

for all  $t_k$  with  $t_i < t_k < t_j$ .

The algorithmic difference between HVG and VG is illustrated in Fig. 6(b). Note that the geometric criterion defined for the HVG algorithm is more ‘‘visibility restrictive’’ than its analogous for the standard VG. That is to say, the nodes within the HVG will have ‘‘less visibility’’ than their counterparts within the VG. It is easily seen that the edge set of the HVG associated with a given time series is a subset of the edge set of the associated VG (i.e., if the horizontal visibility criterion Eq. (104) is fulfilled, then also Eq. (103) holds, but not necessarily vice versa). In addition, VGs are invariant under affine transformations of the entire time series, whereas HVGs are not. One notable advantage of HVGs is that they provide an even higher degree of algorithmic simplicity than standard VGs, resulting in the observation that for certain simple stochastic processes and the quasiperiodic transition route to chaos, some basic graph properties can be calculated analytically [46, 211, 212]. On the other hand, the fact that HVGs typically contain a lower number of edges increases the demands regarding the time series length relatively to the standard VG when using this approach in applications such as tests for time-reversal asymmetry [50].

#### 4.2.3. Other variants of (H)VG

Given the definitions of VG and HVG, the resulting graphs are undirected, unweighted. One straightforward generalization of (H)VGs to directional (H)VGs is to introduce directed edges between vertices, i.e., from the cause at  $t_i$  to the effect at  $t_j > t_i$ . As it will be shown in Sec. 4.5, such directed graphs will give information on time-reversal asymmetry of the considered time series since the connection from  $t_i$  to  $t_j$  does not necessarily coincide with the connection from  $t_j$  to  $t_i$ .

There are some further variants of (H)VG depending on the particular properties of the given time series. For instance, given a binary series, in [213], a simplified VG has been obtained for binary VG and the visibility condition (expression (103)) is reduced to  $x(t_i) + x(t_j) > x(t_k)$  that  $x(t_k) = 0$  for all  $t_k$  such that  $t_i < t_k < t_j$ . The resulting VG from binary series is always connected and undirected, which is more solvable than the standard VGs.

Parametric VG has been proposed in [214, 215], introducing "view angle"  $\alpha$ . When the angle  $\alpha = \pi$ , the parametric version VG and the standard VG are the same. However, the angle  $\alpha = \pi/2$  does not turn into HVG because  $\alpha$  introduces a direction of links in the resulted graphs. The next step is to study the dependence of network structural measures on the parameter  $\alpha$ .

Multiscale limited penetrable HVG algorithm has been proposed in [216, 217], which can be regarded as a continuous construction of HVG based on a proper coarse grained time series. In the original HVG, two time points  $t_i$  and  $t_j$  are connected if no other intermediate points  $t_k$  are larger than  $\min(x(t_i), x(t_j))$ , namely,  $t_k < \min(x(t_i), x(t_j)), i < k < j$ . Now we use a less restrictive criterion, allowing one of  $t_k$  being larger than  $\min(x(t_i), x(t_j))$ , as represented by the new parameter of  $L = 1$ . We allow two points of  $t_k$  larger than  $\min(x(t_i), x(t_j))$  if  $L = 2$ . When increasing  $L$ , there are more edges in the resulting HVG comparing to the standard HVG. The standard VG is recovered from HVG if  $L \rightarrow \infty$ . Furthermore, the limited penetrable VG algorithm has been combined with parametric VG in [218].

The extension of (H)VGs from a univariate time series to scalar fields has been recently reported in [219, 220], which is conceptually closer to the original idea of visibility graphs. In addition, one may reconstruct (H)VGs for a set of ordered data (either descending or increasing ordered) [221].

It is also possible to combine the concepts of VGs with transition networks (Sec. 5), for instance, the so-called visibility graphlet approach has been proposed in [222, 223]. In this case, VGs are constructed respectively for sliding windows and each window is regarded as a network node (phase space partition). The connection between two nodes are decided by the temporal succession. The resulting transition network shows distinct topological features for chaotic series [223].

In all above cases, (H)VGs are reconstructed from a given time series. In [224], *Tsiotas et al.* expanded the VG algorithm to analyze node's attributes of a given graph. More specifically, let us consider a graph  $G(V, E)$  and a node-wise attribute  $Y$  that might be any of the node-wise network measures, for instance, local clustering coefficient or betweenness centralities. The secondary VG analysis for node-wise attributes  $Y$  shows specific ability in pattern recognitions [224].

Other properties of combinatorics have been used to characterize HVGs successfully in [207]. Recently, some analytic results have been obtained for some independent and identically distributed random noise, which has exponential degree distribution [225].

The time complexity of the basic natural VG algorithm is  $O(N^2)$ , which means that it takes a lot of time when dealing with long time series. A faster transform algorithm has been proposed in [226] to reduce the computation time, showing much efficient time complexity  $O(N \log N)$ . Note that for the HVG algorithm, it is not possible to further improve the computation efficiency because the computation complexity has reached the lower bound  $O(N)$ .

### 4.3. Visibility graph properties

#### 4.3.1. Degree distributions

Recent work on (H)VGs has mainly concentrated on the properties of the degree distribution  $p(k)$  resulting from different kind of processes. Specifically, VGs obtained from periodic signals appear as a concatenation of a finite number of network motifs (given that the basic period is an integer multiple of the sampling rate), i.e., have a regular structure with only a few distinct values of the vertex degree. The opposite extreme case, white noise, yields VGs appearing as exponential random graphs, i.e., random networks characterized by an exponential degree distribution. For example, exponential degree distribution has been tested in wind speed records measured in central Argentina [227].

In fractal processes, numerical results suggest that  $p(k)$  exhibits a power law [41],  $p(k) \sim k^{-\gamma}$ . Taking this empirical observation, VG analysis has been suggested to characterize fractional Brownian motions and  $f^\beta$ -noise, finding some heuristic relationship between  $\gamma$  and the process Hurst exponent  $H$  as  $\gamma = 3 - 2H$ ,

and  $\gamma = 5 - 2H$  for fractional Gaussian noise [48, 49]. Depending on the fractal properties of the underlying process, recently an algorithm for constructing VGs from segmented time series has been proposed [228], which estimates power-law exponents reflecting scale-free properties quite well. The idea hinges on a proper choice of the time delay  $\tau$ , which re-samples the original time series resulting in a number of segmentations. However, the number of segments to be used is often not known in the original algorithm. In other words, we do not know the upper bound of  $\tau_{max}$  to terminate the computation. So far, only a heuristic choice of  $\tau_{max}$  has been suggested when network characteristics of segmented time series show more or less convergent behavior [228]. This improved algorithm has been applied to diagnose Autism spectrum disorders [228]. Since there are much concerns regarding the statistical justification of the power laws of VGs, one can directly analyze the degree sequence (instead of the distribution) by detrended fluctuation analysis [229], which quantifies multifractal properties better than the standard VGs analysis.

Some exact results of  $p(k)$  of the HVG associated with generic uncorrelated random series have been obtained in [46]. More specifically, for a bi-infinite time series created from a random variable  $X$  with probability distribution  $f(x)$  with  $x \in [0, 1]$ , it has been proved that the degree distribution of the graph has an exponential form

$$p(k) = \frac{1}{3} \left(\frac{2}{3}\right)^{k-2}, \quad (105)$$

where  $k$  is the degree. Interestingly, every probability distribution  $f(x)$  of uncorrelated random series have the same exponential form. Numerical results of  $f(x)$  from a uniform, a Gaussian and a power law form (e.g.,  $f(x) \sim x^{-2}$ ) show perfect agreements with this theoretical predictions [46]. A general diagrammatic theory has been proposed in [230] to compute  $p(k)$  for any given dynamical process with well defined invariant measure. Taking into account the time information explicitly as in so-called directed HVG (as will be explained below), the outgoing degree distribution  $p_{out}(k) = (1/2)^k$ . Further solvable example are Markovian processes with an integrable invariant measure  $f(x)$ , for instance, the stationary Ornstein-Uhlenbeck process, one dimensional chaotic and quasiperiodic maps of with smooth invariant measure. In addition, the mean degree  $\bar{k}$  of the HVG associated to an uncorrelated random process is then given by

$$\bar{k} = \sum kp(k) = \sum_{k=2}^{\infty} \frac{k}{3} \left(\frac{2}{3}\right)^{k-2} = 4. \quad (106)$$

For an infinite periodic series of period  $T$ , the mean degree  $\bar{k}$  is  $\bar{k} = 4(1 - \frac{1}{2T})$ .

Nevertheless, similar to VGs, HVGs have been successfully applied to studying time series from various fields of sciences. Within the area of the present manuscript, we particularly notice the recent paper by [231] who studied the multifractal properties of some solar flare index in terms of HVG characteristics. The properties of HVG have been tested in river flows [232], showing exponential degree distributions.

In order to understand the hypothetical scale-free property of  $p(k)$  for VGs from fractal records, one has to note that typically, maxima of the time series have visibility contact with more other vertices than other points, i.e., hubs of the network often form at maximum values of the recorded observable. Put it differently, the degree of a vertex in the VG characterizes the maximality property of the corresponding observation in comparison with its neighborhood in the time series. Although local large time series values have better visibility than other small values, hub nodes of large degrees of VGs do not necessarily correspond to higher values, especially when there are some sort of periodic trends in the given data sequence, for instance wind speed records [227, 233]. Therefore, the relationship between maxima time series points and hubs of VGs is not completely general, since there can be specific conditions (e.g., a concave behavior over a certain period of time) which can lead to highly connected vertices that do not coincide with local maxima, for example, in case of a Conway series [41].

In addition, minima of time series provide complementary insight for the understanding of the particular process, for instance, sun spot series [233]. In the standard VGs, the contributions of local minimum values have been somehow largely overlooked by degree distribution  $p(k)$  because minimum values are basically non-hubs. One simple solution is to study the negatively inverted counterpart of the original time series, namely,  $-x(t_i)$ , which quantifies the properties of the local minima [233]. For convenience, we use  $k_{-x}$  and  $p(k_{-x})$  to denote degree sequence and distribution of VG resulted from  $-x(t_i)$ . Here, we remark that this

simple inversion of the time series allows us to create an entirely different complex network. This technique has demonstrated to be useful to understand the long-term behavior of strong minima of the solar cycles [233]. We will review these results in Sec. 6.2.1.

Based on degree distribution  $p(k)$ , we can calculate the graph entropy  $h = -\sum_k p(k) \log p(k)$ , which is used as the approximation to the Shannon entropy  $H$  of the corresponding time series  $x(t)$  [209, 234]. Furthermore, the VG aggregation operator has been proposed in [235, 236]. This operator includes the temporal information in the weights of the aggregation, showing computational simplicity comparing to other traditional aggregation operators.

#### 4.3.2. Stochastic vs. deterministic dynamics

Concerning the HVG, exponential functional forms have been obtained for many random processes, namely,  $p(k) \sim e^{-\lambda k}$ . A scaling factor of  $\lambda_c = \ln(3/2)$  has been found in the case of uncorrelated noise (white noise), which has been further proposed to separate stochastic from chaotic dynamics in the following senses [205, 230, 237]: (i) correlated stochastic series are characterized by  $\lambda > \lambda_c$ , slowly tending to an asymptotic value of  $\ln(3/2)$  for very weak correlations, whereas (ii) chaotic series are often characterized by  $\lambda_{chaos} < \lambda_c$  for decreasing correlations or increasing chaos dimensionality, respectively [205]. In [238], we have provided some further examples supporting argument (i). Meanwhile, we showed some peculiar results indicating that  $\lambda_c$  should not be interpreted as a general critical value separating chaos from noise.

Let us focus on applying (H)VG analysis to auto-regressive (AR) stochastic processes, which often describe certain time-varying processes in nature, economics, etc. The AR model specifies that the output variable depends linearly on its own previous values and on a stochastic term. More specifically,  $\mathbf{x} = [x_1, x_2, \dots, x_i, \dots, i \in \mathbb{Z}]$  is an AR model of order  $p$  denoted as AR( $p$ ) if

$$x_t = \sum_{j=1}^p \varphi_j x_{t-j} + \varepsilon_t, \quad (107)$$

where  $\varphi_j, j \in [1, p]$ , are real-valued coefficients of the model, and  $\varepsilon_t$  is white noise. We further assume that the error terms  $\varepsilon_t$  follow a Gaussian distribution with zero mean and unit variance. Specifically, we perform both VG and HVG analysis for AR(1), namely, (i)  $|\varphi_1| < 1$  for the AR(1) model. It is known that  $\varphi_1 > 0$  corresponds to positive correlation and the correlation length increases when  $\varphi_1$  is increased from 0 to 1. In contrast, anti-correlation is observed for negative coefficient  $\varphi_1$ . Similar H(VG) analysis for the AR(2) model have reported in [238].

In the case of  $\varphi_1 > 0$ , we find that  $p(k)$  approximately follows an exponential distribution. To illustrate this finding, the cumulative degree distributions  $F(k)$  for  $\varphi_1 = 0.3, 0.9$  and  $-0.5$  are shown in Fig. 7A and B, where clear scaling regimes are present in the semi-log plots. Furthermore, when increasing  $\varphi_1$ , in the VG, the exponent  $\lambda$  shows a monotonically decreasing trend (Fig. 7C). In contrast, the value  $\lambda$  for the HVG is increased (Fig. 7D). The result of Fig. 7D confirms the hypothesis stated in [205] the all  $\lambda$  should be larger than  $\lambda_c = \ln(3/2)$  as the correlation length is increased in the case of positively correlated increments.

In turn, when  $\varphi_1 < 0$ , we observe some peculiar results that seem to contradict the hypothesis stated in [205]. According to this hypothesis,  $\lambda$  should be larger than  $\lambda_c$  ( $\lambda > \lambda_c$ ) in stochastic processes in contrast to  $\lambda < \lambda_c$  for chaotic maps. The results of Fig. 7D do not support this claim when  $\varphi_1$  is negative in the AR(1) model. Instead, we find a region where the slope of the exponential degree distribution is smaller than  $\ln(3/2)$  (as highlighted in Fig. 7D). This suggests that the critical value of  $\ln(3/2)$  should not be understood as a general law of separating correlated stochastic from chaotic processes, which requires further investigation.

Working with correlated stochastic time series, further results in [239] do not adequately support the arguments of exponential degree distributions as reported in [205]. More specifically, they have constructed (H)VGs for fractional time series with three different methods, a generic  $1 = 1/f^\beta$  noise with Fourier filtering method, a deterministic fBm process of Weierstrass-Mandelbrot function, and a stochastic fBm process with successive random addition method. The numerical analysis show that VG algorithm may not be a well-defined method to extract correlation information of a time series and its statistics is not essentially the

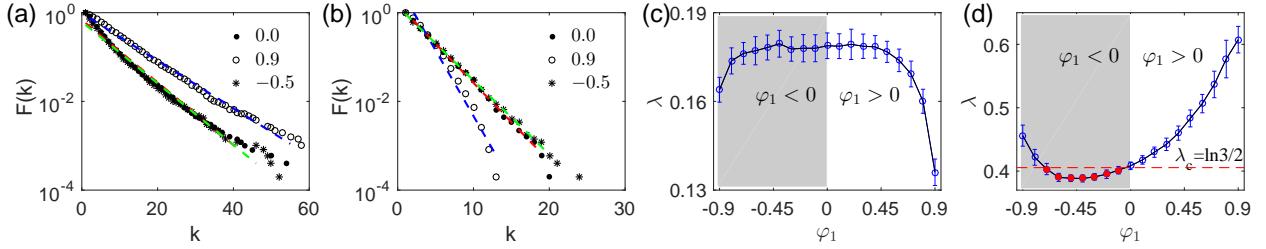


Figure 7: (color online) (A, B) Estimates of  $\lambda$  for approximately exponential degree distributions of AR(1) process. (C, D)  $\lambda$  versus  $\varphi_1$ . (A, C) VG, and (B, D) HVG. When  $\varphi_1 > 0$ ,  $\lambda$  has a decreasing trend in the VG, while in the HVG,  $\lambda$  rises when the correlation length increases. Each dot in panels C and D represents an average over 50 independent random realizations of 5000 data points. In (D),  $\lambda$  values smaller than  $\ln 3/2$  are highlighted by red color.

same as that of the HVG. The degree distributions of HVGs are shown to have parabolic exponential forms with Hurst dependent fitting parameter [239]. Therefore, we conclude that hypothesis of [205, 230, 237] needs careful numerical simulations for proper interpretations.

#### 4.3.3. Local network properties

Here, we consider other local network properties besides degrees and their distributions. In the case of VGs, the local clustering coefficient  $C_i$  and its relationship with the degree  $k_i$  have been numerically studied recently in human heartbeat data [240]. Particularly, it has been observed that  $C(k) \sim k^{-\gamma}$  and  $\gamma = 1$ , pointing to a hierarchical organization of the network [29], since vertices  $i$  with high  $C_i$  and low  $k_i$  (which are most abundant) form densely connected subgraphs, indicating a strong modular structure of the VG.

In the case of HVG associated with an uncorrelated random series, the local clustering coefficient  $C_i$  can be easily deduced by means of geometrical arguments. For a given node  $i$ ,  $C_i$  denotes the rate of nodes connected to  $i$  that are connected between each other. In other words, we have to calculate from a given node  $i$  how many nodes from those visible to  $i$  have mutual visibility (triangles), normalized with the set of possible triangles  $\binom{k}{2}$ , where  $k$  is the degree of node  $i$ . Based on a general rule between degree  $k$  and local clustering coefficient

$$C_i(k) = \frac{k-1}{\binom{k}{2}} = \frac{2}{k}, \quad (108)$$

one obtains the local clustering coefficient distribution  $p(C)$  as

$$p(C) = \frac{1}{3} \left(\frac{2}{3}\right)^{2/C-2}. \quad (109)$$

The above theoretical result has been confirmed by uncorrelated random series [46]. In addition, for HVG of a binary sequence,  $p(C)$  has a simplified expression as reported in [213].

In many cases local maxima of the underlying time series are expected to have large values of betweenness because high values often correspond to hubs in VGs which separate different parts of the series without mutual visibility contact and, thus, act as bottlenecks in the network structure, bundling a large number of shortest paths between vertices at  $t < t_i$  and  $t > t_i$ , respectively. However, in contrast to the degree, betweenness is additionally affected by the vertex' position in the underlying time series due to a simple combinatorial effect: Considering that the majority of shortest paths that cross a vertex  $i$  connect observations before and after  $i$  with each other, there are more possible combinations of such points for  $i$  being close to the middle of the time series than for vertices close to the edges of the record. In this respect, in a VG betweenness centrality of a vertex mixes information on the local maximality of the corresponding observation and its position within the time series.

In the case of closeness centrality the position of a vertex in the time series is even more important in comparison with the value of the underlying observable. Specifically, this measure is strongly determined by the number of vertices to its left and right, respectively. In this spirit, it can be argued that in the middle of the time series, high closeness values are more likely than at its ends. As argued above, a similar (but weaker)

effect contributes to betweenness and - close to the edges of the record - also to the degree (consequently, the highest degree and betweenness values can be taken by other vertices than that corresponding to the global maximum). In contrast, the local clustering coefficient is almost unaffected except for vertices very close to the beginning and end of the time series, since direct connectivity is mainly established between vertices that correspond to observations that are not very distant in time.

We note that boundary effects play an important role in the computation of the above centrality measures as that has been numerically reported in [47]. In particular, it has been shown that the local clustering coefficient is almost unaffected, but the impact of boundaries on the estimated vertex properties is particularly strong for short records, which are typical in geophysical applications. Specifically, degree and other centrality properties of observations close to both ends of a time series are systematically underestimated, which may artificially alter the interpretation of the corresponding results in their geophysical context. Hence, a careful treatment and interpretation of the results of VG analysis is necessary in such cases [47].

#### 4.3.4. Global network properties

In contrast to other approaches to complex network based time series analysis, in a VG the edge density  $\rho$  is a true network characteristic rather than a parameter of the method (cf. [38, 43]) for the corresponding meaning for recurrence networks). Specifically, a maximum edge density of 1 would be present if the underlying time series is globally convex (e.g., of regular parabolic shape), whereas low values indicate a strong fragmentation of the VG and, hence, irregularity of fluctuations of the underlying observable.

For a holistic characterization of a graph,  $C$  and  $\mathcal{L}$  have attracted particular interest, since their common behavior gives rise to a mathematical evaluation of the small-world phenomenon, i.e., the emergence of real-world networks with a high degree of clustering and a short average path length [63].

The topological properties of HVGs constructed from fractional Brownian motions with different Hurst indexes  $H \in (0, 1)$  have been reported in [206]. It is found that the clustering coefficient  $C$  decreases when  $H$  increases.

It can be expected that the value of  $\mathcal{L}$  is large when there are only few edges in the VG (low edge density) and low for a high edge density. Hence, average path length and edge density capture essentially similar properties of the underlying time series. For uncorrelated random series, the mean path length  $\mathcal{L}$  of HVG has a logarithmic scaling relationship with the length of time series  $N$ , in particular,  $\mathcal{L}(N) = 2\ln(N) + 2(\gamma - 1) + O(1/N)$ , where  $\gamma$  is the Euler-Mascheroni constant [46]. In the case of fractional Brownian motions with different Hurst indexes  $H \in (0, 1)$ , and for fixed length of time series of  $N$  points,  $\mathcal{L}$  increases exponentially with  $H$ . In addition,  $\mathcal{L}$  increases linearly with respect to  $N$  when  $H$  is close to 1 and in a logarithmic form when  $H$  is close to 0 [206].

Besides studies on the small-world effect, the assortativity of VGs has recently attracted considerable interest. Specifically, the presence of assortative behavior implies so-called hub attraction, whereas disassortative behavior relates to hub repulsion. It has been shown that the latter is a necessary condition for the emergence of fractal structures in networks [241]. For example, for Brownian motion (a fractal stochastic process), hub repulsion is not present, and the resulting VGs are non-fractal, but show a scaling of the average path length with increasing net-work size as  $\mathcal{L}(N) \sim \log N$ , which is typical for small-world networks. In contrast, for the Conway series (a deterministic fractal) one finds hub repulsion and  $\mathcal{L} \sim N^{-\beta}$ , which implies the presence of a fractal VG (Lacasa et al. 2008). In this respect, the assortativity coefficient or, more specifically, the scaling of the degree correlation determines the fractality of a VG [241], which is an interesting and potentially relevant property when studying fractal time series.

Of course, beyond the aforementioned characteristics there are multiple other measures one could also consider for describing the properties of VGs. This includes also measures characterizing the properties of individual edges as well as the distributions of small subgraphs (motifs). For example, the consideration of four-node subgraphs helps to show different dominant motifs ranks in the VGs of human ventricular time series, which distinguishes ventricular fibrillations from normal sinus rhythms of a subject [242]. Furthermore, the profiles of sequential  $n$ -node motifs of (H)VGs appear with characteristic frequencies which have been computed analytically for certain deterministic and stochastic dynamics [243].

#### 4.3.5. Practical considerations

Many recent publications on (H)VG analysis of time series have particularly made use of data from model systems, which are characterized by rather ideal conditions for statistical analysis. Even for most practical applications presented so far, the properties of the data under study have allowed using this methodological approach without extensive precautions. However, when operating with data obtained in a geophysical context, features challenging basically any kind of time series analysis are often present, including missing data, heteroscedastic “noise”, or even uncertainties in the time domain (the latter being particularly relevant in paleo-climatology). The explicit treatment of the resulting effects on VG properties has not yet been investigated elsewhere [47].

As in the previous section, in the following a corresponding study is presented for the specific case of a Gaussian white noise process as a simple, but still illustrative example. It has to be emphasized that for “real” data characterized by a non-Gaussian probability distribution function, serial dependences, or even (multi-)fractal behavior, the resulting effects could well be much stronger than in this example. A detailed study of the interdependences between such features of the data and the resulting effects of missing data and uncertainties on VG properties is, however, beyond the scope of the presented research.

**Missing data** One important problem of many observational time series - not only in geophysics - is the presence of missing data. Since existing methods of time series analysis typically require a uniform spacing in time, this problem is most often addressed by means of interpolation or sophisticated imputation of the missing observations. In general, there is a great variety of possible approaches for such gap filling, which shall not be further discussed here. Specifically, it is not always a priori clear what method performs best under the specific conditions of the data studied.

As far as VGs are concerned, the problem of missing data has not yet been explicitly addressed. Unlike many other approaches of time series analysis VGs do not explicitly require uniform sampling. Hence, missing data could be ignored when performing a corresponding analysis. However, if it is known that there must have been an observation at a given time, it could be conceptually problematic to neglect this information in the analysis. From a broader perspective, it can be argued, however, that this argument applies to all kinds of time series, since values of the considered observable (with a continuous-time variability) taken in between two subsequent observations remain always unknown, but could have a certain impact on the results of the analysis.

Looking at the issue of missing values from a complementary perspective, one can reinterpret this problem as an attack to (or just failure of) the complex network represented by the visibility graph. In complex network theory, the impact of such attacks on various types of networks has been intensively studied under the aspect of safety and robustness of infrastructures (e.g., Albert et al. 2000, Holme et al. 2002). In general, one has to distinguish random failures (corresponding to randomly missing values) from intentional attacks, which typically affect the network hubs. Since for a VG, these hubs correspond to the maxima of the underlying time series, this effect is particularly relevant for certain types of censored data, e.g., in case of measurement failures due to the limited detection range of a measurement device. Since it is known that attacks on hubs typically have a more severe effect on the network architecture than other vertices, censoring can strongly alter the properties of the resulting VGs. However, even a random removal can have notable consequences for the VG properties on both global and local scale.

In order to illustrate the effect of missing data on the properties of VGs, in the following, two different types of treatment are studied, which can be considered as opposite extreme cases. On the one hand, missing data will be simply neglected in the generation of the VG. On the other hand, since there is no information about the magnitude of the missing values, it can be a more honest solution to consider the VG as being fragmented into pieces corresponding to times before and after the missing observation, i.e., regarding the VG becoming decomposed into mutually disconnected subgraphs. It should be noted, however, that the latter approach results in the emergence of additional boundary effects. It should be emphasized that sophisticated gap filling by means of interpolation or imputation will be most likely a better strategy in many practical applications.

**Homo- and heteroscedastic uncertainties** In a similar way as for the treatment of missing values, the influence of measurement uncertainties on the resulting VG properties can be studied. For convenience,

homoscedastic uncertainties are modeled as an additional additive Gaussian white noise component ([figure 8?](#)), whereas the heteroscedastic case is studied by considering multiplicative noise with a reasonable, simple analytical distribution ([Figure 9?](#)). Again, it is found that in both cases the signal-to-noise ratio has a considerable effect by systematically shifting the distributions of vertex properties obtained for the original data towards those expected for the noise process. Note that since in the considered numerical example both signal and noise originated from mutually independent Gaussian white noise processes, there is a saturation of the KS statistics for moderate noise at values corresponding to the variance of VG properties for independent realizations of the same “signal” process.

**Uncertain timings** In full analogy to the case of uncertainties in the observable  $x$ , one can study the impact of uncertain timings  $t$  on the properties of the resulting VGs. The latter is a wide-spread problem particularly in the analysis of paleoclimate time series [244]. Since in the construction of VGs both observable and time enter in terms of an inequality defined by a linear relationship, it is not surprising that uncertain timing can indeed have a similar effect on the VG properties as uncertainties in the measurement itself. Figure 10 displays the corresponding results for a realization of Gaussian white noise originally observed with regular spacing, with the timings being corrupted later as

$$\tilde{t}_i = t_i + \Delta t(|1 - 2\eta_i| - 0.5). \quad (110)$$

Here,  $\eta_i$  are independent realizations of a random variable with uniform distribution in  $[0, 1]$ , and  $\Delta t$  is the spacing between subsequent observations in the original data set. Note that this specific form of the time-scale corruption, which allows preserving the temporal order of observations, has been inspired by the tent map as a paradigmatic nonlinear mapping often used as an illustrative example in complex systems sciences. The obtained results demonstrate that the distribution of vertex properties of a VG are indeed affected by modifications of the time-scale, however, the changes are considerably smaller than for noisy corruptions of the measurements themselves. The reason for this is that the modification used here has been restricted by the normal sampling interval, whereas the changes induced by additive and multiplicative noise allowed for comparably larger modifications in the data.

**Point process** Point processes are ubiquitous in geophysics, for instance, seismic magnitude series [245]. A point process is often characterized by random time occurrence of the events and these events are clusterized because the events are neither Poissonian nor regularly distributed over time. To check the effects of irregular timing on degree distributions, they construct two VGs from (1) the seismic series of the original random occurrence times, and the series (2) that has been substituted by regular conventional time unit. Interestingly, almost identical results have been obtained for the degree distributions, which suggests that the effects of irregular timing are not crucial.

The trivial connection of neighboring points in time in the (H)VG enhances the signature of structures due to autocorrelations in the record under study. Although this might be desirable for (H)VGs since some of their respective network properties are explicitly related with the presence of serial dependences (e.g., the typical scale of the degree distribution of HVGs, cf. [46]), there could be situations in which one is interested in removing the corresponding effects. In such cases, it is possible to introduce a minimum time difference for two observations to be connected in the network for removing the effect of slowly decaying auto-dependences, which would correspond to the Theiler window in other concepts of nonlinear time series analysis ([246]).

**General question: What information can be gained from VGs? Which networks properties are (when) useful to study?**

#### 4.4. Bivariate visibility graph methods

Despite the success, the range of applicability of (H)VGs methods has been mainly limited to univariate time series, although the most challenging problems in the area of nonlinear science concern systems that are described by multivariate time series. We can do synchronization analysis when generalizing (H)VG analysis from a univariate to bivariate time series [247, 248]. Furthermore, cross visibility algorithm helps to understand coupling and information transfer between two time series [249]. We summarize some different approaches to characterize bivariate (H)VGs.

#### 4.4.1. Visibility graph similarity

Based on the definition of HVGs, Lacasa *et al.* [163] proposed to transforming a multidimensional time series into an appropriately defined multiplex visibility graph. New information can be extracted from the original multivariate time series, with the aims of describing signals in graph-theoretical terms or to construct novel feature vectors to feed automatic classifiers in a simple, accurate and computationally efficient way.

Consider a  $M$ -dimensional real valued time series  $\{\mathbf{x}(t)\}_{t=1}^n$ , with  $\mathbf{x}(t) = (x^{[1]}(t), x^{[2]}(t), \dots, x^{[M]}(t)) \in \mathbb{R}^M$  for any value of  $t$ , measured empirically or extracted from a  $M$ -dimensional, either stochastic or deterministic dynamical system. An  $M$ -layer multiplex visibility graph  $\mathcal{M}$  is then constructed, where layer  $\alpha$  corresponds to the HVG associated to the time series of state variable  $\{x^{[\alpha]}(t)\}_{t=1}^N$ . Note that  $\mathcal{M}$  is represented by the vector of adjacency matrices of its layers  $\mathcal{A} = \{A^{[1]}, A^{[2]}, \dots, A^{[M]}\}$ , where  $A^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$  is the adjacency matrix of layer  $\alpha$ . Such a mapping builds a bridge between multivariate series analysis and the recent developments in the theory of multilayer networks [69], making it possible to employ the structural descriptors introduced to study multiplex networks as a toolbox for the characterisation of multivariate signals.

Two measures have been proposed to capture, respectively, the abundance of single edges across layers and the presence of inter-layer correlations of node degrees [163], which help to characterize information shared across variables (layers) of the underlying high dimensional system. Simply speaking, we compute these two measures by Eqs. (56, 57) based on the adjacency matrices of HVGs. More specifically, the first measure is the average edge overlap (Eq. (57)), which computes the expected number of layers of the multiplex on which an edge is present. Note that  $\omega$  takes values in  $[1/M, 1]$  and in particular  $\omega = 1/M$  if each edge  $(i, j)$  exists in exactly one layer, i.e. if there exist a layer  $\alpha$  such that  $a_{ij}^{[\alpha]} = 1$  and  $a_{ij}^{[\beta]} = 0 \forall \beta \neq \alpha$ , while  $\omega = 1$  only if all the  $M$  layers are identical. As a consequence, the average edge overlap of a multiplex VG can be used as a proxy of the overall coherence of the original multivariate time series, with higher values of  $\omega$  indicating high correlation in the microscopic structure of the signal.

The second measure proposed in [163] is to quantify the presence of interlayer correlation between the degrees of the same node at two different layers. More specifically, given a pair of layers  $\alpha$  and  $\beta$  of  $\mathcal{M}$ , respectively characterized by the degree distributions  $P(k^{[\alpha]})$  and  $P(k^{[\beta]})$ , the interlayer correlation is defined by the mutual information  $I_{\alpha,\beta}$  (Eq. (56)) where  $P(k^{[\alpha]}, k^{[\beta]})$  is the joint probability to find a node having degree  $k^{[\alpha]}$  at layer  $\alpha$  and degree  $k^{[\beta]}$  at layer  $\beta$ . The higher  $I_{\alpha,\beta}$  the more correlated the degree distributions of the two layers and therefore the structure of the associated time series. Then the average of  $I_{\alpha,\beta}$  over every pair of layers of  $\mathcal{M}$  gives a scalar variable  $I = \langle I_{\alpha,\beta} \rangle_{\alpha,\beta}$ , which captures the amount of information flow in the multivariate time series.

Note that the second measure above gives a weighted correlation matrix of  $M \times M$  and each entry is represented by  $I_{\alpha,\beta}$ . That means the original  $M$ -dimensional time series is transformed to a weighted graph of  $M$  nodes, where each node represents one layer and the weights of the edge denote the magnitude of mutual information computed from the associated (H)VG degree distributions.

#### 4.4.2. Joint and excess degrees

In addition, we introduce some network-theoretic quantities to quantify the asymmetries in bivariate time series [210]. Following the notations as described for multiplex (H)VG, we restrict our analysis with 2-dimensional time series which resulting two layers  $\alpha$  and  $\beta$ . For  $\{x^{[\alpha]}(t)\}_{t=1}^N$  and  $\{x^{[\beta]}(t)\}_{t=1}^N$ , we again denote two (H)VGs with adjacency matrices  $A^{[\alpha]}$  and  $A^{[\beta]}$ , respectively. Note that the sets of vertices are the same for both subgraphs, with differences exclusively in the set of edges.

Based on the thus obtained (H)VGs, we proceed as follows:

- From the two (H)VGs, we have two sets of neighbors,  $\mathcal{N}_{x^{[\alpha]}, x^{[\beta]}}(t) = \left\{ A_{t,t_j}^{[\alpha], [\beta]} \equiv 1, t_j \in \{1, \dots, N\} / \{t\} \right\}$  for each time  $t \in \{1, \dots, N\}$ . The *degree sequences* are then defined as

$$k_{[\alpha], [\beta]}(t) = \#\mathcal{N}_{[\alpha], [\beta]}(t) = \sum_{t_j} A_{t,t_j}^{[\alpha], [\beta]}(t). \quad (111)$$

## 2. The joint degree sequence

$$k^{joint}(t) = \# (\mathcal{N}_{[\alpha]}(t) \cap \mathcal{N}_{[\beta]}(t)) = \sum_{t_j} A_{t,t_j}^{[\alpha]}(t) \cdot A_{t,t_j}^{[\beta]}(t). \quad (112)$$

gives the number of common neighbors of a vertex corresponding to time  $t$  in both sequences.

Notably, we can define  $k^{joint}(t)$  as the degree sequence of a *joint (horizontal) visibility graph* combining the visibility criteria for two distinct time series. Here, the adjacency matrix is defined by the point-wise multiplication of the individual (H)VGs' adjacency matrices. This idea is conceptually related to the concept of joint recurrence plots encoding the simultaneous recurrence of two dynamical systems in their respective phase spaces [85]. As in the latter case, generalizing joint (H)VGs to the case of more than two time series is straightforward, but will not be considered here given the bivariate nature of the data under study.

3. In a similar spirit as the joint degree sequence, we can quantify the number of edges associated with time  $t$ , which connect to vertices contained in  $\mathcal{N}_{x^{[\alpha]}}(t)$  but *not* in  $\mathcal{N}_{x^{[\beta]}}(t)$ , or vice versa. More specifically,

$$k_{[\alpha]}^O(t) = \# (\mathcal{N}_{[\alpha]}(t) \cap \overline{\mathcal{N}_{[\beta]}(t)}) = \sum_{t_j} A_{t,t_j}^{[\alpha]}(t) \cdot (1 - A_{t,t_j}^{[\beta]}(t)) \quad (113)$$

$$k_{[\beta]}^O(t) = \# (\mathcal{N}_{[\beta]}(t) \cap \overline{\mathcal{N}_{[\alpha]}(t)}) = \sum_{t_j} A_{t,t_j}^{[\beta]}(t) \cdot (1 - A_{t,t_j}^{[\alpha]}(t)) \quad (114)$$

(where  $\overline{\mathcal{N}_{[\alpha],[\beta]}(t)} = \{1, \dots, T\} \cap \{\mathcal{N}_{[\alpha],[\beta]}(t) \cup t\}$  is the complementary set of  $\mathcal{N}_{[\alpha],[\beta]}(t)$ ) measures the number of neighbors that belong *only* to  $\mathcal{N}_{[\alpha]}(t)$  or  $\mathcal{N}_{[\beta]}(t)$ , respectively. In what follows,  $k_{[\alpha],[\beta]}^O(t)$  will be referred to as the *conditional degree sequences*. By definition,

$$k_{[\alpha],[\beta]}^O(t) = k_{[\alpha],[\beta]}(t) - k^{joint}(t). \quad (115)$$

Based on the latter definitions, we can proceed in a similar way as in Eq. (115) and compute the following properties:

$$\Delta k(t) = k_{[\alpha]}^O(t) - k_{[\beta]}^O(t) = k_{[\alpha]}(t) - k_{[\beta]}(t) \quad (116)$$

$$\Delta_{rel}k(t) = \Delta k(t) / (k_{[\alpha]}(t) + k_{[\beta]}(t)). \quad (117)$$

The *excess degree*  $\Delta k(t)$  quantifies how much “more convex” the fluctuations of  $A^{[\alpha]}$  are in comparison with  $A^{[\beta]}$  around a given time  $t$  (i.e., how many more or less visibility connections the observation of  $A^{[\alpha]}$  at time  $t$  obeys in comparison with  $A^{[\beta]}$ ). By additionally considering the *relative excess degree*  $\Delta_{rel}k(t)$  normalized by the sum of the individual degrees, we obtain a measure that does not exhibit marked sensitivity with respect to the actual degrees  $k_{[\alpha],[\beta]}$ , which may considerably vary over time according to the statistical and dynamical characteristics of the data.

In Sec. 6.2.2, we will summarize the applications of  $\Delta k(t)$  and  $\Delta_{rel}k(t)$  to characterize the North–South asymmetry of solar activity [210], which provide many nonlinear properties that have not been obtained by other methods. Notably, our approach is conceptually related with recently developed (H)VG-based tests for time series irreversibility, which compare (among others) degree distributions obtained when considering edges to past and future observations separately [50].

## 4.5. Decomposition of visibility graphs

### 4.5.1. Time-directed visibility graphs and characterizations

So far in the literature the family of visibility graphs are undirected, as visibility did not have a predefined temporal arrow. These can be made directed by again assigning to the links a time arrow, which result in the so called directed VGs (DVGs) and directed HVGs (DHVGs). Accordingly, a link between  $i$  and  $j$  (where time ordering yields  $i < j$ ) generates an outgoing link for  $i$  and an ingoing link for  $j$ . Therefore,

the degree  $k(t)$  of the node  $t$  is now split into an ingoing degree  $k_{in}(t)$ , and an out-going degree, such that  $k(t) = k_{in}(t) + k_{out}(t)$ . The ingoing degree  $k(t)$  is defined as the number of links of node  $t$  with other past nodes associated with data in the series (that is, nodes with  $t' < t$ ). Conversely, the outgoing degree  $k_{out}(t)$ , is defined as the number of links with future nodes. **A graph illustration of the method?** Then we define the *in* and *out* (or ingoing and outgoing) degree distributions of a DVG (DHVG) as  $P_{out}(k) \equiv P(k_{out} = k)$  and  $P_{in}(k) \equiv P(k_{in} = k)$ , respectively. An important property at this point is that the ingoing and outgoing degree sequences are interchangeable under time series reversal.

Given the adjacency matrix  $A_{ij}$  of a (H)VG, the degree  $k_i = \sum_j A_{ij}$  measures the number of edges incident to a given vertex  $i$ . Then, Donges *et al.* [50] used different notations for in- and out-degree sequences. More specifically, they decompose this quantity  $k_i$  for a vertex corresponding to a measurement at time  $t_i$  into contributions due to other vertices in the past and future of  $t_i$ ,

$$k_i^r = \sum_{j < i} A_{ij}, \quad (118)$$

$$k_i^a = \sum_{j > i} A_{ij} \quad (119)$$

with  $k_i = k_i^r + k_i^a$ , being referred to as the *retarded* and *advanced degrees*, respectively, in the following. Note that  $k_i^r$  and  $k_i^a$  correspond to the respective in- and out-degrees of time-directed (H)VGs as recently defined in [51]. While the degrees of an individual vertex can be significantly biased due to the finite data [47], the resulting frequency distributions of retarded and advanced degrees are equally affected.

The local clustering coefficient  $\mathcal{C}_i = \binom{k_i}{2}^{-1} \sum_{j,k} A_{ij} A_{jk} A_{ki}$  is another vertex property of higher order characterising the neighbourhood structure of vertex  $i$  [31]. Here, for studying the connectivity due to past and future observations separately, we define the *retarded* and *advanced local clustering coefficients*

$$\mathcal{C}_i^r = \binom{k_i^r}{2}^{-1} \sum_{j < i, k < i} A_{ij} A_{jk} A_{ki}, \quad (120)$$

$$\mathcal{C}_i^a = \binom{k_i^a}{2}^{-1} \sum_{j > i, k > i} A_{ij} A_{jk} A_{ki}. \quad (121)$$

Hence, both quantities measure the probability that two neighbours in the past (future) of observation  $i$  are mutually visible themselves. Note that the decomposition of  $\mathcal{C}_i$  into retarded and advanced contributions is not as simple as for the degree and involves degree-related weight factors and an additional term combining contributions from the past and future of a given vertex.

Finally, we note that other measures characterising complex networks on the local (vertex/edge) as well as global scale could be used for similar purposes as those studied in this work. However, since path-based network characteristics (e.g., closeness, betweenness, or average path length) cannot be easily decomposed into retarded and advanced contributions, the approach followed here is mainly restricted to neighbourhood-based network measures like degree, local and global clustering coefficient, or network transitivity. As a possible solution, instead of decomposing the network properties, the whole edge set of a (H)VG could be divided into two disjoint subsets that correspond to visibility connections forwards and backwards in time, as originally proposed by Lacasa *et al.* [51]. For these directed (forward and backward) (H)VGs, also the path-based measures can be computed separately and might provide valuable information.

#### 4.5.2. Tests for time series irreversibility

Testing for nonlinearity of time series has been of great interest. Various approaches have been developed for identifying signatures of different types of nonlinearity as a necessary precondition for the possible emergence of chaos. Since linearity of Gaussian processes directly implies time-reversibility [250–252], nonlinearity results (among other features) in an asymmetry of certain statistical properties under time-reversal [253]. Therefore, studying reversibility properties of time series is an important alternative to the direct quantitative assessment of nonlinearity [254]. In contrast to classical higher-order statistics requiring surrogate data techniques [253], most recently developed approaches for testing irreversibility have been based on symbolic dynamics [255–257] or statistical mechanics concepts [258–260].

The time series reversibility has the following definition: a time series  $\mathcal{S} = \{x_1, x_2, \dots, x_n\}$  is called statistically time reversible if the time series  $\mathcal{S}^* = \{x_{-1}, x_{-2}, \dots, x_{-n}\}$  has the same joint distribution as  $\mathcal{S}$ .

Therefore, time reversibility implies stationarity [251]. By this definition, time series reversibility reduces to the equivalence between forward and backward statistics and hence, nonstationary series are infinitely irreversible and therefore  $\mathcal{S}$  and  $\mathcal{S}^*$  have different statistics that increase over time [250].

In many applications, one can actually quantify different kinds of time asymmetries in the underlying dynamics on nonstationary processes. Following the previous work [51], the topological properties of (H)VGs associated to several types of nonstationary processes have been proposed to quantify the different degrees of irreversibility of several nonstationary processes [261]. Furthermore, they take advantage of the fact that the topological properties of these graphs are effectively invariant under time shift for large classes of nonstationary processes, which allows to introduce the concept of visibility graph stationarity. This in turn allows to compare to extract meaningful information on the time asymmetry of nonstationary processes.

More general, Lacasa *et al.* [261] defined time series reversibility in terms of (H)VGs in the following: a time series  $\mathcal{S} = \{x(t)\}_{t=1}^n$  is said to be (order  $p$ ) (H)VG reversible if and only if, for large  $n$ , the order  $p$  block ingoing and outgoing degree distribution estimates of the (H)VG associated to  $\mathcal{S}$  are asymptotically identical, i.e.,

$$P_{in}(k_1 k_2 \dots k_p) = P_{out}(k_1 k_2 \dots k_p). \quad (122)$$

This property yields that the ingoing and outgoing degree sequences of the original and time reversed series have the same distribution, namely,

$$P_{in}(k)[\mathcal{S}] = P_{out}(k)[\mathcal{S}^*]; P_{out}(k)[\mathcal{S}] = P_{in}(k)[\mathcal{S}^*], \quad (123)$$

where  $\mathcal{S}^* = \{x_{n+1-t}\}_{t=1}^n$  represents the time reversed series. For time series, we assess how close the system is to reversibility by quantifying the distance between  $P_{in}$  and  $P_{out}$ .

The distance between the *in* and *out* degree distributions has been calculated by the Kullback-Leibler divergence (KLD) [51, 261], which is introduced in information theory as a measure to distinguish between two probability distributions. More specifically, the KLD between these two distributions is

$$D[P_{out}(k) \| P_{in}(k)] = \sum_k P_{out}(k) \log \frac{P_{out}(k)}{P_{in}(k)}. \quad (124)$$

This measure vanishes if and only if the outgoing and ingoing degree probability distributions of a time series are identical. Then the (H)VG reversibility is redefined if the following expression holds:

$$\lim_{n \rightarrow \infty} D[P_{out}(k) \| P_{in}(k)] = 0. \quad (125)$$

Truly irreversible process have positive values of Eq. (124) in the limit of large  $n$  [51].

Lacasa *et al.* [51, 261] conjecture that the information stored in the *in* and *out* distributions takes into account the amount of time irreversibility of the associated series. More precisely, they claim that this can be measured, in a first approximation, as the distance (in a distributional sense) between the *in* and *out* degree distributions ( $P_{in}(k)$  and  $P_{out}(k)$ ). If needed, higher order measures can be used, such as the corresponding distance between the *in* and *out* degree-degree distributions ( $P_{in}(k, k')$  and  $P_{out}(k, k')$ ). These are defined as the *in* and *out* joint degree distributions of a node and its first neighbors, describing the probability of an arbitrary node whose neighbor has degree  $k'$  to have degree  $k$ . Namely, we compare the outgoing degree distribution in the actual (forward) series  $P_{k_{out}} = P_{out}(k)$  with the corresponding probability in the time-reversed (or backward) time series, which is equal to the probability distribution of the ingoing degree in the actual process  $P_{k_{out}} = P_{in}(k)$ .

Therefore, by calculating Eq. (124), Lacasa *et al.* [51] have shown that one can correctly distinguish between reversible and irreversible stationary time series, including analytical and numerical studies of its performance for: (i) reversible stochastic processes (uncorrelated and Gaussian linearly correlated), (ii) irreversible stochastic processes, (iii) reversible (conservative) and irreversible (dissipative) chaotic maps, and (iv) dissipative chaotic maps in the presence of noise.

Notice that the majority of previous methods to estimate time series irreversibility generally proceed by first making a (somewhat ad hoc) local symbolization of the series, coarse-graining each of the series data

into a symbol (typically, an integer) from an ordered set [255–260]. The method based on directed (H)VGs here lacks an ad hoc symbolization process, which may in principle take into account multiple scales. The unnecessary requirement of symbolization is desirable if we want to tackle complex signal and hence it can be applied directly to any kind of real-valued time series.

We note that (H)VG reversibility varies depending on the detailed properties of particular processes [261], which calls for careful interpretations. For instance, both analytical calculations and numerical simulations show that unbiased additive random walks, while nonstationary, are both (H)VG stationary and (H)VG time reversible. On the other hand, biased memoryless additive random walks are HVG irreversible with finite irreversibility measures that quantify the degree of time asymmetry, while these are still VG reversible, as VGs are invariant under superposition of linear trends in the original data. Numerics suggest that HVGs can capture, for both finite and infinite series size, the irreversible nature of non-Markovian additive random walks, whereas VGs are only able to do so for finite series. For multiplicative random walks, the processes are HVG reversible if the process is akin to an unbiased additive process in logarithmic space, and time irreversible if the process reduces to a biased additive process in logarithmic space. Finally, the VGs capture the time irreversible character of multiplicative random walks, yielding finite values in the unbiased case and asymptotically diverging quantities in the biased case. Furthermore, these conclusions are based on the limit of infinitely long time series  $n \rightarrow \infty$  and finite size time series always yields finite, non-null values of HVG and VG irreversibility [262], which needs a proper test justifying the statistical significance.

While the results of KLD measure for reversible and irreversible dynamics quantitatively differ in several orders of magnitude, a statistical test is required. Lacasa *et al.* proposed to address the statistical significance by surrogate techniques as follows: one first proceeds to shuffle the series under study in order to generate a randomized resampled data set with the same underlying probability density. This resampled series, whose irreversibility measure is asymptotically null, is considered as the null hypothesis of the test. Taking a slightly different algorithm, Donges *et al.* [50] have thoroughly extended this idea and provided a set of rigorous statistical tests for time series irreversibility, which can be formulated based on both standard and horizontal VGs and utilise different network properties. Specifically, for both VGs and HVGs, network degrees as well as local clustering coefficients can be decomposed into contributions from past and future observations (Eqs. (118-121)), which allows studying some of the time series' statistical properties under time-reversal. They find statistically significant deviations between the distributions of time-ordered vertex properties for nonlinear systems for which the absence of time-reversal symmetry is known, but not for linear systems.

Time-irreversibility of a stationary stochastic process or time series  $\{x_i\}$  requires that for arbitrary  $n$  and  $m$ , the tuples  $(x_n, x_{n+1}, \dots, x_{n+m})$  and  $(x_{n+m}, x_{n+m-1}, \dots, x_n)$  have the same joint probability distribution [251]. Instead of testing this condition explicitly (which is practically unfeasible in most situations due to the necessity of estimating high-dimensional probability distribution functions from a limited amount of data), for detecting time series irreversibility it can be sufficient to compare the distributions of certain statistical characteristics obtained from both vectors (e.g., [263]). Following the decomposition of vertex properties into time-directed contributions proposed above, (H)VG-based methods appear particularly suited for this purpose. Specifically, in the following we will utilise the frequency distributions  $p(k^r)$  and  $p(k^a)$  ( $p(\mathcal{C}^r)$  and  $p(\mathcal{C}^a)$ ) of retarded and advanced vertex properties as representatives for the statistical properties of the time series when viewed forward and backward in time.

In the case of time-reversibility, we conjecture that both sequences  $\{k_i^r\}$  and  $\{k_i^a\}$  (or  $\{\mathcal{C}_i^r\}$  and  $\{\mathcal{C}_i^a\}$ ) should be drawn from the same probability distribution, because the visibility structure towards the past and future of each observation has to be statistically equivalent. In turn, for an irreversible (i.e., nonlinear) process, we expect to find statistically significant deviations between the probability distributions of retarded and advanced characteristics.

As an alternative to the Kullback-Leibler distance between the empirically observed distribution functions used by Lacasa *et al.* [51], we propose utilising some standard statistics for testing the homogeneity of the distribution of random variables between two independent samples. In this framework, rejecting the null hypothesis that  $\{k_i^r\}$  and  $\{k_i^a\}$  ( $\{\mathcal{C}_i^r\}$  and  $\{\mathcal{C}_i^a\}$ ) are drawn from the same probability distribution, respectively, is equivalent to rejecting the null hypothesis that the time series under investigation is reversible. Since for sufficiently long time series (representing the typical dynamics of the system under study), the

available samples of individual vertex properties approximate the underlying distributions sufficiently well, we can (despite existing correlations between subsequent values) consider the Kolmogorov-Smirnov (KS) test for testing this null hypothesis. Specifically, a small  $p$ -value of the KS test statistic (e.g.,  $p < 0.05$ ) implies that the time series has likely been generated by an irreversible stochastic process or dynamical system. Even more, these  $p$ -values are distribution-free in the limit of  $N \rightarrow \infty$ . Neglecting possible effects of the intrinsic correlations between the properties of subsequent vertices on the estimated  $p$ -values (which shall be addressed in future research), this implies that we do *not* need to construct surrogate time series for obtaining critical values of our test statistics as in other irreversibility tests. Note that other (not network-related) statistical properties sensitive to the time-ordering of observations could also be exploited for constructing similar statistical tests for time series irreversibility [50]. **a figure emphasize the statistical significance aspect? both  $k$  and  $C$  series for this purpose.** Recently, a combination of Kullback-Leibler distance between the ingoing and outgoing degree sequences and the so-called inversion number of the permutation of the original time series has been proposed to characterize the asynchronous patterns of time irreversibility [264].

Utilising standard as well as horizontal VGs for discriminating between the properties of observed data forwards and backwards in time has at least two important benefits: (i) Unlike for some classical tests (e.g., [253]), the reversibility properties are examined without the necessity of constructing surrogate data. Hence, the proposed approach saves considerable computational costs in comparison with such methods and, more importantly, avoids the problem of selecting a particular type of surrogates. Specifically, utilising the KS test statistic or a comparable two-sample test for the homogeneity (equality) of the underlying probability distribution functions directly supplies a  $p$ -value for the associated null hypothesis that the considered properties of the data forward and backward in time are statistically indistinguishable. (ii) The proposed approach is applicable to data with non-uniform sampling (common in areas like palaeoclimate [47] or astrophysics) and marked point processes (e.g., earthquake catalogues [245]). For such data, constructing surrogates for nonlinearity tests in the most common way using Fourier-based techniques is a challenging task, which is avoided by (H)VG-based methods.

We emphasize that our method exploits the time-information explicitly used in constructing (H)VGs. Other existing time series network methods (e.g., recurrence networks [38, 42, 43]) not exhibiting this feature cannot be used for the same purpose. Furthermore, there are methodological questions such as the impacts of sampling, observational noise, and intrinsic correlations in vertex characteristics as well as a systematic comparison to existing methods for testing time series irreversibility that need to be systematically addressed in future research. Furthermore, (H)VG-based methods are generally faced with problems such as boundary effects and the ambiguous treatment of missing data [47], which call for further investigations.

However, path-based measures of (H)VGs are known to be strongly influenced by boundary effects [47], so that they could possibly lose their discriminative power for irreversibility tests. In addition, irreversibility tests have been conducted for various real valued time series. Examples include neuro-physiological EEG recordings [50], mean temperature anomaly series [265], financial time series [266], oil-water two phase flows [267], meteorological stream flow fluctuation [268], correlated fractal processes [262], seismic sequences of Mexican subduction zone [269].

## 5. Transition networks (Reik)

The construction of the transition network depends on a proper phase space partition. For instance, we first mesh the phase space with box of equal sizes following the traditional idea of fractal dimension computations [7, 147] or complexity measures [270]. Then each partition is labeled with  $\pi_i$  and regarded as a vertex in the network. The connectivity between partition  $\pi_i$  and  $\pi_j$  is then represented by the transition frequency following the temporal order of observations. Alternatively, one may define a partition as a “state” based on a symbolic representation of time series, for instance, ordinal patterns of the series [28, 271]. Transforming the time series into a transition network is a process of mapping the temporal information into a Markov chain to obtain a compressed or simplified representation of the original dynamics.

### 5.1. Markov chains

In the terminology of Markov chain modeling [52], these partition algorithms generally map time series data to a Markov chain by defining nodes as some set of states that span the time series points, and allocating directed edges based on temporal succession that represent transitional probabilities based on the source data. Therefore, indicators of dynamics are derived from the transition matrix [52].

### 5.2. Symbolic encoding of time series

Coarse-graining the range of values in a time series into a suitable set of classes (“symbols”)  $\{\pi_1, \dots, \pi_k\}$  allows considering the transition probabilities  $w_{\alpha,\beta} = p(x_{i+1} \in \pi_\beta | x_i \in \pi_\alpha)$  between these classes in terms of a weighted and directed network. Therefore, the first step of this approach is to applying a suitable symbolic discretization to the phase space of the studied system. The next step is to explicitly use the temporal order of observations, i.e., their connectivity, to represent causality relationships contained in the dynamics of the observed dynamical system, which hence results in a time directed transition network. The directed network is represented by a weighted matrix  $W = \{w_{\alpha,\beta}\}, \alpha, \beta \in [1, \dots, k]$ . Note that for a trajectory that does not leave a finite volume in phase space, there is only a finite number of discrete “states”  $\pi_i$  with a given minimum size in phase space. This implies the presence of absorbing or recurrent states in the resulting transition network.

The transition probability approach is well suited for identifying such “states” (i.e. regions in phase space) that have a special importance for the causal evolution of the studied system in terms of betweenness centrality  $b_v$  and related measures. Moreover, the resulting networks do not only depend on a single parameter, but on the specific definition of the full set of classes. Note, however, that coarse graining might be a valid approach in case of noisy real-world time series, where extraction of dynamically relevant information hidden by noise can be supported by grouping the data. In contrast to the other approaches for constructing complex networks from time series, the topology of transition networks depends on the specific choice of discretization. In the next sections, we give some examples.

#### 5.2.1. Threshold-based coarse-graining

Note that meshing the phase space with equal size  $\varepsilon$  is a threshold-based coarse-graining of the original system. Note that quantile mapping discretizing continuous time series shares much similarity with the threshold based coarse graining technique [272]. In [38], we have illustrated an example in obtaining a transition network from time series of the Lorenz system. The transition behavior among different partitions is characterized by state transition matrix  $W$ , which is weighted and each entry  $W_{ij}$  is estimated by the transition frequency from state  $i$  to  $j$ . Traditionally, the transition entropy is simply the Shannon entropy of the weight matrix  $W$ , which is used to characterize the changes of the underlying system.

We note that the temporal information is lost after the coarse graining and the transition frequency matrix  $W$  is estimated over the entire time series. Therefore, the resulting ordinal transition network is static representations [273]. Inspired by the current ideas of temporal networks, Weng *et al.* proposed to construct a temporal network from time series by unfolding temporal information  $t$  into an additional topological dimension. More specifically, a transition from node  $i$  to  $j$  is established whenever the trajectory flow performs a transition from  $i$  to  $j$  at time  $t$  which is denoted as  $(i \rightarrow j; t)$ . By adding the additional time axis to the transition route, the consecutive memory network is constructed by introducing the memory factor  $\tau$  and Weng *et al.* further proposed the memory entropy analysis to characterize the memory effect of the observed time series. The identified memory effect can accurately differentiate various types of time series including white noise,  $1/f$  noise, AR model, periodic and chaotic time series.

#### 5.2.2. Order pattern-based coarse graining

There is a growing number of works in transforming time series into networks by ordinal partitions of time series [28, 53, 274–276]. A series of systematic investigations of ordinal methods has been conducted in irregularly sampled time series [56–58], which shows high potential for studies of experimental observation data from climate sciences [59]. In this method, the first step is to embed a one-dimensional time series  $\{x(t)\}$  into phase space by using techniques from traditional time delay embedding, i.e., a proper choice of

embedding dimension  $D_x$  and time delay  $\tau$ . Then, embedded points in phase space are mapped to nodes in the network space according to its rank order and links are allocated between nodes based on temporal succession on the trajectory. In Fig. 8, we show an example of ordinal partition network using the algorithms of [28].

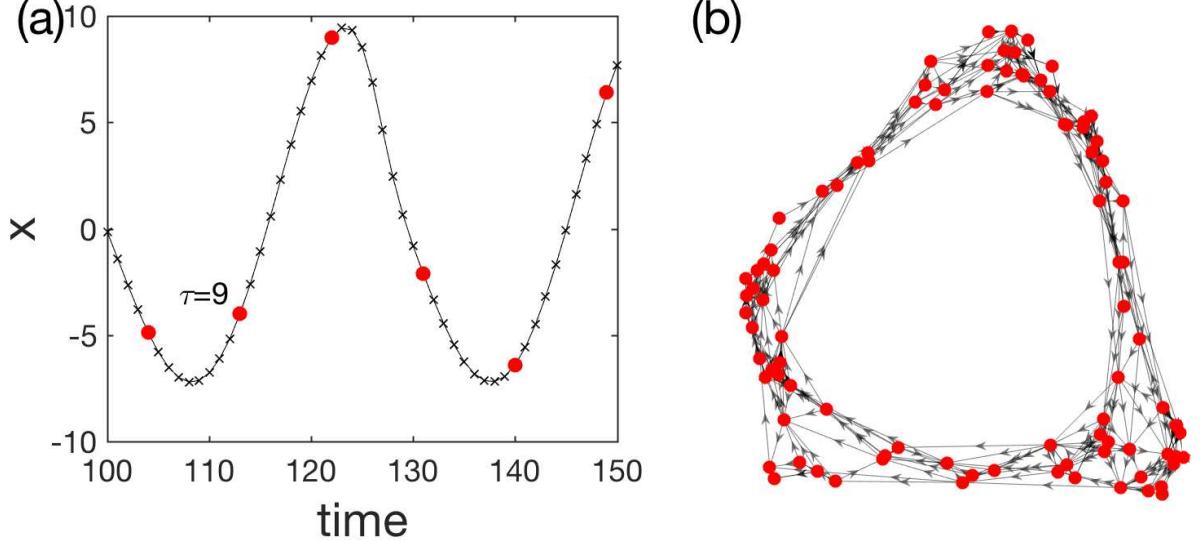


Figure 8: (a) illustration of permutation symbols from a time series of the Rössler attractor ( $a = 0.165$ ). Assume  $\tau = 9$  and  $D_x = 6$ . The embedded vector is highlighted by red color  $\vec{x}_{104} = \{x_{104}, x_{113}, x_{122}, x_{131}, x_{140}, x_{149}\}$  and its corresponding pattern is defined by the rank ordering  $\pi_{104} = \{5, 1, 2, 4, 6, 3\}$ . (b) ordinal pattern transition network (isolated vertices and self-loops are excluded for the visualization purpose). Directed edges are denoted by arrows. Reproduced from [28].

In [28], McCullough *et al* illustrated the construction algorithm in detail by the Rössler system and find that periodic dynamics translate to ring structures whereas chaotic time series translate to band or tube-like structures – thereby indicating that this algorithm generates networks whose structure is sensitive to system dynamics. Furthermore, it is demonstrated that simple network measures including the mean out degree and variance of out degrees can track changes in the dynamical behaviour in a manner comparable to the largest Lyapunov exponent [28]. Therefore, these results show that measures of transition networks have the potential to be useful as indicators for dynamical discrimination and for detecting change points.

Note that the embedding dimension  $D_x$  and time delay  $\tau$  are two important parameters for constructing ordinal partition networks, in particular having crucial impacts on the appearance of forbidden order patterns [53, 57, 58]. The selection of time delay  $\tau$  must be chosen in relation to the sampling rate for continuous systems. The authors proposed using a time delay  $\tau > 1$  [277]. In [28], they proposed to select these two parameters by traditional methods used for time series embedding, for instance, the first zero of the autocorrelation of the time series because it provides a sufficiently good phase space reconstruction for the Rössler time series. While there does not yet exist a robust metric for determining the correct choice of  $D_x$ , in the cases time series of the Rössler system they found that networks with the most visually intuitive structure often correspond with peak values of degree variance with respect to  $D_x$ . In addition, it was demonstrated that the range  $6 \leq D_x \leq 10$  was the most useful when using simple network measures to track changes in dynamics [28]. Note that the choice of  $D_x$  also determines the level of simplification of the original phase space by ordinal partitions.

Most of the recent works have focused only on univariate time series  $\{x(t)\}$ . However, the generalization to multivariate time series remains largely untouched. Most of the observable phenomena in the empirical sciences are of a multivariate nature. For instance, assets in stock markets are observed simultaneously and their joint development is analyzed to better understand tendencies. In climate science, multiple observations (temperature, pressure, precipitation, human activities etc, from different locations) are the basis of reliable

$\Pi$	$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$	$\pi_7$	$\pi_8$
$\Delta x$	$\pi_x^1, +$	$\pi_x^1, +$	$\pi_x^1, +$	$\pi_x^1, +$	$\pi_x^0, -$	$\pi_x^0, -$	$\pi_x^0, -$	$\pi_x^0, -$
$\Delta y$	$\pi_y^1, +$	$\pi_y^1, +$	$\pi_y^0, -$	$\pi_y^0, -$	$\pi_y^1, +$	$\pi_y^1, +$	$\pi_y^0, -$	$\pi_y^0, -$
$\Delta z$	$\pi_z^1, +$	$\pi_z^0, -$						

Table 3: Order patterns in three dimensional time series  $(x(t), y(t), z(t))$ .

predictions for the future climate conditions.

In [55], they propose to construct ordinal partition transition networks from multivariate (high dimensional) data. Given a scalar time series  $\{x(t)\}$  which is produced by a deterministic dynamical system, the order structure of the time series depends on the embedding dimension  $D_x$  and time delay  $\tau$ . Let us start with embedding dimension  $D_x = 2$ . Neglecting equality, we have two relations between  $x(t)$  and  $x(t + \tau)$ , namely, two symbol sequences representing order patterns  $\pi_x$ :

$$\pi_x(t) = \begin{cases} 1 & \text{if } x(t) < x(t + \tau), \\ 0 & \text{if } x(t) > x(t + \tau). \end{cases} \quad (126)$$

In [274], Small *et al* used a fixed lag  $\tau = 1$  for embedding and we follow this idea. By this choice, the order pattern  $\pi_x^1$  captures the increasing trend, respectively,  $\pi_x^0$  corresponds to the decreasing trend of the time series. This definition is equivalent to considering the signs of the increments  $\Delta x(t) = x(t + 1) - x(t)$  by a first-order difference of the original series.

Generalizing the above idea to the case of three dimensional time series  $(x(t), y(t), z(t))$ , we first obtain the increment series  $(\Delta x(t), \Delta y(t), \Delta z(t))$ . Then the order patterns are defined by the combinations of signs of  $\Delta x(t), \Delta y(t)$  and  $\Delta z(t)$ . In particular, the ordinal pattern  $\Pi(t) \in (\pi_1, \dots, \pi_i), i = 1, \dots, 8$  of a three dimensional time series  $(x(t), y(t), z(t))$  is enumerated in Tab. 3. Therefore, the dimension of order pattern  $\Pi(t)$  for an  $n$ -dimensional time series  $(\{x_1\}(t), \dots, \{x_n\}(t))$  is  $D = 2^n$  since each component has either increasing or decreasing trend at time  $t$ .

Noting that we consider the increments between two consecutive time points of each measurement in the space of multi measurements, which captures the dynamic properties of the multi-variate time series in its associated velocity space (difference space). Therefore, time delay  $\tau$  in the order pattern definition (Eq. (126)) has rather a different interpretation with the time delay that is often used in embedding. We can certainly generalize the discussion to the case of time delays larger than 1 (i.e.,  $\tau > 1$ ) and embedding dimension  $D_x > 2$  for each variable (measurement), but we think that the physical meaning in terms of dynamics becomes ambiguous for multivariate time series.

In addition, the resulting ordinal pattern transition network utilizes nullclines to obtain phase space partitions of the systems. As shown in Fig. 9(a), the chaotic Rössler system is color coded by the ordinal pattern partitions and the corresponding transition network in shown in Fig. 9b.

In order to emphasize the importance of non-self transitions between ordinal patterns, we remove the self-loops, which is typical of most research work on complex networks [30]. Furthermore, we remove self-loops before computing the weighted matrix  $W$  to keep the normalization  $\sum_{i,j} w_{ij} = 1$ . Note that self-loops should not be expected with large amounts in stochastic processes.

Given the observations that different occurrence frequencies of ordinal patterns  $p(\pi_i)$  and their transitions  $w_{i,j}$ , two Shannon entropies are defined as

$$\mathcal{H}_O = - \sum_{i=1}^{2^n} p(\pi_i) \log_2 p(\pi_i), \quad (127)$$

$$\mathcal{H}_T = - \sum_{i,j=1}^{2^n} w_{ij} \log_2 w_{ij}. \quad (128)$$

In the terminologies of [275],  $\mathcal{H}_O$  characterizes the vertex (node) complexity and  $\mathcal{H}_T$  is for the edge (link) transitional complexity, both of which have been shown useful for characterizing synchronization transition

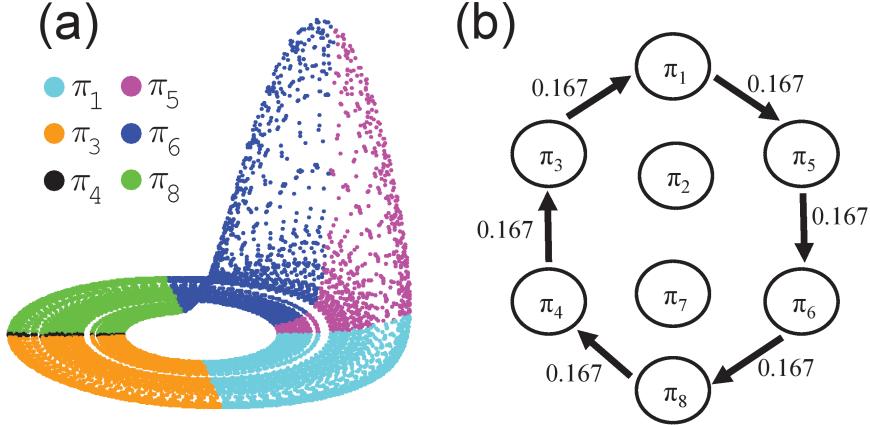


Figure 9: (a) Rössler attractor in phase space color coded by order patterns ( $a = 0.165$ ), (b) ordinal pattern transition network (self-loops are excluded).

dynamics [55]. Note that Eq. (128) measures the transitional complexity of the ordinal patterns, but in a slightly different way of normalizations as used in [275, 276].

### 5.2.3. Cross and joint ordinal transition networks

The method of [55] has been further generalized to construct cross and joint ordinal partition transition networks for two coupled systems [278]. Let us start with an example of a single chaotic Rössler system as represented by three variables  $(x_1(t), y_1(t), z_1(t))$ . The ordinal pattern transition network is reconstructed based on the signs of the increments of each variable  $(\Delta x_1(t), \Delta y_1(t), \Delta z_1(t))$ , where  $\Delta x_1(t) = x_1(t+1) - x_1(t)$ ,  $\Delta y_1(t) = y_1(t+1) - y_1(t)$ , and  $\Delta z_1(t) = z_1(t+1) - z_1(t)$ . The definition of patterns  $\Pi(t) \in (\pi_1, \dots, \pi_8)$ ,  $i = 1, \dots, 8$  are enumerated in Tab. 3. For two coupled systems, we have further time series from the other system as represented by  $(x_2(t), y_2(t), z_2(t))$ .

A cross ordinal pattern transition network (COPT) compares the relative speeds between two systems by the signs of  $(\Delta x_1(t) - \Delta x_2(t))$ ,  $(\Delta y_1(t) - \Delta y_2(t))$  and  $(\Delta z_1(t) - \Delta z_2(t))$ . The pattern definitions of a COPT are shown in Tab. 4. An example of COPT is reconstructed for two coupled Rössler system in a non-sync regime [278], which is shown in Fig. 10(a). Considering the effects of the different magnitudes of the three variables,

$\Pi$	$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$	$\pi_7$	$\pi_8$
$\Delta x_1 - \Delta x_2$	+	+	+	+	-	-	-	-
$\Delta y_1 - \Delta y_2$	+	+	-	-	+	+	-	-
$\Delta z_1 - \Delta z_2$	+	-	+	-	+	-	+	-

Table 4: Pattern definitions of a COPT. Note that “+” means a positive value while “-” is for a negative value.

we also compute an *alternative* COPT by replacing  $\Delta x_1(t) - \Delta x_2(t)$  by  $\Delta x_1(t)/x_1(t) - \Delta x_2(t)/x_2(t)$ , respectively,  $\Delta y_1(t) - \Delta y_2(t)$  by  $\Delta y_1(t)/y_1(t) - \Delta y_2(t)/y_2(t)$ , and  $\Delta z_1(t) - \Delta z_2(t)$  by  $\Delta z_1(t)/z_1(t) - \Delta z_2(t)/z_2(t)$ . An example of the alternative COPT is shown in Fig. 10(b). Comparing Fig. 10(a) to 10(b), the alternative COPT reflects better the non-coherent transitions between ordinal patterns since the coupling strength is in the non-synchronization regime ( $\kappa_1 = 0$  and  $\kappa_2 = 0.01$ ).

In an analogy, a joint ordinal pattern transition networks (JOPT) compares the relative speeds between two systems by the signs of  $\Delta x_1(t) \cdot \Delta x_2(t)$ ,  $\Delta y_1(t) \cdot \Delta y_2(t)$  and  $\Delta z_1(t) \cdot \Delta z_2(t)$  and the pattern definitions of a JOPT are summarized in Tab. 5. An example of JOPT is shown in Fig. 10(c). In contrast to cross ordinal patterns, we notice that the joint ordinal patterns represent whether the respective variables of two systems show the same trend of changes or not, regardless of the magnitudes of the respective variables.

The ideas of both COPT and JOPT have been particularly applied to analyze synchronization transitions

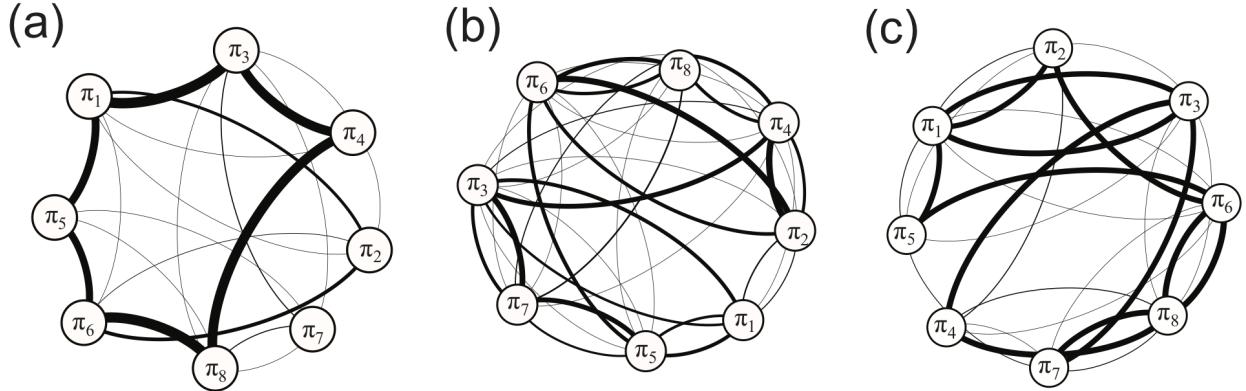


Figure 10: Cross and joint ordinal pattern transition networks, which are reconstructed from two coupled Rössler systems in the non-sync regime [278]. (a) Cross ordinal pattern transition network (COPT), (b) an alternative version of COPT, and (c) joint ordinal pattern transition network (JOPT). The directions of links have been suppressed for better visualizations. Reproduced from [278].

$\Pi$	$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$	$\pi_7$	$\pi_8$
$\Delta x_1 \cdot \Delta x_2$	+	+	+	+	-	-	-	-
$\Delta y_1 \cdot \Delta y_2$	+	+	-	-	+	+	-	-
$\Delta z_1 \cdot \Delta z_2$	+	-	+	-	+	-	+	-

Table 5: Pattern definitions of a JOPT. Note that “+” means a positive value while “−” is for a negative value.

[278]. Note that a COPT and JOPT are two slightly different ways to construct networks from multivariate time series, providing complementary information. The ordinal patterns of a COPT are defined by considering the signs of the difference of  $\Delta\vec{x}_1 - \Delta\vec{x}_2$  between two subsystems. In contrast, the ordinal patterns of a JOPT are defined by the signs of the product of  $\Delta\vec{x}_1 \cdot \Delta\vec{x}_2$ . It is certain that the amplitudes of oscillations of different variables influence directly the definition of a COPT. However amplitudes become not important for a JOPT because only the signs of the product are considered. In addition, it is straightforward to generalize the ideas of JOPTs from two to three (or even  $n$ ) coupled subsystems with an extended number of pattern definitions. However, it remains to be a big challenge for constructing a COPT for three coupled subsystems.

#### 5.2.4. Other approaches

For one dimensional symbol sequence, one may construct a directed symbol transition network [279]. Working with experimental data of *RR*-intervals from cardiac regulations, Makowiec *et. al* proposed to construct transitions networks from the subsequent increment series  $\Delta RR_n = RR_n - RR_{n-1}$  [280–285]. In this series of work, they have demonstrated that transition network approaches are a powerful tool in quantifying the unique properties of the *RR*-interval time series of patients after heart transplant surgery.

When constructing ordinal partition transition networks by a sliding window scheme [274], the ordinal pattern of the windowed sequence corresponds to one node of the network. Since the amplitude information is neglected, this approach may be combined with a transition network, where the nodes of the network are the binned amplitudes of the time series [277]. More specifically, each time step  $n$  is associated with a symbol-pair containing the amplitude information  $\alpha(n)$  and the ordinal pattern  $\pi(n)$ . The former is calculated by binning the time series in the interval  $[\min(\{y_n\}), \max(\{y_n\})]$  into  $Q$  equal regions.  $\alpha(n)$  is then simply the bin number of  $y_n$ . Respectively,  $\pi(n)$  is the ordinal pattern of the embedded vector  $[y_n, y_{n+\tau}, \dots, y_{n+(L-1)\tau}]$ . The symbol-pair at step  $n$ ,  $(\alpha(n), \pi(n))$ , is then one node of the network and it is connected by a directed link to the symbol-pair  $(\alpha(n+1), \pi(n+1))$  of the successive time step. Furthermore, this algorithm has been combined with recurrence networks and surrogate networks, which

has shown power in detecting weak nonlinearities in time series [286].

Coarse graining a financial time series, one may focus on the particular up-down behaviour of the volatility stock index series as presented in [287, 288]. More specifically, they symbolize time series by the parameter  $\theta = \arctan \Delta x / \Delta t$ , which characterizes the local increasing/decreasing velocity of the measurement. Then, the local velocities are coarse grained into four states ( $R, r, d, D$ ), which correspond to violent-up meta, common-up meta, common down meta, and violent down meta patterns respectively. They found that the topological important nodes of the resulting transition networks play important roles in both information control and transport of stock market [288].

In a series of works by Gao *et al.*[289–291], they proposed a linear regression patterns transmission algorithm which captures the evolution of linear regression of bivariate time series. This algorithm has been demonstrated to be a useful tool to show the correlation mode transmission in crude oil spot price and future price [292]. Based on reduced autoregressive models generated from time series, a proper directed transition network reconstruction algorithm has been proposed in [293], such that the delay information has been successively captured by the transition behavior of the resulting network. Furthermore, proper surrogate method is required when extending these ideas from a univariate to multivariate time series analysis [294].

### 5.3. Network interpretation of Markov chains

## 6. Applications (Reik + Jonathan)

In this section, we illustrate the application of the above discussed methods on selected examples from theory and real-world research questions. We focus on recurrence network approaches, (horizontal) visibility graphs, and transition network approaches since those methods have found a much wider and deeper applications in diverse research fields.

### 6.1. Recurrence networks

We separate this section in two subsections, corresponding to numerical examples and real-world applications respectively.

#### 6.1.1. Numerical examples

*Illustration of network properties in phase space:* . RN approaches can be useful for disentangling different dynamical regimes not only in dissipative, but also in Hamiltonian systems. For dissipative systems, characteristics based on recurrence networks have recently attracted much interest for discriminating qualitatively different types of dynamics in terms of measures of complexity, dynamical invariants, or even structural characteristics of the underlying attractor's geometry in phase space. More specifically, local vertex-wise network characteristics of time series can be visualized in the corresponding phase space [43]. For instance, phase spaces of discrete Hénon map, the chaotic Rössler and Lorenz system have been color coded by vertex degrees  $k$ , local clustering coefficient  $C_v$  and betweenness  $b_v$ , respectively [43]. Furthermore, global network measures of transitivity  $\mathcal{T}$ , average path length  $\mathcal{L}$  have been applied to identify dynamical transitions in the Logistic map when the control parameter is changed [42].

In [157], we demonstrate that the validity of the RN approach for achieving the same goals in low-dimensional Hamiltonian systems. Using the standard map as a paradigmatic example, we perform recurrence network analysis for distinguishing regular and chaotic orbits co-existing in the same phase space. Specifically, we show that sticky orbits of the standard map can have a distinct geometric organization that can be detected reliably by recurrence network analysis of relatively short time series (e.g.,  $N = 1,000$  or  $5,000$  points). Let us restrict our attention to the standard map

$$\mathbf{v}(t) : \begin{cases} y_{n+1} = y_n + \frac{\kappa}{2\pi} \sin(2\pi x_n), \\ x_{n+1} = x_n + y_{n+1}, \end{cases} \mod 1 \quad (129)$$

with  $\kappa$  denoting the system's single control parameter, and  $\mathbf{v}_n = (x_n, y_n)$  being the state vector of the system at its  $n$ -th iteration. This model is probably the best-studied chaotic Hamiltonian map and can be interpreted as a Poincaré section of a periodically kicked rotor [295, 296].

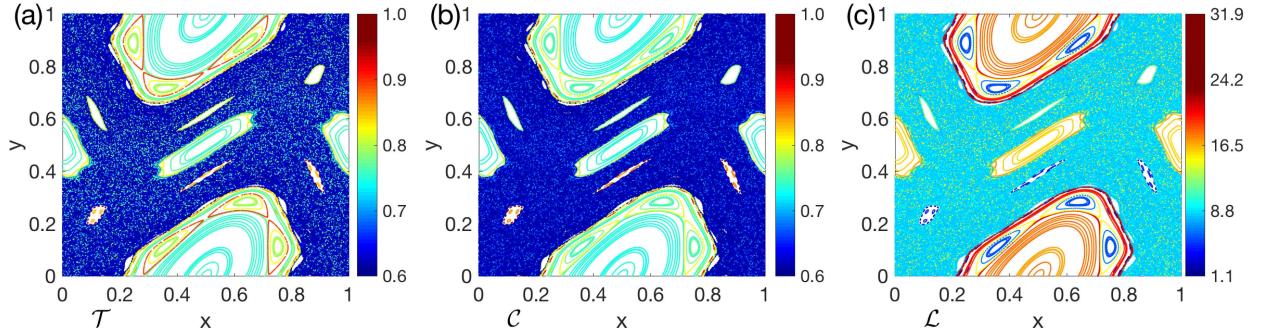


Figure 11: Phase space of the standard map (Eq. (129)) characterized by recurrence statistics for the standard map using fixed  $RR = 0.02$ . (a)  $\mathcal{T}$ , (b)  $\mathcal{C}$ , (c)  $\mathcal{L}$ . Reproduced from [157].

The whole phase space presents a complicated mixture of domains of chaotic trajectories coexisting with domains of regular ones. The regular component consists of both periodic and quasi-periodic trajectories, while the irregular one contains one or more chaotic orbits. A typical chaotic trajectory needs a long time to fill its corresponding domain in phase space. Due to the existence of periodic islands, once a chaotic orbit gets close to such an island, it can stay close to it and be almost regular in its motion for a rather long time. After this transient period it escapes again to the large chaotic sea. Such a long-term confinement of the trajectory close to the regular domain is commonly referred to as stickiness [296, 297], which has been accepted as a fundamental property of many Hamiltonian systems.

To visualize the phase space, we choose 200 initial conditions distributed randomly within the domain of definition of the standard map,  $(x, y) \in [0, 1] \times [0, 1]$ . Here, we use the canonical parameter value of  $\kappa = 1.4$  in Eq. (129). All trajectories are computed for 5000 time steps. When aiming for a quantitative comparability of RN characteristics (which can depend on  $RR$ ), we suggest to adaptively choose  $\varepsilon$  such that the  $RR$  has the same fixed value [157]. The corresponding results for the three global RN measures  $\mathcal{T}, \mathcal{C}$  and  $\mathcal{L}$  are shown in Figs. 11. In this figure, for each random realization of 5000 time series points, we use the same color coding which is determined by the RN measures.

The overall structure of the phase space with its intermingled regular and irregular components is captured by all these RN measures as shown in Figs. 11. Further dynamic and geometric measures are discussed in [157]. Specifically, quasi-periodic trajectories are characterized by larger values of network transitivity  $\mathcal{T}$  and  $\mathcal{C}$ , while filling chaotic ones have smaller  $\mathcal{T}$  and  $\mathcal{C}$  (Fig. 11(a, b)).

Traditionally a chaotic orbit can fill the complete domain (as  $t \rightarrow \infty$ ), regular ones are distinct and mutually nested, which results in the chaotic domain is often larger than the regular one. The RN measure  $\mathcal{L}$  depends clearly on the size of the orbit. Therefore, the pattern of  $\mathcal{L}$  (Fig. 11(c)) is influenced by thresholds. In turn, when fixing  $RR$  the effect of different spatial distances on the estimated RN average path lengths  $\mathcal{L}$  is essentially removed (Fig. 11(c)).

*Characterize non-stationary systems.* While the aforementioned results have been obtained for stationary systems, i.e., independent realizations of the system at fixed parameter values, tracing temporal changes in dynamical complexity of non-stationary systems is another interesting field of application with numerous examples in the real-world. Using model systems with drifting parameters such as the Lorenz [128] or Rössler systems, it is possible to systematically evaluate the performance of RN characteristics in a sliding windows framework, underlining their capabilities for discriminating between qualitatively different types of dynamics and different degrees of complexity in non-stationary (transient) runs as well. For the example of a linearly drifting control parameter  $c$  of the Rössler system, we find that the values at which bifurcations between periodic and chaotic behavior occur in the non-stationary system do well coincide with the numerically estimated bifurcation points of the autonomous system inferred from, indicating that in the considered example, transient dynamics close to the bifurcation points does not play a major role as long as the considered RNs are still sufficiently large to obtain a reliable statistics.

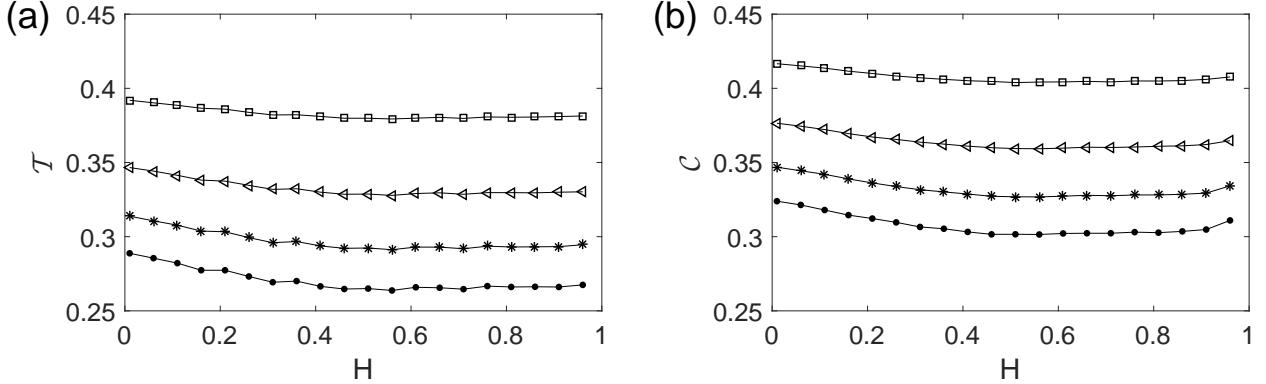


Figure 12: Dependence of (a) RN transitivity  $T$  and (b) global clustering coefficient  $C$  for fGn on the Hurst exponent  $H$  for different embedding dimensions ( $m = 3$ :  $\square$ ,  $m = 4$ :  $\triangleleft$ ,  $m = 5$ :  $*$ ,  $m = 6$ :  $\bullet$ ), taken over 200 independent realizations and using a RN edge density of  $\rho = 0.03$ . The embedding delay has been kept at the same value for all realizations with the same  $H$  according to the de-correlation time  $\tau_{0.1}$ . In all cases,  $N = 2^{12}$ . Reproduced from [150].

There is another category of non-stationary process of long-term correlated stochastic dynamics, for instance, fractional Brownian motion (fBm) as we discussed in Sec. 3.6, which needs special care when applying recurrence based network analysis. In the case of non-stationary fBm, the fundamental concepts of phase space reconstruction and low-dimensional dynamics do not apply anymore [150]. Therefore, the results presented in [151] hold only for the particular choices of the algorithmic parameters (for instance, length of time series, embeddings etc), showing limited physical interpretations. One solution to the problem could be transforming the process in a way so that it becomes stationary [150]. In recent applications to non-stationary real-world time series [127, 128], the authors have removed non-stationarities in the mean by removing averages taken within sliding windows from the data. In the particular case of fBm, the underlying stochastic process can be transformed into a stationary one by a first-order difference filter, i.e., by considering its increments  $x_{i+1} - x_i$ . The transformed series is commonly referred to as fractional Gaussian noise (fGn) in analogy with the classical Brownian motion arising from an aggregation of Gaussian innovations. Notably, fGn retains the long-range correlations and Gaussian probability density function (PDF) from the underlying fBm process.

Because of its stationarity, for fGn the embedding parameters can be chosen more properly than for fBm. Following the discussion in Sec. 3.6, we choose embedding delay  $\tau$  according to the autocorrelation function (ACF). In the case for  $H < 0.5$  where  $H$  is the Hurst exponent of the process, the estimated ACF drops to negative value at lag one resulting from subsequent values are negatively correlated for the anti-persistent process. Therefore, we choose  $\tau = 1$  for  $H < 0.5$ . In contrast, for  $H > 0.5$  we use the de-correlation time  $\tau_{0.1}$  as an estimator for embedding delay  $\tau$ , which increases with rising  $H$  as one would expect since larger  $H$  indicates a longer temporal range of correlations. The embedding dimension  $m$  is chosen via the FNN method. Unlike for fBm, our results suggest that the optimal value  $m$  rises with an increasing length of the time series. In general, considerably higher values of  $m$  are suggested than for fBm, which matches the theoretical expectations more closely. However, due to the finite sample size, we still find a vanishing FNN rate at a finite embedding dimension, which is probably related to a lack of proper neighbors when high dimensions are considered.

We have demonstrated in [27] that the RN characteristics transitivity  $T$  and global clustering coefficient  $C$  provide relevant information for characterizing the geometry of the resulted RNs, which has been numerically supported for various deterministic-chaotic systems. Furthermore, the theory presented in [131] shows that the corresponding considerations can be extended to any kind of process with a given state density  $p(\mathbf{x})$ . Here, we exemplify these considerations for the case of fGn and examine how the transitivity properties of RNs arising from such stationary long-range correlated stochastic processes depend on the characteristic Hurst exponent. From the numerical perspective, we show the dependence of the results on the embedding dimension  $m$  explicitly.

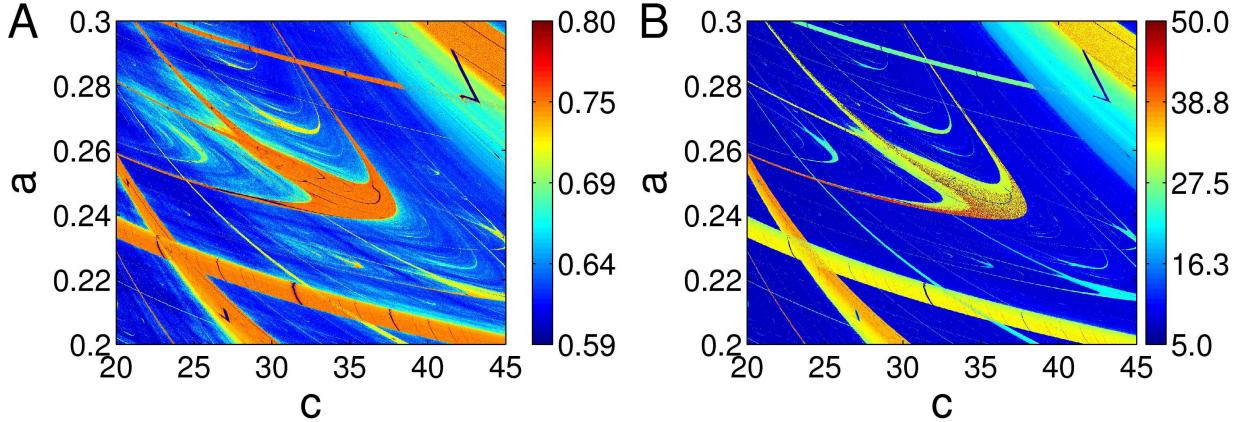


Figure 13: RN transitivity  $\hat{\mathcal{T}}$  (A) and average path length  $\hat{\mathcal{L}}$  (B) for a two-dimensional intersection ( $a = b$ ) of the three-dimensional parameter space of the Rössler system (Eq. ??), displaying “shrimp” structures (i.e., self-similar periodic windows with complex shape). For details, see [143].

For  $H > 0.5$ , Fig. 12 shows that for a given embedding dimension  $m$ , both  $\mathcal{T}$  and  $\mathcal{C}$  do not depend on  $H$ , which is expected since the  $m$ -dimensional Gaussian PDF of the process does not depend on  $H$  [131, 150]. Some minor deviation from the constant values can be observed at  $H$  close to 1, i.e., close to the non-stationary limit case represented by  $1/f$ -noise, which might be due to numerical effects [150].

For  $H < 0.5$ , both  $\mathcal{T}$  and  $\mathcal{C}$  rise with decreasing  $H$ . The reason for this behavior is that  $\tau = 1$  is the recommended, but still not “optimal” embedding delay for anti-persistent processes. Specifically, the closer  $H$  approaches 0, the stronger is the anti-correlation at lag one. This means that with the same embedding delay  $\tau = 1$ , the smaller  $H$  the stronger are the mutual correlations between the different components of the embedding vector. As a consequence, the state vectors do not form a homogeneous  $m$ -dimensional Gaussian PDF with independent components in the reconstructed phase space, but are stretched and squeezed along certain directions, so that the resulting geometric structure appears significantly lower-dimensional than  $m$ . More numerical considerations have been discussed in [150], for instance, systematical biases when  $H$  is close to 0 and finite sample size  $N$ .

*Characterize parameter space in the Rössler system.* In order to illustrate the performance of RN transitivity  $\mathcal{T}$  and average path length  $\mathcal{L}$  as tracers for qualitative changes in the dynamics of complex systems, we briefly recall results originally obtained by the authors [143]. In the latter work, the RN properties have been successfully used to discriminate between periodic and chaotic solutions in a two-dimensional subspace of the Rössler system. As Fig. 13 reveals, there are sequences of transitions between periodic and chaotic solutions. Specifically, we clearly see from the figure that the periodic windows are characterized by higher values of  $\mathcal{T}$  and  $\mathcal{L}$  than the chaotic solutions, which is in agreement with the general considerations discussed above. Specifically, for the periodic windows, we find  $\hat{\mathcal{T}}$  close to 0.75, the theoretical value for periodic dynamics (i.e., a system with effective dimension of 1).

In a similar way, we may use the RN framework for capturing the signatures of qualitative changes in the attractor’s shape and invariant density as a single control parameter is varied systematically. In a previous study using the Rössler system, we have investigated the RN properties across the transition from the classical phase-coherent Rössler attractor to the non-coherent funnel regime [144]. Our results indicate that phase coherence – in a similar spirit as fractal dimension – should be characterized from a geometric rather than a dynamics viewpoint. However, as of today there is no single RN-based index for phase coherence that has been explicitly derived from theoretical considerations.

While the aforementioned results have been obtained for stationary systems, i.e., independent realizations of the system at fixed parameter values, tracing temporal changes in dynamical complexity of non-stationary systems is another interesting field of application with numerous examples in the real-world. Using model

systems with drifting parameters such as the Lorenz [38] or Rössler systems (see Fig. ??), it is possible to systematically evaluate the performance of RN characteristics in a sliding windows framework, underlining their capabilities for discriminating between qualitatively different types of dynamics and different degrees of complexity in non-stationary (transient) runs as well. For the example of a linearly drifting control parameter  $c$  of the Rössler system (Fig. ??), we find that the values at which bifurcations between periodic and chaotic behavior occur in the non-stationary system do well coincide with the numerically estimated bifurcation points of the autonomous system inferred from Fig. 13, indicating that in the considered example, transient dynamics close to the bifurcation points does not play a major role as long as the considered RNs are still sufficiently large to obtain a reliable statistics.

coupling directions and synchronization analysis.

### 6.1.2. Real-world applications

Although much recent work on RNs and multivariate generalizations thereof has been focused on the development of the theoretical framework and its numerical exploration using simple low-dimensional model systems, there have already been some first successful applications to characterizing system's properties from experimental or observational time series. For example, the successful application of RN to predict protein structural classes has been reported in [298].

As Norbert suggested, we need 1-2 explicit examples showing the applications in this subsection. Please make suggestions.

*Applications in climatology:* . One important field of recent applications is paleoclimatology, which has already been taken as an illustrative example in the seminal paper by Marwan *et al.* [42]. The corresponding study was later extended to some systematic investigation of the temporal variability profile of RN-based complexity measures for three marine sediment records of terrigenous dust flux off Africa during the last 5 million years. Donges *et al.* [128] argued that RNs can be used for characterizing dynamics from non-uniformly sampled or age-uncertain data, since this methodological approach does not make explicit use of time information. In turn, due to the necessity of using time-delay embedding, there is implicit time information entering the analysis, which has been recognized but widely neglected in previous works. Notably, disregarding age uncertainty and sampling heterogeneity appears a reasonable approximation only in cases where the distribution of instantaneous sampling rates remains acceptably narrow.

The results of Donges *et al.* [127] pointed to the existence of spatially coherent changes in the long-term variability of environmental conditions over Africa, which have probably influenced the evolution of human ancestor species. Specifically, RN transitivity and average path length have been interpreted as indicators for “climate regularity” (i.e., the complexity of fluctuations as captured by the transitivity dimensions) and “abrupt dynamical changes”, respectively. By identifying three time intervals with consistent changes of the RN properties obtained from spatially widely separated records, it has been possible to attribute the corresponding long-term changes in the dynamics to periods characterized by known or speculated mechanisms for large-scale climate shifts such as changes in the Indian ocean circulation patterns, the intensification of the atmospheric Walker circulation, or changes in the dominant periodicity of Northern hemispheric glacial cycles. Moreover, Donges *et al.* [128] demonstrated a good robustness of the results of RN analysis obtained in a sliding windows framework when varying the corresponding parameters (e.g., window size or embedding delay) over a reasonable range.

As another methodological step towards better understanding climatic mechanisms, we have used two speleothem records for studying interdependencies between the two main branches of the Asian summer monsoon (the Indian and East Asian summer monsoon) by means of inter-system recurrence network approaches [167]. For this purpose, they selected two data sets of oxygen isotope anomalies from speleothems obtained from two caves in China and the Oman, respectively, which can be considered as proxies for the annual precipitation and, hence, the overall strength of the two monsoon branches over the last about 10,000 years. The asymmetries of the IRN cross-transitivities and global cross-clustering coefficients provided clear evidence for a marked influence of the Indian summer monsoon on the East Asian branch rather than vice versa, which is in good agreement with existing climatological theories. As a subsequent extension of this

work, we emphasize the possibility of repeating the same kind of analysis in a sliding windows framework, thereby gaining information on possible temporal changes of the associated climatic patterns during certain time periods as recently revealed using correlation-based complex network analysis applied to a larger set of speleothem records from the Asian monsoon domain [299].

In order to characterize dynamical complexity associated with more recent environmental variability, Lange and Böse [300, 301] used RQA as well as RN analysis for studying global photosynthetic activity from remote sensing data in conjunction with global precipitation patterns. Specifically, they studied 14-years long time series (1998-2011) of the fraction of absorbed photosynthetically active radiation with a spatial resolution of  $0.5^\circ$  around the Earth and a temporal sampling of about ten days, providing time series of  $N = 504$  data points. Their results revealed very interesting spatial complexity patterns, which have been largely, but not exclusively determined by the amplitude of the annual cycle of vegetation growth in different ecosystems.

*Applications in fluid dynamics:* . In a series of papers, Gao *et al.* investigated the emerging complexity of dynamical patterns in two-phase gas-liquid or oil-water flows in different configurations using RN techniques. Bifurcation scenarios from slugs to bubbles of a two phase flow of water-air occurring in a circular horizontal mini-channel has been recently analyzed by recurrence plots and recurrence network approaches in [302]. In general, multiple sensors measuring fluctuations of electrical conductance have been used for obtaining signals that are characteristic for the different flow patterns. For gas-liquid two-phase upward flows in vertical pipes, different types of complex networks generated from observational data have been proposed, among which the degree correlations (assortativity) of RNs was proven to be particularly useful for distinguishing between qualitatively different flow types [75, 303, 304]. One may also construct a directed weighted RN [305–309]. For oil-water two-phase upward flows in a similar configuration, the global clustering coefficient of RNs reveals a marked increase in dynamical complexity (detectable in terms of a decreasing  $\hat{C}$ ) as the flow pattern changes from slug flow over coarse to very finely dispersed bubble flow [185, 310]. In case of oil-water two-phase flows in inclined pipes [311], the motif distributions of RNs (specifically, the frequency distributions of small subgraphs containing exactly four vertices) revealed an increasing degree of heterogeneity, where the motif ranking was conserved in all experimental conditions, whereas the absolute motif frequency dramatically changed. The corresponding results were independently confirmed using some classical measures of complexity, which indicated increasing complexity in conjunction with increasing heterogeneity of the RN motif distributions. Finally, for characterizing horizontal oil-water flows [308], RN and inter-system RN analysis were combined for studying conductance signals from multiple sensors. Specifically, cross-transitivity was found a useful measure for tracing the transitions between stable stratified and unstable states associated with the formation of droplets. Furthermore, Gao *et al.* [186, 187, 216, 312] further extended these ideas to construct multivariate weighted recurrences networks from multi-channel measurements from different oil-water flow patterns.

*Applications in electrochemistry:* . Zou *et al.* [145] studied the complexity of experimental electrochemical oscillations as one control parameter of the experiments (temperature) was systematically varied. By utilizing a multitude of complementary RN characteristics, they could demonstrate a systematic rise in dynamical complexity as temperature increased, but an absence of a previously speculated phase transition [313] separating phase-coherent from noncoherent chaotic oscillations. The latter results were independently confirmed using other classical indicators for phase coherence, as well as studies of a corresponding mathematical model of the specific electrochemical processes.

*Applications in medicine:* . Finally, there have been a couple of successful applications in a medical context. Marwan *et al.* [314] demonstrated that the global clustering coefficients of RNs obtained from heartbeat intervals, diastolic and systolic blood pressure allowed a reliable identification of patients with pre-eclampsia, a cardiovascular disease during pregnancy with a high risk of fetal and maternal morbidity. Their results were further improved by Ramírez *et al.* [315, 316] who considered combinations of various RN-based network characteristics. In a similar spirit as for cardiovascular diseases, recent results point to the capability of RN characteristics for discriminating between the EEG signals of healthy and epileptic patients [317].

## 6.2. Visibility graphs

(H)VGs approaches have been applied so far for studying energy dissipation rates in fully developed turbulence [239, 318, 319], financial data [49, 320, 321], physiological time series [48, 199, 240, 322–324], seizure detections by EEG signals [325–327], cardiorespiratory interaction signals [328], alcoholism identification by EEG signals[329]. In the geoscientific context, [330] studied the time series of annual US landfalling hurricane counts. Subsequently, [331] presented a study on daily streamflow series from the US and China. VGs analysis has been reported in air temperature data from China [332], wind speed records from central Argentina [227], quarterly macroeconomic series of China [333]. Telesca *et al.*[245] used VGs for studying seismic activity in Italy, and Corinth rift in western central Greece [334]. Nonlinear features of seismic time series have been recently reviewed in [335]. Both  $k$  nearest neighbors network and VGs analysis show almost the same qualitative behavior and allow to reveal the underlying system dynamics in turbulent heated jets [336]. Motif structures and subgraphs of VGs have been used to human ventricular fibrillation (ECG) time series [122, 242], ECG diagnosis of epilepsy [337], air traffic flow data [338]. Simple topological measures such as the diameter, average path length, modularity, clustering coefficient, density and hierarchical organizations of networks have been used to characterize different dynamic properties between atmospheric and oceanic variables [339]. In addition, the (H)VGs have been proposed to predict catastrophes of a non-autonomous network which derived from a marine system [340], which demonstrates that the topological characteristics like average degrees of the networks do show pronounced signatures at the onset of catastrophes. Fractal characteristics of (H)VGs have been reported in fractional Brownian motions, which has been recently extended to multiparticle emission data in high energy heavy-ion collisions [341], which shows consistent power law degree distributions as compared to the results as obtained by the traditional sandbox algorithm. In [216], a slight modification of HVG algorithm has been proposed to extract the multiscale properties of time series from oil-water two phase flow signals. In the case of intermittent time series, some phenomenology theories have been obtained in order to link the laminar episodes and chaotic bursts with the connectivity of the resulting HVGs [342, 343].

Here, we illustrate two specific examples showing the applications of (H)VG analysis to time series of sunspots.

### 6.2.1. Example I: sunspot numbers

Solar activity is characterized by complex dynamics superimposed to an almost periodic, about 11 years cycle. Zou *et al.* [233] performed the VGs analysis on both the daily and monthly sunspot series. The natural VGs focus on the effects of the local maxima on the resulting graphs. In order to disclose the contributions of local minima to the VGs, they proposed two ways to construct the network: one is from the original observable measurements and the other is from a negative-inverse-transformed series. In the particular case of sunspot series, local minima play important roles in forming the increasing and decreasing phases of the solar cycles. In [233], they found that the degree distribution of the derived networks for the strong maxima has clear non-Gaussian properties, while the degree distribution for minima is bimodal.

Here, let us construct VGs for both the International Sunspot Number (ISN) [344] (see [233] for more consistent results that are based on the sunspot area (SSA) series). We perform VG analysis for both monthly and daily sunspot series, which yields, respectively, month-to-month and day-to-day correlation patterns of the sunspot activities. The degree sequence  $k_i = \sum_j A_{i,j}$  reflects the maximal visibility of the corresponding observation in comparison with its neighbors in the time series. Furthermore, the degree distribution is denoted as  $p(k)$ . In the case of sunspot time series, one is often required to investigate what contributions local minimum values make to the network – something that has been largely overlooked by the traditional VGs. One simple solution is to study the negatively inverted counterpart of the original time series, namely,  $-x(t_i)$ , which quantifies the properties of the local minima. We use  $k_{-x}$  and  $p(k_{-x})$  to denote the case of  $-x(t_i)$ . Here, we remark that this simple inversion of the time series allows us to create an entirely different complex network.

Figure 14(a,b) show the degree distributions  $p(k)$  of the VGs derived from the ISN  $x(t_i)$  with heavy-tails corresponding to hubs of the graph, which clearly deviates from Gaussian properties. In contrast,  $p(k_{-x})$  of the negatively inverted sunspot series  $-x(t_i)$  shows a completely different distribution, consisting of a

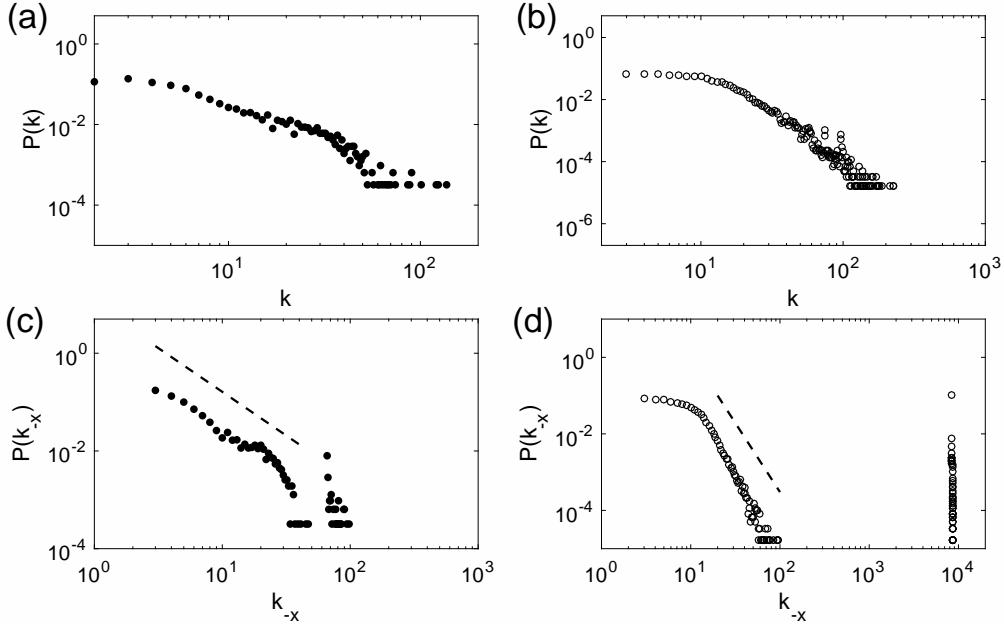


Figure 14: Degree distribution  $p(k)$  of VGs from monthly (a,c) and daily data (b,d). (a,b) is for  $x(t_i)$ , and (c,d)  $-x(t_i)$ . One would suspect a fit to the first part of  $p(k_{-x})$  yields that the slope of dashed line in (c) is 1.79, and that of (d) is 3.61, *but all p-values are 0, rejecting the hypothetical power laws.*

bimodal property (Figure 14c,d), extra large degrees are at least two orders of magnitude larger than most of the vertices (Figure 14(d)). Since well-defined scaling regimes are absent in either  $p(k)$  or  $p(k_{-x})$  (nor do they appear in the cumulative distributions, see more details of the statistical tests in [233]), we may reject the hypothetical power laws.

Based on the degree sequences  $k_x$  and  $k_{-x}$ , Zou *et al.* [233] further investigated the long term variations of local maxima/minima of the sunspot series. They found that the positions of strong maxima are largely homogeneously distributed over the time domain, while that of the strong minima are much more clustered in the time axis. This suggests that the strong active regions of sunspots appear more or less independently of each other, in contrast, the process of the strong minima has a positive long term correlation. The hidden regularity of the time positions of maxima/minima has been characterized by the waiting time distribution: the interval between two successive events is called the waiting-time. The distribution of the waiting time between two subsequent maxima sunspot deviates significantly from an exponential function [233]. In contrast, the waiting times between subsequent strong minima have a power-law distribution where a statistically significant scaling regime has been found. These results of the difference between maxima and minima could be used for evaluating models for solar activity because they reflect important properties that are not included in other measures reported in the literature.

VGs for sunspot series show rich community structures, each of which mainly consists of the temporal information of two consecutive solar cycles. The solar cycle of approximately 11-years yields that most of the temporal points of the decreasing phase of one solar cycle are connected to those points of the increasing phase of the next cycle in the resulting VGs [233]. When the sunspot number reaches a stronger but more infrequent extreme maximum, we have inter-community connections, since they have a better visibility contact with more neighbors than other time points – hence, forming hubs in the graph. The inter-community connections extend over several consecutive solar cycles encompassing the temporal cycle-to-cycle information. In Fig. 15(a), we highlight some hubs of large degrees ( $k_i > 15$ ), which have been suggested to identify solar cycles [233]. In addition, there are strong positive correlations between large degrees  $k_i$  and high betweenness  $b_i$ , which further characterizes the node's ability to transport information from one place to another along the shortest path as shown in Fig. 15(b).

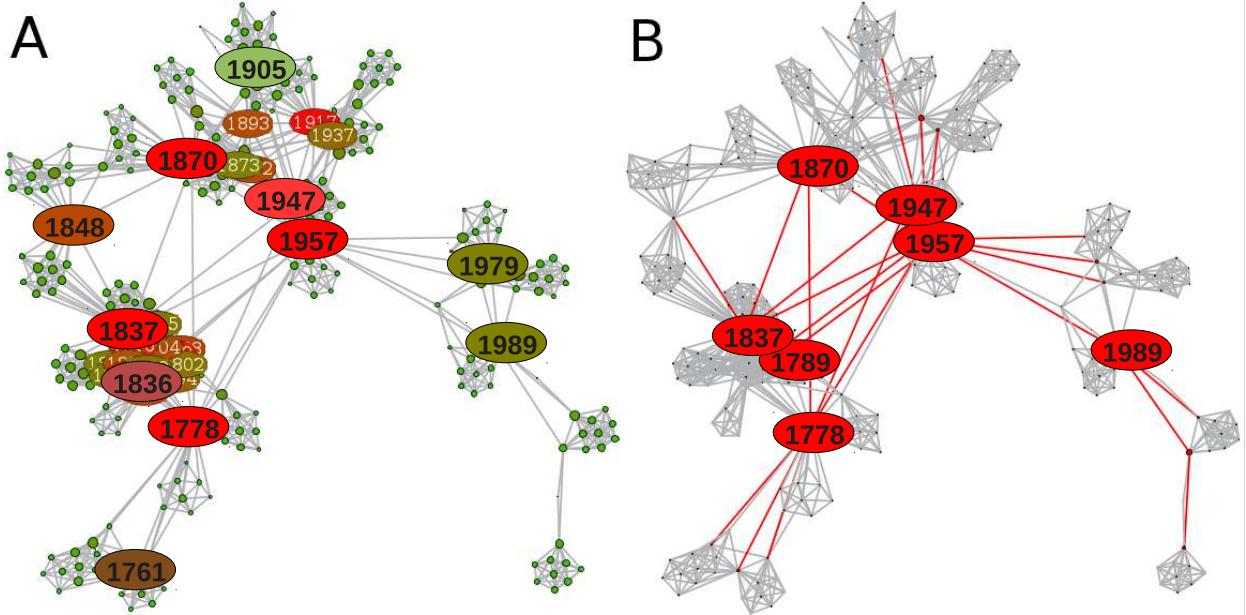


Figure 15: Network representations of the VG constructed from the annual sunspot numbers of the entire series. Highlighted visible nodes are: (A) large degrees ( $k_i > 15$ ), and (B) high betweenness centrality ( $b_i > 0.2$ ).

### 6.2.2. Example II: Asymmetry of sunspots

In the previous Sec. 6.2.1, we performed (H)VG analysis for the sunspots series. In addition, in the line of this research, another important features of sunspots is the presence of a marked, time-varying hemispheric asymmetry, which have not yet been completely uncovered. The hemispheric asymmetry of solar activity manifests itself in the statistical properties of a variety of activity indicators such as sunspot numbers, areas and spatial distribution, the numbers of flares and coronal mass ejections, solar radio and X-ray flux, etc., and has been recognized to vary on multi-decadal time scales (see, e.g.,[345–352], and references therein). Notably, it is commonly believed that the observed distinct hemispheric asymmetry is an intrinsic property associated with the underlying solar magnetic field dynamics, which in turn serves as the driver of solar activity responsible for particle and electromagnetic emissions directly affecting the Earth. Traditionally, the North–South asymmetry of solar activity has been mainly defined in terms of amplitude differences between the hemispheric values of different properties [345]. Recently, it has been argued that phase information should be explicitly taken into account as well, noting that a mutual time shift between the activity cycles observed separately at both solar hemispheres could provide a significant contribution to the observed asymmetry [347, 348]. However, even despite these methodological advances, properly quantifying the North–South asymmetry is a challenging problem by itself. Specifically, the complex dynamics of the entire solar activity cycles calls for replacing traditional linear statistical approaches by methods originated in the field of nonlinear dynamics.

In [210], Zou *et. al* proposed (H)VGs analysis to study the asymmetric distributions of the sunspots over the solar surface. They have argued that both VGs and HVGs provide complementary information on hemispheric asymmetries in dynamical properties. More specifically as we discussed in Sec. 4.4.2, the excess degree  $\Delta k(t)$  (Eq. (116)) and the relative excess degree  $\Delta_{rel}k(t)$  (Eq. (117)) have been proposed to characterize the possible asymmetric properties for (H)VGs that are reconstructed from bivariate time series, which resulted two interacting layers  $\alpha$  and  $\beta$ . These two measures are based on the computations of joint degree  $k^{joint}(t)$  (Eq. 112) and the conditional degree sequences  $k_{[\alpha],[\beta]}^Q(t)$  (Eq. 115). Here we review some results based on  $\Delta k(t)$  and  $\Delta_{rel}k(t)$  and some more details are retrieved in [210]. We emphasize that the absolute excess degree can be easily interpreted in terms of inter-hemispheric differences, whereas the relative excess degree partially corrects for the skewness effect and allows quantitatively assessing the

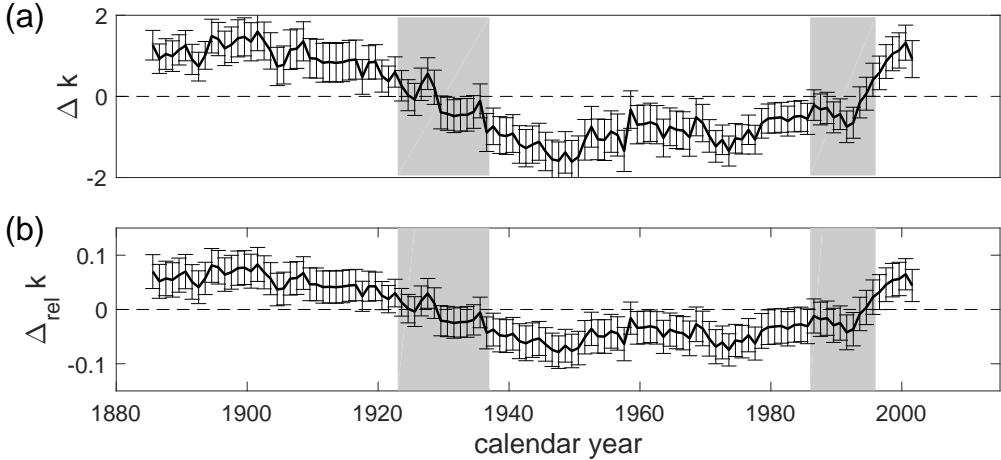


Figure 16: (a) absolute and (b) relative excess degrees obtained from the VGs of  $A_{N,S}$  computed over the sliding windows with a width of  $w = 270$  months and a mutual overlap of 12 months. Error bars display mean values and standard deviations within a given time window centered at the respective point in time. Gray areas mark those time intervals where the sign of the excess degree changes. Reproduced from [210].

relevance of differences between the degree sequences of both hemispheres.

First we construct the (H)VGs for monthly hemispheric sunspot area series  $A_N(t)$  and  $A_S(t)$ , yielding the degree sequences  $k_N(t)$  and  $k_S(t)$ , respectively. The about 11 years solar activity cycle will be clearly visible in the degree sequences and, hence, the joint and conditional degree sequences. Therefore in [210], the long-term asymmetric distribution behavior of the sunspots has been captured by utilizing a sliding window technique that averages the degree sequence over some time period. In all following considerations the window size has been chosen as  $w = 270$  months, with a mutual overlap of 12 months between subsequent time windows. This specific choice of the window size covers about one full period of the solar magnetic field polarity cycle (approximately 22 years). Note that there are no marked changes in the long-term variability of the (H)VG-based characteristics for  $w$  being between about 180 and 400 months. Figure 16 shows the mean features associated with the degree sequences for our sliding windows, together with the associated window-wise standard deviations. Our results reveal two transitions between periods of positive and negative mean (absolute and relative) excess degrees, which take place at about 1925–1935 (from higher degrees in the Northern Hemisphere to those in the Southern one) and 1985–1995 (vice versa). Furthermore, positive (negative) excess degrees imply higher mean degrees in the Northern (Southern) Hemisphere. These observations have been partially explained by the strong asymmetry of the probability distributions of sunspots over the north and south hemispheres, for instance, very high positive skewness. Notably, absolute and relative excess degrees exhibit qualitatively the same long-term variability.

The corresponding analysis by mean of HVGs is shown in Figure 17, which reveals some interesting facts: first of all, all degree-related quantities obey considerably lower values and weaker overall variability than for the VG. This is to be expected since the HVG is a subgraph of the VG. However, while the absolute degree values in the HVG typically reduce by a factor of about 2–4 in comparison with the VG, the absolute excess degrees are by more than one order of magnitude smaller (Fig. 17).

Moreover, for the HVG-based excess degree  $\Delta k(t)$  we do not find comparably clear indications for transitions between time periods with clear hemispheric predominance as for the VG (Fig. 17(a)). The only notable exception is the time period between about 1925 (corresponding to the formerly identified first transition in the VG) and 1950, where the excess degree of the HVG is significantly negative (as also observed before for the VG). Specifically, the transition in the hemispheric predominance reflected by the VGs' conditional degree sequences coincides with a sharp drop in the corresponding series for the HVG at about 1925, whereas the end of the period of significantly negative excess degrees in the HVG at about 1950 accompanies the termination of the gradual downward trend of the excess degree obtained from the

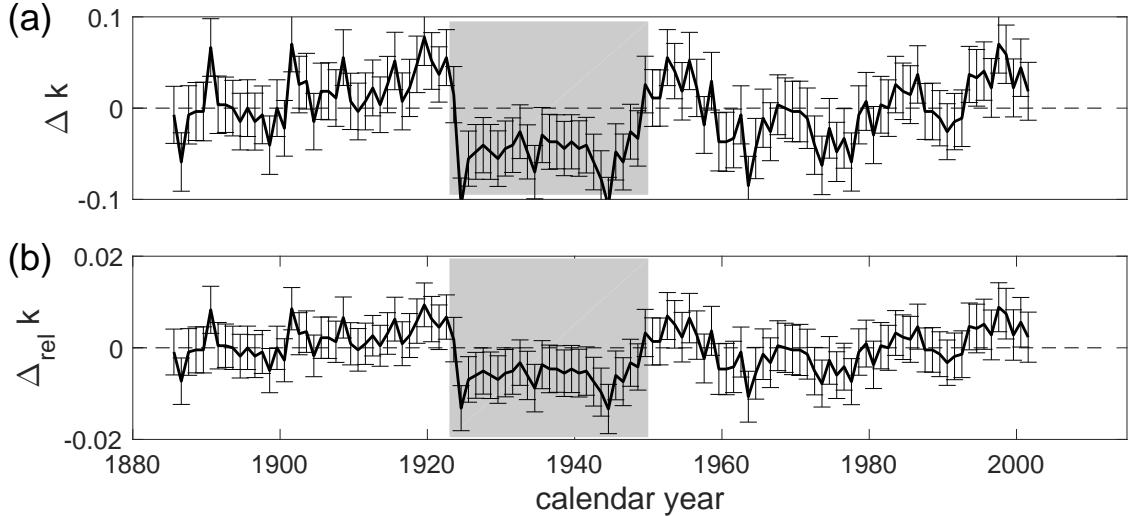


Figure 17: As in Fig. 16 for the corresponding HVG properties. The gray area indicates the only considerably long time interval with statistically significant hemispheric asymmetry of the mean conditional degree. Reproduced from [210].

VGs (Fig. 17(b)). Taken together, we interpret these findings such that the effect of the asymmetry of the hemispheric sunspot area values mostly dominates possible variations in dynamical characteristics. However, to this end we tentatively conclude that parts of the observed long-term changes of the VG-based excess degree cannot be explained by combining the corresponding changes in skewness and HVG-based excess degree (i.e., distribution and dynamics, respectively). One possible reason for this could be complex changes in the PDF of the sunspot areas, which go beyond fluctuations in skewness, but yet have a significant effect on the resulting VGs' properties.

In summary, we conclude that temporal changes in the hemispheric predominance of the graph properties lag those directly associated with the total hemispheric sunspot areas. These findings open a new dynamical perspective on studying the North–South sunspot asymmetry, which needs to be further explored in future work.

### 6.3. Transition networks

Depending on the particular symbolic representations of time series, there are various applications of transition network approaches to real time series. For instance, co-movement time series of economic growth and high-end talent development efficiency [340]. Here we focus on the application of ordinal pattern transition network approach as proposed by McCullough *et al.* in [28], where they applied this analysis to experimental time series generated by a diode resonator circuits. They argue that the network size, mean shortest path length, and network diameter are highly sensitive to the interior crisis captured in this particular data set. Meanwhile, the ordinal pattern partition networks have been reconstruct from Electrocardiogram (ECG) data from patients with a variety of heart conditions [53]. Network measures of mean degrees, entropies of the set of ordinal patterns and the number of non-occurring ordinal patterns have been computed for the resulting transition networks, showing statistically significant difference between healthy patients and several groups of unhealthy patients with varying heart conditions.

In [275], McCullough *et al* have introduced to compute both local and global out-link entropies of ordinal transition networks to quantify the complexity of temporal structure in the networks from time series. The numerical comparative investigation in the Rössler system has demonstrated that these complexity measures track dynamical changes through period doubling and periodic windows over a range of the bifurcation parameter. Furthermore, the analysis has been applied to time series of electrocardiograms (ECGs). More specifically, complexity measures are able to capture the unique properties, discriminating between

short-time ECG recordings characterized by normal sinus rhythm(NSR), ventricular tachycardia (VT) and ventricular fibrillation (VF). The global node out-link entropy of each time series is computed for both a short and a long time embedding lag and the resulting two-dimensional vector constitutes a measure of multiscale complexity description. In addition, the ordinal network analysis is performed to characterize age-related effects in interbeat interval dynamics from ECGs.

Another application of ordinal transition network method is to characterize synchronization transitions. In [55], they proposed to construct ordinal pattern partition transition networks for multivariate time series. This approach has been demonstrated to be useful for identifying dynamical regimes shifts and characterizing routes to phase synchronization. In particular, they consider three diffusively coupled Rössler systems via  $x$  component [111], which read

$$\begin{pmatrix} \dot{x}_k \\ \dot{y}_k \\ \dot{z}_k \end{pmatrix} = \begin{pmatrix} -\omega_k y_k - z_k + \sum_{l \neq k} \kappa_{k,l}(x_l - x_k) \\ \omega_k x_k + 0.165 y_k \\ 0.4 + z_k(x_k - 8.5) \end{pmatrix}, \quad (130)$$

where  $k = 1, 2, 3$  and  $\kappa$  is the coupling strength. We consider non-identical oscillators by choosing  $\omega_1 = 0.98, \omega_2 = 1.02, \omega_3 = 1.06$ . The oscillator  $k = 2$  is bidirectionally coupled to both  $k = 1$  and  $k = 3$ , whereas there is no direct coupling between  $k = 1$  and  $k = 3$ . The Eqs. (130) are numerically integrated by the fourth-order Runge Kutta method with integration step  $h = 0.01$ . We construct ordinal pattern transition networks from the  $x_k$  components, namely,  $(x_1, x_2, x_3)$  and the definitions of patterns have been summarized in Tab. 3. The results are shown in Fig. 18, which have been averaged over 50 random initial conditions when integrating Eqs. (130).

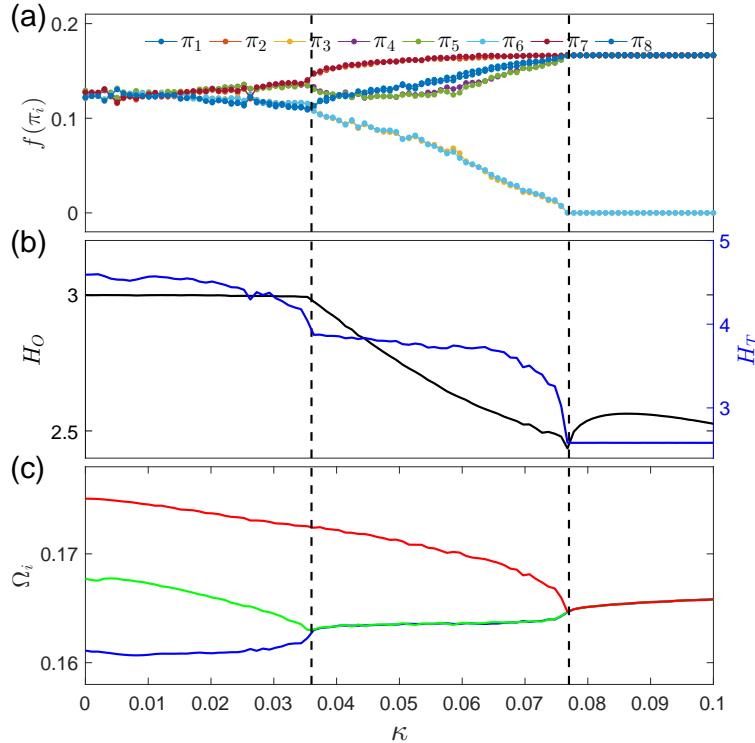


Figure 18: Phase synchronization transitions of three coupled Rössler systems. (a) frequency of each ordinal pattern  $f(\pi_i)$ , (b) entropy values  $H_O$  (Eq. (127)) and  $H_T$  (Eq. (128)), (c) mean rotation frequency  $\Omega_i$  of each oscillator. Subsystem  $k_1$  and  $k_2$  are synchronized at  $\kappa_{c1} = 0.036$ , and  $k_3$  joins the synchronization only at a stronger coupling strength  $\kappa_{c2} = 0.077$ . Both critical coupling values are highlighted by vertical dashed lines.

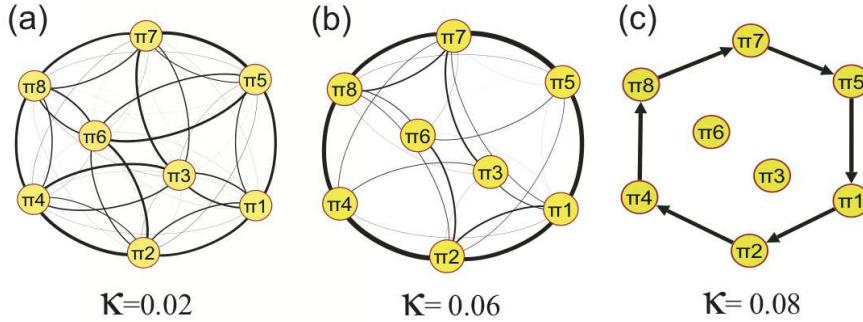


Figure 19: Ordinal transition networks on the path to phase synchronization. (a) non-sync regime of  $\kappa = 0.02 < \kappa_{c1}$ , (b) oscillators  $k = 1$  and  $k = 2$  are phase synchronized, but not with  $k = 3$ ,  $\kappa = 0.06 \in [\kappa_{c1}, \kappa_{c2}]$ , (c) all three oscillators are phase locked  $\kappa = 0.08 > \kappa_{c2}$ . Thickness of links are determined by the transition frequencies. In (a) and (b), link arrows are suppressed.

In the regime of no synchrony ( $\kappa < \kappa_{c1} = 0.036$ ), three oscillators evolve almost independently such that all ordinal patterns have the same frequencies of 0.125. There are rather small gradual changes only when  $\kappa$  approaches to  $\kappa_{c1}$  (Fig. 18(a)). The entropy value  $\mathcal{H}_T$  is more sensitive to these gradual changes showing a pronounced decreasing trend, while  $\mathcal{H}_O$  seems to be a constant (Fig. 18(b)). The average rotation frequencies  $\Omega_k$  of each oscillator are shown in (Fig. 18(c)), which confirms no synchrony in this coupling regime.

In the regime that phase synchronization appears between oscillators  $k = 1$  and  $k = 2$ , but not with  $k = 3$  ( $\kappa \in [\kappa_{c1}, \kappa_{c2}] = [0.036, 0.077]$ ), we observe monotonic increasing trends for order patterns  $\pi_1$ ,  $\pi_2$ ,  $\pi_7$ , and  $\pi_8$  (Fig. 18(a)). In addition, we find relatively slower increasing trends for patterns of  $\pi_4$  and  $\pi_5$ . In contrast, some monotonic decreasing trends are found for  $\pi_3$  and  $\pi_6$ . The changes in the frequencies of order patterns are captured by both entropy values  $\mathcal{H}_O$  and  $\mathcal{H}_T$ , showing gradual decreasing trends (Fig. 18(b)). The average rotation frequencies  $\Omega_k$  are shown in Fig. 18(c), where  $k = 1$  and  $k = 2$  are phase locked to the same rotation frequency but not with  $k = 3$ .

In the regime with all oscillators in phase synchronization ( $\kappa > \kappa_{c2} = 0.077$ ), we find that frequencies of patterns  $\pi_1$ ,  $\pi_2$ ,  $\pi_4$ ,  $\pi_5$ ,  $\pi_7$ ,  $\pi_8$  converge to the same value  $f(\pi_i) = 1/6$ , while  $\pi_3$  and  $\pi_6$  are absent (Fig. 18(a)). In other words, forbidden patterns of  $\pi_3$  and  $\pi_6$  are observed if all oscillators are synchronized. The entropy  $\mathcal{H}_O$  shows parabola-like trends (increasing first and then decreasing slowly), but  $\mathcal{H}_T$  is a constant of 2.585 (Fig. 18(b)). All mean rotation frequencies converge to the same value since three oscillators are phase locked (Fig. 18(c)).

In the process from non-synchrony to phase synchronization, the transition networks have experienced rather random transitions between all possible ordinal patterns to a state of transitions between a limited number of ordinal patterns as shown in Fig. 19. In addition, we find  $\pi_3$  and  $\pi_6$  are forbidden patterns if all three oscillators are synchronized.

## 7. Software implementation – pyunicorn (Jonathan+Norbert)

In this chapter, we introduce the Python software package **pyunicorn**, which implements methods from both complex network theory and nonlinear time series analysis, and unites these approaches in a performant, modular and flexible way [60]. Here, we mainly present a brief introduction of **pyunicorn** and a discussion of software structure and related computational issues. More details of the illustrative examples have been presented in [60]. Although in the tutorial of [60], the work flow of using **pyunicorn** is mainly illustrated drawing upon examples from climatology, the package is applicable to all fields of study where the analysis of (big) time series data is of interest, e.g. neuroscience [32, 160, 353].

**pyunicorn** is intended as an integrated container for a host of conceptionally related methods which have been developed and applied by the involved research groups for many years. Its aim is to establish a shared infrastructure for scientific data analysis by means of complex networks and nonlinear time series

analysis and it has greatly taken advantage from the backflow contributed by users all over the world. The code base has been fully open sourced under the BSD 3-Clause license.

`pyunicorn`'s capabilities for analyzing and modeling complex networks are described including general networks, spatial networks, networks of interacting networks or multiplex networks and node-weighted networks. More specifically, as shown in Fig. shows the software architecture of `pyunicorn` displayed as a Unified Modeling Language (UML) diagram of class relationships. `pyunicorn` presents methods for constructing and analyzing functional networks from fields of multiple time series, including cases demonstrating the application of climate network and coupled climate network analysis, methods for performing non-linear time series analysis using recurrence plots, recurrence networks and visibility graphs. `pyunicorn` also presents methods for generating surrogate time series, which are useful for both functional network and network-based time series analysis.

Table 6: Structure of the `pyunicorn` software package listing the most important classes belonging to each submodule (selection for brevity).

core	funcnet	climate	timeseries	utils
Network	CouplingAnalysis	ClimateNetwork	RecurrencePlot	mpi
GeoNetwork	CouplingAnalysisPurePython	CoupledClimateNetwork	CrossRecurrencePlot	navigator
InteractingNetworks		TsonisClimateNetwork	JointRecurrencePlot	
ResNetwork		SpearmanClimateNetwork	RecurrenceNetwork	
Data		MutualInfoClimateNetwork	InterSystemRecurrenceNetwork	
Grid		ClimateData	JointRecurrenceNetwork	
			VisibilityGraph	
			Surrogates	

The `pyunicorn` library is fully object-oriented and its inheritance and composition hierarchy reflects the relationships between the analysis methods. It consists of five subpackages (Tab. 6):

- core** This name space contains the basic building blocks for general network analysis and modeling and is accessible after calling `import pyunicorn`. The backbone `Network` class provides numerous standard and advanced complex network statistics, measures and generative models as well as import and export capabilities to GraphML, GML, NCOL, LGL, DOT, DIMACS and other formats. `Grid` and `GeoNetwork` extend these functionalities with respect to spatio-temporally embedded networks, which can be imported from and exported to ASCII and NetCDF files via the `Data` class. `InteractingNetworks` provides advanced methods designed for networks of networks (or multiplex networks), while `ResNetwork` specializes in power grids transporting electric currents and related infrastructure networks.
- funcnet** Advanced tools for construction and analysis of general (non-climate) functional networks will be accommodated here. So far, `CouplingAnalysis` calculates cross-correlation, mutual information, mutual sorting information and their respective surrogates for large arrays of scalar time series.
- climate** Building on top of `GeoNetwork` and `Data`, the `ClimateNetwork` class and its children facilitate the construction and analysis of functional networks representing the statistical interdependency structure within a field of time series, based on similarity measures such as lagged linear Pearson or Spearman correlation and mutual information. `CoupledClimateNetwork` extends this capability to the study of interrelationships between two distinct fields of observables.
- timeseries** This subpackage provides various tools for the analysis of non-linear dynamical systems and uni- and multivariate time series. Apart from visibility graphs with time-directed measures (`VisibilityGraph` class), the focus lies on recurrence-based methods derived from the `RecurrencePlot` class. These include single, joint and cross-recurrence plots as well as standard, joint and inter-system recurrence networks, supporting time-delay embedding and several phase space metrics and are equipped with common measures of recurrence quantification analysis. `Surrogates` allow testing for the statistical

significance of similarity measures by generating surrogate data sets under miscellaneous constraints from observable time series.

**utils** Currently this includes MPI parallelization support and an experimental interactive network navigator.

`pyunicorn` is conveniently applicable to research domains in science and society as different as neuroscience, infrastructure and climatology. Most computationally demanding algorithms are implemented in fast compiled languages on sparse data structures, allowing the performant analysis of large networks and time series data sets. The software's modular and object-oriented architecture enables the flexible and parsimonious combination of data structures, methods and algorithms from different fields. For example, combining complex network theory (`core.Network` class) and recurrence plots (`timeseries.RecurrencePlot` class) yields recurrence network analysis (`timeseries.RecurrenceNetwork` class), a versatile framework that opens new perspectives for nonlinear time series analysis and the study of complex dynamical systems in phase space. Another example are climate networks (`climate.ClimateNetwork` class), an approach bringing together ideas and concepts from complex network theory and classical eigen analysis of climatological data (e.g. empirical orthogonal function analysis) [60].

Along these lines, `pyunicorn` has the potential to facilitate future methodological developments in the fields of network theory, time series analysis and complex systems science by synthesizing existing elements and by adding new methods and classes that interact with or build upon preexisting ones. Nonetheless, we urge users of the software to ensure that such developments are theoretically well-founded and explicable as well as motivated by well-posed and relevant research questions to produce high-quality research.

Besides `pyunicorn`, we suggest that some other software packages are available, for instance, the MATLAB toolbox CRPtool that allows to get the adjacency matrix (or recurrence matrix). Furthermore, this toolbox provides the computation of several network measures like clustering coefficients  $C$  and transitivity  $T$ . In addition, we recommend some network visualization toolboxes, including Gephi or Networkx which can be used to analyze the networks (when the recurrence matrix is available).

## 8. Conclusions and future perspectives (All)

### 8.1. Conclusions and discussion

Comments by Jonathan:

- What is the added value of network methods compared to standard methods?
- Why do network methods typically seem to work better with small data sets than standard methods? Does this have to do with  $N^2$  of information units (although not independent) being created from  $N$  data points through building an adjacency matrix, leading to more robust statistics?
- What are common applications of network methods? Such as identifying regime shifts / transitions / tipping points in time series, distinguishing different dynamical regimes, analysis of dynamical system structure in phase space, prediction of some types... What more?

We might probably start with taking these points already in the introduction (as motivating questions beyond the methodological interest) and take them up for more detailed discussion at the end of the review.

For recurrence network approaches, there are some evident questions, the most relevant being about the invariance of findings under variation of the threshold value  $\varepsilon$  since this value determines the link density of the network and all network characteristics become trivial in the limit of full connectivity [2].

### 8.2. Future perspectives

Some of the following topics should be paid for more attention, which are currently under investigations. Furthermore, the questions as discussed in the previous sections should be integrated into the following working packages.

1. Evolving network analysis for time series remains largely open, for instance, by temporal network approaches.

There is one natural extension of network science from static to dynamic analysis, which build models by evolving networks that change as a function of time. The generalization to evolving network analysis for time series remains largely open, which however almost all real world networks evolve over time, either by adding or removing nodes or links over time. Evolving network concepts build on established network theory and are now being introduced into studying networks in many diverse fields [354]. Taking social networks as examples, people make and lose friends over time, thereby creating and destroying edges, and some people become part of new social networks or leave their networks, changing the nodes in the network.

The most common way of characterizing evolving networks is to treat evolving networks as successive snapshots of a static network. In analyzing time series, we often use running (sliding) window technique and network analysis is performed for each window. This set of all individual static windows compose a motion perspective over time. Dynamical changes or regime shifts are then tracked by the variations of network properties. Unfortunately, the sliding window technique to a motion picture also reveals the main difficulty with this approach, namely, empirical choices of time window lengths and the window overlapping which are rarely suggested by the network analysis. Using extremely small window sizes between two consecutive snapshot preserves resolution, but may actually obscure wider trends which only become visible over longer timescales. Conversely, using larger window sizes loses the temporal order of events within each snapshot. Therefore, we need careful choices of the appropriate timescale for dividing the evolution of a network into static snapshots, which have certain effects on proper interpretations of the results [128, 210].

Anyway, such a sliding window technique does not cover all aspects of the temporal structures of networks [354]. More generally, we need to include an additional time dimension to quantify the detailed information on the temporal sequences of the network structures, for instance, timings when the edges are active or not. The time ordering have important effects that can not be captured by static networks, especially when considering dynamical processes on top of networks. When coming to time series analysis, Weng *et al* [273] proposed to transform time series into temporal networks by encoding temporal information into an additional topological dimension of the graph, which captures the “lifetime” of edges. We note that a proper combination of ordinal pattern transition network approach (for instance, short-term transition networks) may provide the temporal information for this problem since the transition matrix describes the probability of future evolution directions of the trajectory in phase space.

## 2. Multilayer and multiplex network analysis for multiple time scale time series.

In this work, we have provided a review on the existing methods in reconstructing multilayer and multiplex networks from time series, for instance, multiplex recurrence networks, multiplex visibility graphs, inter-system cross recurrence networks, and joint recurrence networks etc. However, most of these methods are appropriate for stationary time series as we have discussed in Sec. 3.6. When monitoring complex physical systems over time, one often finds multiple phenomena in the data that work on different time scales. For example, observations are collected on a smaller time scale, and also exhibit time series behavior over a larger time scale, which is rather typical for data of climate sciences. Another prominent example from neuroscience is the recording of spiking activity of individual neurons (discrete event series) and local field potentials (time continuous measurement). Variability of the data on the smaller scale can obscure the time series behavior of the data on the larger scale, making it more difficult to identify the larger scale trends. If one is interested in analyzing and modeling these individual phenomena, it is crucial to recognize the multiple time scales in the construction of multilayer and multiplex networks from time series.

## 3. Bringing complex network methods to understanding complexity from various disciplines.

In future, we need to demonstrate the existing methods by more applications to various disciplines, particularly providing deeper insights to the existing knowledge of the respective topics. Depending on the particular working topic, it is crucial to make good use of the network analysis to extract some features that are not easily captured by most of the standard methods.

#### 4. The inverse problem of regeneration of time series from networks.

Most of the existing works focus on the proper transformation methods mapping a time series into network representations. However, regeneration of time series from a network as denoted by the adjacency matrix remains a big challenge, which certainly has many applications. In networks, the order of the vertices can be exchanged without affecting the network topology. But for this reconstruction of the trajectory the temporal order of the nodes is required. Depending on different network representations for time series, regeneration algorithms should be different. Recently these methods have been focused on mainly on recurrences [77, 78, 80],  $k$  nearest neighbor networks [81, 82], and ordinal transition networks [79]. For all these different algorithms, there are several important algorithmic parameters that have to be chosen empirically in order to guarantee consistent topology between the reconstructed time series and the original system. The computational complexity of each algorithm has to be evaluated in the future work. One application of such regeneration algorithms is to perform surrogate analysis, for example, to test the statistical significance of the results obtained for the original time series. Therefore, we have to take into account the proper choice of null hypothesis while proposing algorithms regenerating time series from networks.

#### 5. Building network models for time series prediction.

Time series modeling and forecasting has attracted a great number of researchers' attention, which is the core of nonlinear time series analysis [2, 3]. To this end, we build a proper model to forecast the system's future behavior, given a sequence of observations of one or a few time variable characteristics. Most of methods originated from nonlinear dynamics are state-space models, which build local model in 'patches' of a reconstructed state space and then use that model to make the prediction of the next point, which remains an active area of research [2]. In the case of stochastic processes, delay vector strategies have been further proposed to generalize the state-space models to short-term prediction. We note that most of existing network approaches to nonlinear time series analysis have been focusing on characterizing network features of phase space. Time series modeling and prediction have not been reported in the literature yet.

### Acknowledgements

The reported development of the recurrence network framework has been a community effort. Among other colleagues, we particularly acknowledge important contributions and discussions by Jobst Heitzig, Michael Small, Zonghua Liu, Shuguang Guan, Henk A. Dijkstra, Kira Rehfeld, Nora Molkenthin, Hanna C.H. Schultz, Alraune Zech, Jan H. Feldhoff, Aljoscha Rheinwalt, Hannes Kutza, Alexander Radebach, Alexej Gluschkow, Paul Schultz, Stefan Schinkel, Wolfram Barfuss, Marc Wiedermann, Jakob Runge, Qing Yi Feng, Liubov Tupikina, Veronika Stolbova, and **xxxxxx**. Financial support of this work has been granted by the Natural Science Foundation of Shanghai (Grant No. 17ZR1444800), the German Research Association via the IRTG 1740 "Dynamical phenomena in complex networks", the German Federal Ministry for Education and Research (Young Investigators Group CoSy-CC<sup>2</sup>, grant no. 01LN1306A), **xxxxxxxx**. Numerical codes used for estimating recurrence network properties can be found in the software package **pyunicorn** [60], which is available at <http://tocsy.pik-potsdam.de/pyunicorn.php>. The distribution includes an extensive online documentation system with the detailed API documentation also being available in the PDF format [? ]. The software description in this article as well as in the supplemental material [? ] are based on the **pyunicorn** release version 0.5.0.

- [1] V. Mayer-Schönberger, K. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Eamon Dolan/Houghton Mifflin Harcourt, 2013.
- [2] E. Bradley, H. Kantz, Nonlinear time-series analysis revisited, *Chaos* 25 (9) (2015) 097610. [arXiv:1503.07493](https://arxiv.org/abs/1503.07493), doi:10.1063/1.4917289.
- [3] H. Kantz, T. Schreiber, *Nonlinear time series analysis*, 2nd Edition, Cambridge University Press, 2004.
- [4] N. H. Packard, J. P. Crutchfield, J. D. Farmer, R. S. Shaw, Geometry from a Time Series, *Phys. Rev. Lett.* 45 (9) (1980) 712–716. doi:10.1103/PhysRevLett.45.712.
- [5] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, L. S. Tsimring, The analysis of observed chaotic data in physical systems, *Rev. Mod. Phys.* 65 (4) (1993) 1331–1392. doi:10.1103/RevModPhys.65.1331.

- [6] J. C. Sprott, Chaos and Time-Series Analysis, Oxford University Press, Oxford, 2003.
- [7] P. Grassberger, I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.* 50 (5) (1983) 346–349. doi:10.1103/PhysRevLett.50.346.
- [8] S. Boccaletti, J. Kurths, G. Osipov, D. Valladares, C. Zhou, The synchronization of chaotic systems, *Phys. Rep.* 366 (1-2) (2002) 1–101. doi:10.1016/S0370-1573(02)00137-0.
- [9] A. Pikovsky, M. Rosenblum, J. Kurths, Synchronization – A Universal Concept in Nonlinear Sciences, Cambridge University Press, 2001.
- [10] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and dynamics, *Phys. Rep.* 424 (4-5) (2006) 175–308. doi:10.1016/j.physrep.2005.10.009.
- [11] T. Stankovski, T. Pereira, P. V. E. McClintock, A. Stefanovska, Coupling functions: Universal insights into dynamical interaction mechanisms, *Rev. Mod. Phys.* 89 (4) (2017) 045001. arXiv:1706.01810, doi:10.1103/RevModPhys.89.045001.
- [12] S. Frenzel, B. Pompe, Partial Mutual Information for Coupling Analysis of Multivariate Time Series, *Phys. Rev. Lett.* 99 (20) (2007) 204101. doi:10.1103/PhysRevLett.99.204101.
- [13] M. Vejmelka, M. Paluš, Inferring the directionality of coupling with conditional mutual information, *Phys. Rev. E* 77 (2) (2008) 026214. doi:10.1103/PhysRevE.77.026214.
- [14] C. W. J. Granger, Investigating Causal Relations by Econometric Models and Cross-spectral Methods, *Econometrica* 37 (1969) 424–438.
- [15] M. Dhamala, G. Rangarajan, M. Ding, Estimating Granger Causality from Fourier and Wavelet Transforms of Time Series Data, *Phys. Rev. Lett.* 100 (1) (2008) 018701. doi:10.1103/PhysRevLett.100.018701.
- [16] T. Schreiber, Measuring Information Transfer, *Phys. Rev. Lett.* 85 (2) (2000) 461–464. doi:10.1103/PhysRevLett.85.461.
- [17] R. Q. Quiroga, J. Arnhold, P. Grassberger, Learning driver-response relationships from synchronization patterns, *Phys. Rev. E* 61 (5) (2000) 5142–5148. doi:10.1103/PhysRevE.61.5142.
- [18] M. G. Rosenblum, A. S. Pikovsky, Detecting direction of coupling in interacting oscillators, *Phys. Rev. E* 64 (4) (2001) 045202. doi:10.1103/PhysRevE.64.045202.
- [19] M. G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, R. Mrowka, Identification of coupling direction: Application to cardiorespiratory interaction, *Phys. Rev. E* 65 (4) (2002) 041909. doi:10.1103/PhysRevE.65.041909.
- [20] D. A. Smirnov, R. G. Andrzejak, Detection of weak directional coupling: Phase-dynamics approach versus state-space approach, *Phys. Rev. E* 71 (3) (2005) 036207. doi:10.1103/PhysRevE.71.036207.
- [21] M. Paluš, M. Vejmelka, Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections, *Phys. Rev. E* 75 (5) (2007) 056211. doi:10.1103/PhysRevE.75.056211.
- [22] M. C. Romano, M. Thiel, J. Kurths, C. Grebogi, Estimation of the direction of the coupling by conditional probabilities of recurrence, *Phys. Rev. E* 76 (3) (2007) 036211. doi:10.1103/PhysRevE.76.036211.
- [23] A. Bahraminasab, F. Ghasemi, A. Stefanovska, P. V. E. McClintock, H. Kantz, Direction of Coupling from Phases of Interacting Oscillators: A Permutation Information Approach, *Phys. Rev. Lett.* 100 (8) (2008) 084101. doi:10.1103/PhysRevLett.100.084101.
- [24] J. Nawrath, M. C. Romano, M. Thiel, I. Z. Kiss, M. Wickramasinghe, J. Timmer, J. Kurths, B. Schelter, Distinguishing Direct from Indirect Interactions in Oscillatory Networks with Multiple Time Scales, *Phys. Rev. Lett.* 104 (3) (2010) 038701. doi:10.1103/PhysRevLett.104.038701.
- [25] M. Ding, Y. Chen, S. L. Bressler, Granger Causality: Basic Theory and Application to Neuroscience, in: B. Schelter, M. Winterhalder, J. Timmer (Eds.), *Handb. Time Ser. Anal.*, Wiley-VCH Verlag GmbH & Co. KGaA, 2007, pp. 437–460.
- [26] K. Hlavackovaschindler, M. Palus, M. Vejmelka, J. Bhattacharya, Causality detection based on information-theoretic approaches in time series analysis, *Phys. Rep.* 441 (1) (2007) 1–46. doi:10.1016/j.physrep.2006.12.004.
- [27] R. V. Donner, J. Heitzig, J. F. Donges, Y. Zou, N. Marwan, J. Kurths, The geometry of chaotic dynamics a complex network perspective, *Eur. Phys. J. B* 84 (4) (2011) 653–672. doi:10.1140/epjb/e2011-10899-1.
- [28] M. McCullough, M. Small, T. Stemler, H. H.-C. Iu, Time lagged ordinal partition networks for capturing dynamics of continuous dynamical systems, *Chaos* 25 (5) (2015) 053101. arXiv:arXiv:1501.06656v1, doi:10.1063/1.4919075.
- [29] R. Albert, A.-L. Barabasi, Statistical mechanics of complex networks, *Rev. Mod. Phys.* 74 (1) (2002) 47–97. doi:10.1103/RevModPhys.74.47.
- [30] L. da F. Costa, F. A. Rodrigues, G. Travieso, P. R. Villas Boas, Characterization of complex networks: A survey of measurements, *Adv. Phys.* 56 (1) (2007) 167–242. doi:10.1080/00018730601170527.
- [31] M. E. J. Newman, The structure and function of complex networks, *SIAM Rev.* 45 (2) (2003) 167–256. doi:10.1137/S003614450342480.
- [32] E. Bullmore, O. Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems, *Nat. Rev. Neurosci.* 10 (2009) 186–198. doi:10.1038/nrn2575.
- [33] C. Zhou, L. Zemanova, G. Zamora, C. C. Hilgetag, J. Kurths, Hierarchical organization unveiled by functional connectivity in complex brain networks, *Phys. Rev. Lett.* 97 (23) (2006) 238103. doi:10.1103/PhysRevLett.97.238103.
- [34] C. Zhou, L. Zemanova, G. Zamora-Lopez, C. C. Hilgetag, J. Kurths, Structure-function relationship in complex brain networks expressed by hierarchical synchronization, *New J. Phys.* 9 (6) (2007) 178. doi:10.1088/1367-2630/9/6/178.
- [35] A. A. Tsonis, P. J. Roebber, The architecture of the climate network, *Physica A* 333 (2004) 497–504. doi:10.1016/j.physa.2003.10.045.
- [36] J. F. Donges, Y. Zou, N. Marwan, J. Kurths, The backbone of the climate network, *Europhys. Lett.* 87 (4) (2009) 48007. doi:10.1209/0295-5075/87/48007.
- [37] J. F. Donges, Y. Zou, N. Marwan, J. Kurths, Complex networks in climate dynamics - Comparing linear and nonlinear network construction methods, *Eur. Phys. J. ST* 174 (2009) 157–179. doi:10.1140/epjst/e2009-01098-2.
- [38] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. Xiang, J. Kurths, Recurrence-based time series

- analysis by means of complex network methods, *Int. J. Bifurc. Chaos* 21 (04) (2011) 1019–1046. [arXiv:1010.6032](#), doi:[10.1142/S0218127411029021](#).
- [39] J. Zhang, M. Small, Complex Network from Pseudoperiodic Time Series: Topology versus Dynamics, *Phys. Rev. Lett.* 96 (23) (2006) 238701. doi:[10.1103/PhysRevLett.96.238701](#).
  - [40] X. Xu, J. Zhang, M. Small, Superfamily phenomena and motifs of networks induced from time series, *Proc. Natl. Acad. Sci.* 105 (50) (2008) 19601–19605. doi:[10.1073/pnas.0806082105](#).
  - [41] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J. C. Nuno, From time series to complex networks: The visibility graph, *Proc. Natl. Acad. Sci.* 105 (13) (2008) 4972–4975. [arXiv:0810.0920](#), doi:[10.1073/pnas.0709247105](#).
  - [42] N. Marwan, J. F. Donges, Y. Zou, R. V. Donner, J. Kurths, Complex network approach for recurrence analysis of time series, *Phys. Lett. A* 373 (46) (2009) 4246–4254. [arXiv:0907.3368](#), doi:[10.1016/j.physleta.2009.09.042](#).
  - [43] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, J. Kurths, Recurrence networks—a novel paradigm for nonlinear time series analysis, *New J. Phys.* 12 (3) (2010) 033025. [arXiv:0908.3447v1](#), doi:[10.1088/1367-2630/12/3/033025](#).
  - [44] J. Zhang, J. Sun, X. Luo, K. Zhang, T. Nakamura, M. Small, Characterizing pseudoperiodic time series through the complex network approach, *Phys. D* 237 (22) (2008) 2856–2865. doi:[10.1016/j.physd.2008.05.008](#).
  - [45] Y. Yang, H. Yang, Complex network-based time series analysis, *Physica A* 387 (5–6) (2008) 1381–1386. doi:[10.1016/j.physa.2007.10.055](#).
  - [46] B. Luque, L. Lacasa, F. Ballesteros, J. Luque, Horizontal visibility graphs: Exact results for random time series, *Phys. Rev. E* 80 (4) (2009) 046103. [arXiv:1002.4526](#), doi:[10.1103/PhysRevE.80.046103](#).
  - [47] R. V. Donner, J. F. Donges, Visibility graph analysis of geophysical time series: Potentials and possible pitfalls, *Acta Geophys.* 60 (3) (2012) 589–623. doi:[10.2478/s11600-012-0032-x](#).
  - [48] L. Lacasa, B. Luque, J. Luque, J. C. Nuño, The visibility graph: A new method for estimating the Hurst exponent of fractional Brownian motion, *EPL (Europhysics Lett.)* 86 (3) (2009) 30001. doi:[10.1209/0295-5075/86/30001](#).
  - [49] X.-H. Ni, Z.-Q. Jiang, W.-X. Zhou, Degree distributions of the visibility graphs mapped from fractional Brownian motions and multifractal random walks, *Phys. Lett. A* 373 (42) (2009) 3822–3826. [arXiv:0812.2099](#), doi:[10.1016/j.physleta.2009.08.041](#).
  - [50] J. F. Donges, R. V. Donner, J. Kurths, Testing time series irreversibility using complex network methods, *EPL (Europhysics Lett.)* 102 (1) (2013) 10004. [arXiv:1211.1162](#), doi:[10.1209/0295-5075/102/10004](#).
  - [51] L. Lacasa, A. Nuñez, É. Roldán, J. M. R. Parrondo, B. Luque, Time series irreversibility: a visibility graph approach, *Eur. Phys. J. B* 85 (6) (2012) 217. [arXiv:1108.1691](#), doi:[10.1140/epjb/e2012-20809-8](#).
  - [52] G. Nicolis, A. G. Cantu, C. Nicolis, Dynamical aspects of interaction networks, *Int. J. Bifurc. Chaos* 15 (11) (2005) 3467–3480. doi:[10.1142/S0218127405014167](#).
  - [53] C. W. Kulp, J. M. Chobot, H. R. Freitas, G. D. Sprechini, Using ordinal partition transition networks to analyze ECG data, *Chaos* 26 (7) (2016) 073114. doi:[10.1063/1.4959537](#).
  - [54] C. Bandt, B. Pompe, Permutation Entropy: A Natural Complexity Measure for Time Series, *Phys. Rev. Lett.* 88 (17) (2002) 174102. doi:[10.1103/PhysRevLett.88.174102](#).
  - [55] J. Zhang, J. Zhou, M. Tang, H. Guo, M. Small, Y. Zou, Constructing ordinal partition transition networks from multivariate time series, *Sci. Rep.* 7 (1) (2017) 7795. doi:[10.1038/s41598-017-08245-x](#).
  - [56] C. W. Kulp, J. M. Chobot, B. J. Niskala, C. J. Needham, Using forbidden ordinal patterns to detect determinism in irregularly sampled time series, *Chaos* 26 (2) (2016) 023107. doi:[10.1063/1.4941674](#).
  - [57] M. McCullough, K. Sakellariou, T. Stemler, M. Small, Counting forbidden patterns in irregularly sampled time series. I. The effects of under-sampling, random depletion, and timing jitter, *Chaos* 26 (12) (2016) 123103. doi:[10.1063/1.4968551](#).
  - [58] K. Sakellariou, M. McCullough, T. Stemler, M. Small, Counting forbidden patterns in irregularly sampled time series. ii. reliability in the presence of highly irregular sampling, *Chaos* 26 (12). doi:[10.1063/1.4970483](#).
  - [59] D. Eroglu, F. H. McRobie, I. Ozken, T. Stemler, K. H. Wyrwoll, S. F. Breitenbach, N. Marwan, J. Kurths, See-saw relationship of the Holocene East Asian-Australian summer monsoon, *Nat. Commun.* 7 (2016) 1–7. doi:[10.1038/ncomms12929](#).
  - [60] J. F. Donges, J. Heitzig, B. Beronov, M. Wiedermann, J. Runge, Q. Y. Feng, L. Tupikina, V. Stolbova, R. V. Donner, N. Marwan, H. A. Dijkstra, J. Kurths, Unified functional network and nonlinear time series analysis for complex systems science: The pyunicorn package, *Chaos* 25 (11) (2015) 113101. [arXiv:1507.01571](#), doi:[10.1063/1.4934554](#).
  - [61] R. Cohen, S. Havlin, Complex networks: Structure, robustness and function, Cambridge University Press, Cambridge, 2010.
  - [62] M. E. J. Newman, Networks: An introduction, Oxford University Press, Oxford, 2010.
  - [63] D. J. Watts, S. H. Strogatz, Collective dynamics of “small-world” networks, *Nature* 393 (6684) (1998) 440–442. doi:[10.1038/30918](#).
  - [64] A. Barrat, M. Weigt, On the properties of small-world network models, *Eur. Phys. J. B* 13 (2000) 547–560. doi:[10.1007/s100510050067](#).
  - [65] P. Erdős, A. Rényi, On random graphs. i, *Publicationes Mathematicae* 6 (1959) 290?297.
  - [66] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, *Science* 286 (5439) (1999) 509–512. doi:[10.1126/science.286.5439.509](#).
  - [67] A. Clauset, C. R. Shalizi, M. E. J. Newman, Power-Law Distributions in Empirical Data, *SIAM Rev.* 51 (4) (2009) 661–703. doi:[10.1137/070710111](#).
  - [68] M. E. J. Newman, Assortative mixing in networks, *Phys. Rev. Lett.* 89 (20) (2002) 208701. doi:[10.1103/PhysRevLett.89.208701](#).
  - [69] S. Boccaletti, G. Bianconi, R. Criado, C. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, M. Zanin, The structure and dynamics of multilayer networks, *Phys. Rep.* 544 (1) (2014) 1–122.

- [doi:10.1016/j.physrep.2014.07.001](https://doi.org/10.1016/j.physrep.2014.07.001).
- [70] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, S. Havlin, Catastrophic cascade of failures in interdependent networks, *Nature* 464 (7291) (2010) 1025–1028. [doi:10.1038/nature08932](https://doi.org/10.1038/nature08932).
  - [71] J. Zhang, J. Sun, X. Xu, M. Small, Time series classification by complex network transformation, in: 2008 Int. Symp. Nonlinear Theory its Appl., 2008, pp. 90–93.
  - [72] Y. Shimada, T. Kimura, T. Ikeguchi, Analysis of Chaotic Dynamics Using Measures of the Complex Network Theory, in: Artif. Neural Networks - ICANN 2008, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 61–70. [doi:10.1007/978-3-540-87536-9\\_7](https://doi.org/10.1007/978-3-540-87536-9_7).
  - [73] J. Wu, H. Sun, Z. Gao, Mapping to Complex Networks from Chaos Time Series in the Car Following Model, in: B. Mao, Z. Tian, H. Huang, Z. Gao (Eds.), Traffic Transp. Stud., American Society of Civil Engineers, Reston, VA, 2008, pp. 397–407. [doi:10.1061/40995\(322\)37](https://doi.org/10.1061/40995(322)37).
  - [74] M. Small, J. Zhang, X. Xu, Transforming Time Series into Complex Networks, in: J. Zhou (Ed.), Complex Sci. First Int. Conf. Complex 2009. Shanghai, China, Febr. 2009. Revis. Pap. Part 2, Springer, Berlin, 2009, Ch. Transformi, pp. 2078–2089. [doi:10.1007/978-3-642-02469-6\\_84](https://doi.org/10.1007/978-3-642-02469-6_84).
  - [75] Z.-K. Gao, N.-D. Jin, Flow-pattern identification and nonlinear dynamics of gas-liquid two-phase flow in complex networks, *Phys. Rev. E* 79 (6) (2009) 1–14. [doi:10.1103/PhysRevE.79.066303](https://doi.org/10.1103/PhysRevE.79.066303).
  - [76] A. S. L. O. Campanharo, M. I. Sirer, R. D. Malmgren, F. M. Ramos, L. A. N. Amaral, Duality between Time Series and Networks, *PLoS One* 6 (8) (2011) e23378. [doi:10.1371/journal.pone.0023378](https://doi.org/10.1371/journal.pone.0023378).
  - [77] Y. Hirata, S. Horai, K. Aihara, Reproduction of distance matrices and original time series from recurrence plots and their applications, *Eur. Phys. J. Spec. Top.* 164 (1) (2008) 13–22. [doi:10.1140/epjst/e2008-00830-8](https://doi.org/10.1140/epjst/e2008-00830-8).
  - [78] Y. Hirata, A. Oda, K. Ohta, K. Aihara, Three-dimensional reconstruction of single-cell chromosome structure using recurrence plots, *Sci. Rep.* 6 (2016) 34982. [doi:10.1038/srep34982](https://doi.org/10.1038/srep34982).
  - [79] M. McCullough, K. Sakellariou, T. Stemler, M. Small, Regenerating time series from ordinal networks, *Chaos* 27 (3) (2017) 035814. [doi:10.1063/1.4978743](https://doi.org/10.1063/1.4978743).
  - [80] M. Thiel, M. C. Romano, J. Kurths, How much information is contained in a recurrence plot?, *Phys. Lett. A* 330 (5) (2004) 343–349. [doi:10.1016/j.physleta.2004.07.050](https://doi.org/10.1016/j.physleta.2004.07.050).
  - [81] A. Khor, M. Small, Examining k-nearest neighbour networks: Superfamily phenomena and inversion, *Chaos* 26 (4) (2016) 043101. [doi:10.1063/1.4945008](https://doi.org/10.1063/1.4945008).
  - [82] L. Hou, M. Small, S. Lao, Dynamical Systems Induced on Networks Constructed from Time Series, *Entropy* 17 (12) (2015) 6433–6446. [doi:10.3390/e17096433](https://doi.org/10.3390/e17096433).
  - [83] J. Liu, S. T. Shi, J. C. Zhao, Comparison study of typical algorithms for reconstructing time series from the recurrence plot of dynamical systems, *Chinese Phys. B* 22 (1) (2013) 1–7. [doi:10.1088/1674-1056/22/1/010505](https://doi.org/10.1088/1674-1056/22/1/010505).
  - [84] J. P. Zbilut, A. Giuliani, C. L. Webber Jr., Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification, *Phys. Lett. A* 246 (1-2) (1998) 122 – 128. [doi:10.1016/S0375-9601\(98\)00457-5](https://doi.org/10.1016/S0375-9601(98)00457-5).
  - [85] M. C. Romano, M. Thiel, J. Kurths, W. von Bloh, Multivariate Recurrence Plots, *Physics Letters A* 330 (3–4) (2004) 214–223. [doi:10.1016/j.physleta.2004.07.066](https://doi.org/10.1016/j.physleta.2004.07.066).
  - [86] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems, *Europhys. Lett.* 4 (9) (1987) 973–977. [doi:10.1209/0295-5075/4/9/004](https://doi.org/10.1209/0295-5075/4/9/004).
  - [87] N. Marwan, M. C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems, *Phys. Rep.* 438 (5–6) (2007) 237–329. [doi:10.1016/j.physrep.2006.11.001](https://doi.org/10.1016/j.physrep.2006.11.001).
  - [88] J. P. Zbilut, C. L. Webber Jr., Embeddings and delays as derived from quantification of recurrence plots, *Physics Letters A* 171 (3–4) (1992) 199–203. [doi:10.1016/0375-9601\(92\)90426-M](https://doi.org/10.1016/0375-9601(92)90426-M).
  - [89] L. L. Trulla, A. Giuliani, J. P. Zbilut, C. L. Webber Jr., Recurrence quantification analysis of the logistic equation with transients, *Physics Letters A* 223 (4) (1996) 255–260. [doi:10.1016/S0375-9601\(96\)00741-4](https://doi.org/10.1016/S0375-9601(96)00741-4).
  - [90] A. M. Fraser, H. L. Swinney, Independent coordinates for strange attractors from mutual information, *Phys. Rev. A* 33 (2) (1986) 1134–1140. [doi:10.1103/PhysRevA.33.1134](https://doi.org/10.1103/PhysRevA.33.1134).
  - [91] M. B. Kennel, R. Brown, H. D. I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Phys. Rev. A* 45 (6) (1992) 3403–3411. [doi:10.1103/PhysRevA.45.3403](https://doi.org/10.1103/PhysRevA.45.3403).
  - [92] F. Takens, Detecting strange attractors in turbulence, in: D. A. Rand, L.-S. Young (Eds.), *Dynamical Systems and Turbulence*, Warwick 1980, Vol. 898 of Lecture Notes in Mathematics, Springer, New York, 1981, pp. 366–381. [doi:10.1007/BFb0091924](https://doi.org/10.1007/BFb0091924).
  - [93] E. Ott, *Chaos in Dynamical Systems*, 2nd Edition, Cambridge University Press, Cambridge, 2002.
  - [94] H. Poincaré, Sur la problème des trois corps et les équations de la dynamique, *Acta Mathematica* 13 (1) (1890) A3–A270.
  - [95] S. Horai, T. Yamada, K. Aihara, Determinism analysis with iso-directional recurrence plots, *IEEJ Transactions on Electronics, Information and Systems* 122 (1) (2002) 141–147.
  - [96] N. Marwan, M. Thiel, N. R. Nowaczyk, Cross Recurrence Plot Based Synchronization of Time Series, *Nonlinear Processes in Geophysics* 9 (3/4) (2002) 325–331.
  - [97] N. Marwan, S. Schinkel, J. Kurths, Significance for a recurrence based transition analysis, in: Proceedings of the 2008 International Symposium on Nonlinear Theory and its Applications NOLTA??08, Budapest, Hungary, September 7-10, 2008, 2008, pp. 412–415.
  - [98] A. Facchini, C. Mocenni, N. Marwan, A. Vicino, E. Tiezzi, Nonlinear time series analysis of dissolved oxygen in the Orbetello Lagoon (Italy), *Ecol. Modell.* 203 (2007) 339–348.
  - [99] G. Litak, J. T. Sawicki, R. Kasparek, Cracked rotor detection by recurrence plots, *Nondestruct. Test. Eval.* 24 (2009) 347–351. [doi:10.1080/10589750802570836](https://doi.org/10.1080/10589750802570836).
  - [100] N. Marwan, M. H. Trauth, M. Vuille, J. Kurths, Comparing modern and Pleistocene ENSO-like influences in NW

- Argentina using nonlinear time series analysis methods, *Clim. Dyn.* 21 (2003) 317–326.
- [101] N. Marwan, J. Kurths, P. Saparin, Generalised Recurrence Plot Analysis for Spatial Data, *Phys. Lett. A* 360 (2007) 545–551.
- [102] A. Giuliani, R. Benigni, J. P. Zbilut, C. Webber, P. Sirabella, A. Colosimo, Nonlinear signal analysis methods in the elucidation of protein sequence–structure relationships, *Chemical Reviews* 102 (5) (2002) 1471–1492. doi:10.1021/cr0101499.
- [103] J. P. Zbilut, J. C. Mitchell, A. Giuliani, N. Marwan, C. L. Webber Jr., Singular Hydrophobicity Patterns and Net Charge: A Mesoscopic Principle for Protein Aggregation/Folding, *Physica A* 343 (2004) 348–358.
- [104] N. Marwan, A Historical Review of Recurrence Plots, *Eur. Phys. J. – Spec. Top.* 164 (2008) 3–12.
- [105] C. L. Webber Jr., N. Marwan, A. Facchini, A. Giuliani, Simpler methods do it better: Success of Recurrence Quantification Analysis as a general purpose data analysis tool, *Phys. Lett. A* 373 (2009) 3753–3756.
- [106] M. Thiel, M. C. Romano, P. L. Read, J. Kurths, Estimation of dynamical invariants without embedding by recurrence plots, *Chaos* 14 (2) (2004) 234–243. doi:10.1063/1.1667633.
- [107] N. Marwan, M. Thiel, N. R. Nowaczyk, Cross Recurrence Plot Based Synchronization of Time Series, *Nonlinear Process. Geophys.* 9 (2002) 325–331.
- [108] M. C. Romano, M. Thiel, J. Kurths, I. Z. Kiss, J. Hudson, Detection of synchronization for non-phase-coherent and non-stationary data, *Europhysics Letters* 71 (3) (2005) 466–472. doi:10.1209/epl/i2005-10095-1.
- [109] M. C. Romano, M. Thiel, J. Kurths, C. Grebogi, Estimation of the direction of the coupling by conditional probabilities of recurrence, *Phys. Rev. E* 76 (3) (2007) 036211. doi:10.1103/PhysRevE.76.036211.
- [110] P. Van Leeuwen, D. Geue, M. Thiel, D. Cysarz, S. Lange, M. C. Romano, N. Wessel, J. Kurths, D. H. Grnemeyer, Influence of paced maternal breathing on fetal – maternal heart rate coordination, *Proc. Natl. Acad. Sci.* 106 (2009) 13661–13666. doi:10.1073/pnas.0901049106.
- [111] J. Nawrath, M. C. Romano, M. Thiel, I. Z. Kiss, M. Wickramasinghe, J. Timmer, J. Kurths, B. Schelter, Distinguishing Direct from Indirect Interactions in Oscillatory Networks with Multiple Time Scales, *Phys. Rev. Lett.* 104 (3) (2010) 1–4. doi:10.1103/PhysRevLett.104.038701.
- [112] T. D. Pham, Fuzzy recurrence plots, *EPL (Europhysics Lett.)* 116 (5) (2016) 50008. doi:10.1209/0295-5075/116/50008.
- [113] P. beim Graben, A. Hutt, Detecting Recurrence Domains of Dynamical Systems by Symbolic Dynamics, *Phys. Rev. Lett.* 110 (15) (2013) 154101. doi:10.1103/PhysRevLett.110.154101.
- [114] D. G. d. B. Costa, B. M. d. F. Reis, Y. Zou, M. G. Quiles, E. E. N. Macau, Recurrence Density Enhanced Complex Networks for Nonlinear Time Series Analysis, *Int. J. Bifurc. Chaos* 28 (01) (2018) 1850008. doi:10.1142/S0218127418500086.
- [115] T. Martinović, G. Zitzlsberger, Highly scalable algorithm for computation of recurrence quantitative analysis, *J. Supercomput.* doi:10.1007/s11227-018-2350-5.
- [116] H. Yang, G. Liu, Self-organized topology of recurrence-based complex networks, *Chaos* 23 (4) (2013) 043116. doi:10.1063/1.4829877.
- [117] C. Liu, W.-X. Zhou, Superfamily classification of nonstationary time series based on DFA scaling exponents, *J. Phys. A Math. Theor.* 43 (49). arXiv:0912.2016, doi:10.1088/1751-8113/43/49/495005.
- [118] Y. Li, H. Cao, Y. Tan, A comparison of two methods for modeling large-scale data from time series as complex networks, *AIP Adv.* 1 (1) (2011) 012103. doi:10.1063/1.3556121.
- [119] Y. Li, H. Caö, Y. Tan, Novel method of identifying time series based on network graphs, *Complexity* 17 (1) (2011) 13–34. doi:10.1002/cplx.20384.
- [120] H. Cao, Y. Li, Unraveling chaotic attractors by complex networks and measurements of stock market complexity, *Chaos* 24 (1) (2014) 013134. doi:10.1063/1.4868258.
- [121] X. Fan, L. Wang, H. Xu, S. Li, L. Tian, Characterizing air quality data from complex network perspective, *Environ. Sci. Pollut. Res.* doi:10.1007/s11356-015-5596-y.
- [122] X. Li, D. Yang, X. Liu, X.-M. Wu, Bridging Time Series Dynamics and Complex Network Theory with Application to Electrocardiogram Analysis, *IEEE Circuits Syst. Mag.* 12 (4) (2012) 33–46. doi:10.1109/MCAS.2012.2221521.
- [123] C. S. Daw, C. E. A. Finney, E. R. Tracy, A review of symbolic analysis of experimental data, *Rev. Scient. Instrum.* 74 (2) (2003) 915–930. doi:10.1063/1.1531823.
- [124] R. Donner, U. Hinrichs, B. Scholz-Reiter, Symbolic recurrence plots: A new quantitative framework for performance analysis of manufacturing networks, *Eur. Phys. J. ST* 164 (2008) 85–104. doi:10.1140/epjst/e2008-00836-2.
- [125] P. Faure, A. Lesne, Recurrence plots for symbolic sequences, *Int. J. Bifurcation Chaos* 20 (6) (2010) 1731–1749. doi:10.1142/S0218127410026794.
- [126] T. D. Pham, From fuzzy recurrence plots to scalable recurrence networks of time series, *EPL (Europhysics Lett.)* 118 (2) (2017) 20003. doi:10.1209/0295-5075/118/20003.
- [127] J. F. Donges, R. V. Donner, M. H. Trauth, N. Marwan, H.-J. Schellnhuber, J. Kurths, Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution, *Proc. Natl. Acad. Sci.* 108 (51) (2011) 20422–20427. doi:10.1073/pnas.1117052108.
- [128] J. F. Donges, R. V. Donner, K. Rehfeld, N. Marwan, M. H. Trauth, J. Kurths, Identification of dynamical transitions in marine palaeoclimate records by recurrence network analysis, *Nonlinear Process. Geophys.* 18 (5) (2011) 545–562. doi:10.5194/npg-18-545-2011.
- [129] M. C. Casdagli, Recurrence plots revisited, *Physica D* 108 (1997) 12–44.
- [130] K. Iwayama, Y. Hirata, K. Takahashi, K. Watanabe, K. Aihara, H. Suzuki, Characterizing global evolutions of complex systems via intermediate network representations, *Sci. Rep.* 2 (1) (2012) 423. doi:10.1038/srep00423.
- [131] J. F. Donges, J. Heitzig, R. V. Donner, J. Kurths, Analytical framework for recurrence network analysis of time series, *Phys. Rev. E* 85 (4). arXiv:arXiv:1203.4701v1, doi:10.1103/PhysRevE.85.046105.

- [132] M. Thiel, M. C. Romano, J. Kurths, M. Rolfs, R. Kliegl, Twin surrogates to test for complex synchronisation, *Europhys. Lett.* 75 (4) (2006) 535–541. doi:10.1209/epl/i2006-10147-0.
- [133] M. C. Romano, M. Thiel, J. Kurths, K. Mergenthaler, R. Engbert, Hypothesis test for synchronization: Twin surrogates revisited, *Chaos* 19 (1) (2009) 015108. doi:10.1063/1.3072784.
- [134] M. Penrose, *Random Geometric Graphs*, Oxford University Press, Oxford, 2003.
- [135] C. Herrmann, M. Barthélémy, P. Provero, Connectivity distribution of spatial networks, *Phys. Rev. E* 68 (2) (2003) 026128. doi:10.1103/PhysRevE.68.026128.
- [136] J. Dall, M. Christensen, Random geometric graphs, *Phys. Rev. E* 66 (1) (2002) 016121. doi:10.1103/PhysRevE.66.016121.
- [137] J. P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.* 57 (1985) 617–656. doi:10.1103/RevModPhys.57.617.
- [138] Y. Zou, J. Heitzig, R. V. Donner, J. F. Donges, J. D. Farmer, R. Meucci, S. Euzzor, N. Marwan, J. Kurths, Power-laws in recurrence networks from dynamical systems, *EPL (Europhysics Lett.)* 98 (4) (2012) 48001. arXiv:1203.3345, doi:10.1209/0295-5075/98/48001.
- [139] R. Jacob, K. P. Harikrishnan, R. Misra, G. Ambika, Measure for degree heterogeneity in complex networks and its application to recurrence network analysis, *R. Soc. Open Sci.* 4 (1) (2016) 160757. arXiv:1605.06607, doi:10.1098/rsos.160757.
- [140] S. Milgram, Small-world problem, *Psychology Today* 1 (1) (1967) 61–67.
- [141] R. V. Donner, J. F. Donges, Y. Zou, N. Marwan, J. Kurths, Recurrence-based evolving networks for time series analysis of complex systems, *Proc. NOLTA 2010* (2010) 87–90.
- [142] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, J. Kurths, Ambiguities in recurrence-based complex network representations of time series, *Phys. Rev. E* 81 (1) (2010) 015101. arXiv:1001.4620, doi:10.1103/PhysRevE.81.015101.
- [143] Y. Zou, R. V. Donner, J. F. Donges, N. Marwan, J. Kurths, Identifying complex periodic windows in continuous-time dynamical systems using recurrence-based methods, *Chaos* 20 (4) (2010) 043130. arXiv:1011.5172, doi:10.1063/1.3523304.
- [144] Y. Zou, R. V. Donner, J. Kurths, Geometric and dynamic perspectives on phase-coherent and noncoherent chaos, *Chaos* 22 (1) (2012) 013115. arXiv:arXiv:1202.4823v1, doi:10.1063/1.3677367.
- [145] Y. Zou, R. V. Donner, M. Wickramasinghe, I. Z. Kiss, M. Small, J. Kurths, Phase coherence and attractor geometry of chaotic electrochemical oscillators, *Chaos* 22 (3) (2012) 033130. doi:10.1063/1.4747707.
- [146] P. Faure, H. Korn, A new method to estimate the Kolmogorov entropy from recurrence plots: its application to neuronal signals, *Physica D* 122 (1-4) (1998) 265–279. doi:10.1016/S0167-2789(98)00177-8.
- [147] P. Grassberger, Generalized dimensions of strange attractors, *Phys. Lett. A* 97 (6) (1983) 227–230. doi:10.1016/0375-9601(83)90753-3.
- [148] D. P. Lathrop, E. J. Kostelich, Characterization of an experimental strange attractor by periodic-orbits, *Phys. Rev. A* 40 (7) (1989) 4028–4031. doi:10.1103/PhysRevA.40.4028.
- [149] F. Strozzi, J. M. Zaldívar, K. Poljansek, F. Bono, E. Gutiérrez, From complex networks to time series analysis and viceversa : application to metabolic networks complex networks time series, Tech. rep., European Commission Joint Research Centre Contact (2009). doi:10.2788/25588.
- [150] Y. Zou, R. V. Donner, J. Kurths, Analyzing long-term correlated stochastic processes by means of recurrence networks: Potentials and pitfalls, *Phys. Rev. E* 91 (2) (2015) 022926. arXiv:1409.3613, doi:10.1103/PhysRevE.91.022926.
- [151] J.-L. Liu, Z.-G. Yu, V. Anh, Topological properties and fractal analysis of a recurrence network constructed from fractional Brownian motions, *Phys. Rev. E* 89 (3) (2014) 032814. doi:10.1103/PhysRevE.89.032814.
- [152] S. Schinkel, O. Dimigen, N. Marwan, Selection of recurrence threshold for signal detection, *Eur. Phys. J. Spec. Top.* 164 (1) (2008) 45–53. doi:10.1140/epjst/e2008-00833-5.
- [153] D. Yang, W.-X. Ren, Y.-D. Hu, D. Li, Selection of optimal threshold to construct recurrence plot for structural operational vibration measurements, *J. Sound Vib.* 349 (2015) 361–374. doi:10.1016/j.jsv.2015.03.046.
- [154] R. Jacob, K. P. Harikrishnan, R. Misra, G. Ambika, Uniform framework for the recurrence-network analysis of chaotic time series, *Phys. Rev. E* 93 (1) (2016) 012202. arXiv:arXiv:1502.03527v1, doi:10.1103/PhysRevE.93.012202.
- [155] D. Eroglu, N. Marwan, S. Prasad, J. Kurths, Finding recurrence networks' threshold adaptively for a specific time series, *Nonlinear Process. Geophys.* 21 (6) (2014) 1085–1092. doi:10.5194/npg-21-1085-2014.
- [156] M. Lin, X. X. Fan, G. Wang, G. Zhao, Network structure entropy and its dynamical evolution for recurrence networks from earthquake magnitude time series, *Eur. Phys. J. B* 89 (5) (2016) 131. doi:10.1140/epjb/e2016-70004-0.
- [157] Y. Zou, R. V. Donner, M. Thiel, J. Kurths, Disentangling regular and chaotic motion in the standard map using complex network analysis of recurrences in phase space, *Chaos* 26 (2) (2016) 023120. arXiv:arXiv:1603.06388v1, doi:10.1063/1.4942584.
- [158] I. Vega, C. Schütte, T. Conrad, Finding metastable states in real-world time series with recurrence networks, *Physica A* 445 (2016) 1–17. arXiv:1404.7807, doi:10.1016/j.physa.2015.10.041.
- [159] M. Thiel, M. C. Romano, J. Kurths, R. Meucci, E. Allaria, F. T. Arecchi, Influence of observational noise on the recurrence quantification analysis, *Phys. D* 171 (2002) 138–152.
- [160] N. P. Subramanyam, J. Hyttinen, Characterization of dynamical systems under noise using recurrence networks: Application to simulated and EEG data, *Phys. Lett. A* 378 (46) (2014) 3464–3474. doi:10.1016/j.physleta.2014.10.005.
- [161] R. Jacob, K. Harikrishnan, R. Misra, G. Ambika, Recurrence network measures for hypothesis testing using surrogate data: Application to black hole light curves, *Commun. Nonlinear Sci. Numer. Simul.* 54 (2018) 84–99. doi:10.1016/j.cnsns.2017.05.018.
- [162] D. Eroglu, N. Marwan, M. Stebich, J. Kurths, Multiplex recurrence networks, *Phys. Rev. E* 97 (1) (2018) 012312. doi:10.1103/PhysRevE.97.012312.
- [163] L. Lacasa, V. Nicosia, V. Latora, Network structure of multivariate time series, *Sci. Rep.* 5 (1) (2015) 15508.

- [arXiv:1408.0925](https://arxiv.org/abs/1408.0925), doi:10.1038/srep15508.
- [164] J. F. Donges, H. C. Schultz, N. Marwan, Y. Zou, J. Kurths, Investigating the topology of interacting networks: Theory and application to coupled climate subnetworks, *Eur. Phys. J. B* 84 (4) (2011) 635–651. [arXiv:1102.3067](https://arxiv.org/abs/1102.3067), doi:10.1140/epjb/e2011-10795-8.
- [165] M. Wiedermann, J. F. Donges, J. Heitzig, J. Kurths, Node-weighted interacting network measures improve the representation of real-world complex systems, *EPL (Europhysics Lett.)* 102 (2) (2013) 28007. [arXiv:1301.0805](https://arxiv.org/abs/1301.0805), doi:10.1209/0295-5075/102/28007.
- [166] J. Feldhoff, Multivariate extensions of recurrence network analysis, Master's thesis, Humboldt University, Berlin (2011).
- [167] J. H. Feldhoff, R. V. Donner, J. F. Donges, N. Marwan, J. Kurths, Geometric detection of coupling directions by means of inter-system recurrence networks, *Phys. Lett. A* 376 (46) (2012) 3504–3513. [arXiv:1301.0934](https://arxiv.org/abs/1301.0934), doi:10.1016/j.physleta.2012.10.008.
- [168] J. H. Feldhoff, R. V. Donner, J. F. Donges, N. Marwan, J. Kurths, Geometric signature of complex synchronisation scenarios, *EPL (Europhysics Lett.)* 102 (3) (2013) 30007. [arXiv:1301.0806](https://arxiv.org/abs/1301.0806), doi:10.1209/0295-5075/102/30007.
- [169] R. Guimerà, M. Sales-Pardo, L. A. N. Amaral, Module identification in bipartite and directed networks, *Phys. Rev. E* 76 (3) (2007) 036102. doi:10.1103/PhysRevE.76.036102.
- [170] M. Kitsak, D. Krioukov, Hidden variables in bipartite networks, *Phys. Rev. E* 84 (2011) 026114. doi:10.1103/PhysRevE.84.026114.
- [171] J.-L. Guillaume, M. Latapy, Bipartite structure of all complex networks, *Inform. Process. Lett.* 90 (5) (2004) 215–221. doi:10.1016/j.ipl.2004.03.007.
- [172] J.-L. Guillaume, M. Latapy, Bipartite graphs as models of complex networks, *Physica A* 371 (2) (2006) 795–813. doi:10.1016/j.physa.2006.04.047.
- [173] S. Fortunato, Community detection in graphs, *Phys. Rep.* 486 (3-5) (2010) 75–174. doi:10.1016/j.physrep.2009.11.002.
- [174] M. J. Barber, Modularity and community detection in bipartite networks, *Phys. Rev. E* 76 (6) (2007) 066102. doi:10.1103/PhysRevE.76.066102.
- [175] T. Murata, Detecting communities from bipartite networks based on bipartite modularities, in: Proceedings of the 2009 International Conference on Computational Science and Engineering - Volume 04, IEEE Computer Society, Washington, DC, USA, 2009, pp. 50–57. doi:10.1109/CSE.2009.81.
- [176] K. Suzuki, K. Wakita, Extracting multi-facet community structure from bipartite networks, in: Proceedings of the 2009 International Conference on Computational Science and Engineering - Volume 04, IEEE Computer Society, Washington, DC, USA, 2009, pp. 312–319. doi:10.1109/CSE.2009.451.
- [177] N. Du, B. Wang, B. Wu, Y. Wang, Overlapping community detection in bipartite networks, in: Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01, IEEE Computer Society, Washington, DC, USA, 2008, pp. 176–179. doi:10.1109/WIAT.2008.98.
- [178] S. Lehmann, M. Schwartz, L. K. Hansen, Biclique communities, *Phys. Rev. E* 78 (1) (2008) 016108. doi:10.1103/PhysRevE.78.016108.
- [179] E. N. Sawardecker, C. A. Amundsen, M. Sales-Pardo, L. A. N. Amaral, Comparison of methods for the detection of node group membership in bipartite networks, *Eur. Phys. J. B* 72 (2009) 671–677. doi:10.1140/epjb/e2009-00397-6.
- [180] P. G. Lind, M. C. González, H. J. Herrmann, Cycles and clustering in bipartite networks, *Phys. Rev. E* 72 (5) (2005) 056127. doi:10.1103/PhysRevE.72.056127.
- [181] P. Zhang, J. Wang, X. Li, M. Li, Z. Di, Y. Fan, Clustering coefficient and community structure of bipartite networks, *Physica A* 387 (27) (2008) 6869–6875. doi:10.1016/j.physa.2008.09.006.
- [182] M. E. J. Newman, Detecting community structure in networks, *Eur. Phys. J. B* 38 (2) (2004) 321–330. doi:10.1140/epjb/e2004-00124-y.
- [183] J. Donges, Functional network macroscopes for probing past and present earth system dynamics: Complex hierarchical interactions, tipping points, and beyond, Ph.D. thesis, Humboldt University, Berlin (2012).
- [184] Y. Zou, M. C. Romano, M. Thiel, N. Marwan, J. Kurths, Inferring indirect coupling by means of recurrences, *Int. J. Bifurc. Chaos* 21 (04) (2011) 1099–1111. doi:10.1142/S0218127411029033.
- [185] Z.-K. Gao, X.-W. Zhang, N.-D. Jin, N. Marwan, J. Kurths, Multivariate recurrence network analysis for characterizing horizontal oil-water two-phase flow, *Phys. Rev. E* 88 (3) (2013) 1–12. doi:10.1103/PhysRevE.88.032910.
- [186] Z.-K. Gao, Y.-X. Yang, Q. Cai, S.-S. Zhang, N.-D. Jin, Multivariate weighted recurrence network inference for uncovering oil-water transitional flow behavior in a vertical pipe, *Chaos* 26 (6) (2016) 063117. doi:10.1063/1.4954271.
- [187] Z.-K. Gao, Y.-X. Yang, L.-S. Zhai, W.-D. Dang, J.-L. Yu, N.-D. Jin, Multivariate multiscale complex network analysis of vertical upward oil-water two-phase flow in a small diameter pipe, *Sci. Rep.* 6 (1) (2016) 20052. doi:10.1038/srep20052.
- [188] M. Paluš, Testing for nonlinearity using redundancies: quantitative and qualitative aspects, *Physica D* 80 (1995) 186–205. doi:10.1016/0167-2789(95)90079-9.
- [189] D. Prichard, J. Theiler, Generalized redundancies for time series analysis, *Physica D* 84 (1995) 476–???. doi:10.1016/0167-2789(95)00041-2.
- [190] J. Zhang, X. Luo, M. Small, Detecting chaos in pseudoperiodic time series without embedding, *Phys. Rev. E* 73 (1) (2006) 016216. doi:10.1103/PhysRevE.73.016216.
- [191] J. Theiler, P. E. Rapp, Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram, *Electroencephalogr. Clin. Neurophysiol.* 98 (3) (1996) 213–222.
- [192] M. Small, D. Yu, R. G. Harrison, Surrogate Test for Pseudoperiodic Time Series Data, *Phys. Rev. Lett.* 87 (18) (2001) 188101. doi:10.1103/PhysRevLett.87.188101.
- [193] F. Emmert-Streib, Parametric construction of episode networks from pseudoperiodic time series based on mutual information, *PLoS One* 6 (12). doi:10.1371/journal.pone.0027733.

- [194] C. Men, J. Wang, Y.-M. Qin, B. Deng, X.-L. Wei, Characterizing electrical signals evoked by acupuncture through complex network mapping: A new perspective on acupuncture, *Comput. Methods Programs Biomed.* 104 (3) (2011) 498–504. doi:10.1016/j.cmpb.2011.08.006.
- [195] M. U. Kobayashi, Y. Saiki, Network analysis of chaotic systems through unstable periodic orbits, *Chaos* 27 (8) (2017) 081103. doi:10.1063/1.4995043.
- [196] P. Cvitanović, Invariant Measurement of Strange Sets in Terms of Cycles, *Phys. Rev. Lett.* 61 (24) (1988) 2729–2732. doi:10.1103/PhysRevLett.61.2729.
- [197] C. Grebogi, E. Ott, J. A. Yorke, Unstable periodic orbits and the dimensions of multifractal chaotic attractors, *Phys. Rev. A* 37 (5) (1988) 1711–1724. doi:10.1103/PhysRevA.37.1711.
- [198] M. Thiel, M. C. Romano, J. Kurths, Spurious structures in recurrence plots induced by embedding, *Nonlinear Dyn.* 44 (1-4) (2006) 299–305. doi:10.1007/s11071-006-2010-9.
- [199] F.-Z. Hou, J. Wang, X.-C. Wu, F.-R. Yan, A dynamic marker of very short-term heartbeat under pathological states via network analysis, *EPL (Europhysics Lett.)* 107 (5) (2014) 58001. doi:10.1209/0295-5075/107/58001.
- [200] S. Hempel, A. Koseska, J. Kurths, Z. Nikoloski, Inner Composition Alignment for Inferring Directed Networks from Short Time Series, *Phys. Rev. Lett.* 107 (5) (2011) 054101. doi:10.1103/PhysRevLett.107.054101.
- [201] T. Lozano-Pérez, M. A. Wesley, An algorithm for planning collision-free paths among polyhedral obstacles, *Communications of the ACM* 22 (10) (1979) 560–570. doi:10.1145/359156.359164.
- [202] G. Nagy, Terrain visibility, *Computers & Graphics* 18 (6) (1994) 763 – 773. doi:[https://doi.org/10.1016/0097-8493\(94\)90002-7](https://doi.org/10.1016/0097-8493(94)90002-7).
- [203] L. De Floriani, P. Marzano, E. Puppo, Line-of-sight communication on terrain models, *Int. J. Geogr. Inf. Syst.* 8 (4) (1994) 329–342. doi:10.1080/02693799408902004.
- [204] A. Turner, M. Doxa, D. O’Sullivan, A. Penn, From Isovists to Visibility Graphs: A Methodology for the Analysis of Architectural Space, *Environ. Plan. B Plan. Des.* 28 (1) (2001) 103–121. doi:10.1068/b2684.
- [205] L. Lacasa, R. Toral, Description of stochastic and chaotic series using visibility graphs, *Phys. Rev. E* 82 (3) (2010) 036120. doi:10.1103/PhysRevE.82.036120.
- [206] W.-J. Xie, W.-X. Zhou, Horizontal visibility graphs transformed from fractional Brownian motions: Topological properties versus the Hurst index, *Physica A* 390 (20) (2011) 3592–3601. arXiv:1012.3850, doi:10.1016/j.physa.2011.04.020.
- [207] G. Gutin, T. Mansour, S. Severini, A characterization of horizontal visibility graphs and combinatorics on words, *Physica A* 390 (12) (2011) 2421–2428. arXiv:1010.1850, doi:10.1016/j.physa.2011.02.031.
- [208] A. Nuñez, L. Lacasa, J. Patricio, B. Luque, Visibility Algorithms: A Short Review, in: *New Front. Graph Theory*, InTech, 2012, Ch. 6, pp. 119–152. doi:10.5772/34810.
- [209] B. Luque, F. J. Ballesteros, A. Robledo, L. Lacasa, Entropy and Renormalization in Chaotic Visibility Graphs, in: *Math. Found. Appl. Graph Entropy*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2016, pp. 1–39. doi:10.1002/9783527693245.ch1.
- [210] Y. Zou, R. V. Donner, N. Marwan, M. Small, J. Kurths, Long-term changes in the northsouth asymmetry of solar activity: a nonlinear dynamics characterization using visibility graphs, *Nonlinear Process. Geophys.* 21 (6) (2014) 1113–1126. doi:10.5194/npg-21-1113-2014.
- [211] B. Luque, M. Cordero-Gracia, M. Gómez, A. Robledo, Quasiperiodic graphs at the onset of chaos, *Phys. Rev. E* 88 (6) (2013) 062918. doi:10.1103/PhysRevE.88.062918.
- [212] B. Luque, F. J. Ballesteros, A. M. Núñez, A. Robledo, Quasiperiodic graphs: Structural design, scaling and entropic properties, *J. Nonlinear Sci.* 23 (2) (2013) 335–342. arXiv:1203.3717, doi:10.1007/s00332-012-9153-2.
- [213] S. Ahadpour, Y. Sadra, Randomness criteria in binary visibility graph and complex network perspective, *Inf. Sci. (Ny)* 197 (2012) 161–176. arXiv:1004.2189, doi:10.1016/j.ins.2012.02.022.
- [214] I. Bezsdudnov, A. Snarskii, From the time series to the complex networks: The parametric natural visibility graph, *Physica A* 414 (2014) 53–60. doi:10.1016/j.physa.2014.07.002.
- [215] A. Snarskii, I. Bezsdudnov, Critical phenomena in the dynamical visibility graph, ArXiv e-prints arXiv:1302.3265.
- [216] Z.-K. Gao, Q. Cai, Y.-X. Yang, W.-D. Dang, S.-S. Zhang, Multiscale limited penetrable horizontal visibility graph for analyzing nonlinear time series, *Sci. Rep.* 6 (1) (2016) 35622. doi:10.1038/srep35622.
- [217] X. Pei, J. Wang, B. Deng, X. Wei, H. Yu, WLPVG approach to the analysis of EEG-based functional brain network under manual acupuncture, *Cogn. Neurodyn.* 8 (5) (2014) 417–428. doi:10.1007/s11571-014-9297-x.
- [218] X. Li, M. Sun, C. Gao, D. Han, M. Wang, The parametric modified limited penetrable visibility graph for constructing complex networks from time series, *Physica A* 492 (2018) 1097–1106. doi:10.1016/j.physa.2017.11.040.
- [219] Q. Xiao, X. Pan, X.-L. Li, M. Stephen, H.-J. Yang, Y. Jiang, J.-Y. Wang, Q.-J. Zhang, Rowcolumn visibility graph approach to two-dimensional landscapes, *Chinese Phys. B* 23 (7) (2014) 078904. doi:10.1088/1674-1056/23/7/078904.
- [220] L. Lacasa, J. Iacovacci, Visibility graphs of random scalar fields and spatial data, *Phys. Rev. E* 96 (1) (2017) 012318. arXiv:1702.07813, doi:10.1103/PhysRevE.96.012318.
- [221] H. Wang, H. Mo, R. Sadiq, Y. Hu, Y. Deng, Ordered visibility graph weighted averaging aggregation operator: A methodology based on network analysis, *Comput. Ind. Eng.* 88 (2015) 181–190. doi:10.1016/j.cie.2015.06.021.
- [222] M. Stephen, C. Gu, H. Yang, Visibility Graph Based Time Series Analysis, *PLoS One* 10 (11) (2015) e0143015. doi:10.1371/journal.pone.0143015.
- [223] S. Mutua, C. Gu, H. Yang, Visibility graphlet approach to chaotic time series, *Chaos* 26 (5) (2016) 053107. doi:10.1063/1.4951681.
- [224] D. Tsiotas, A. Charakopoulos, Visibility in the topology of complex networks, *Physica A* 505 (2018) 280–292. doi:10.1016/j.physa.2018.03.055.
- [225] M. Wang, A. L. M. Vilela, R. Du, L. Zhao, G. Dong, L. Tian, H. E. Stanley, Exact results of the limited pene-

- trable horizontal visibility graph associated to random time series and its application, *Sci. Rep.* 8 (1) (2018) 5130. doi:10.1038/s41598-018-23388-1.
- [226] X. Lan, H. Mo, S. Chen, Q. Liu, Y. Deng, Fast transformation from time series to visibility graphs, *Chaos* 25 (8) (2015) 083105. doi:10.1063/1.4927835.
- [227] J. O. Pierini, M. Lovallo, L. Telesca, Visibility graph analysis of wind speed records measured in central Argentina, *Physica A* 391 (20) (2012) 5041–5048. doi:10.1016/j.physa.2012.05.049.
- [228] M. Ahmadlou, H. Adeli, A. Adeli, Improved visibility graph fractality with application for the diagnosis of Autism Spectrum Disorder, *Physica A* 391 (20) (2012) 4720–4726. doi:10.1016/j.physa.2012.04.025.
- [229] Z. Czechowski, M. Lovallo, L. Telesca, Multifractal analysis of visibility graph-based Ito-related connectivity time series, *Chaos* 26 (2) (2016) 023118. doi:10.1063/1.4942582.
- [230] L. Lacasa, Q. Mary, On the degree distribution of horizontal visibility graphs associated with Markov processes and dynamical systems: Diagrammatic and variational approaches, *Nonlinearity* 27 (9) (2014) 2063–2093. arXiv:1402.5368, doi:10.1088/0951-7715/27/9/2063.
- [231] Z. G. Yu, V. Anh, R. Eastes, Multifractal analysis of solar flare indices and their horizontal visibility graphs, *Nonlinear Process. Geophys.* 19 (6) (2012) 657–665. doi:10.5194/npg-19-657-2012.
- [232] A. Braga, L. Alves, L. Costa, A. Ribeiro, M. de Jesus, A. Tateishi, H. Ribeiro, Characterization of river flow fluctuations via horizontal visibility graphs, *Physica A* 444 (2016) 1003–1011. arXiv:1510.07009, doi:10.1016/j.physa.2015.10.102.
- [233] Y. Zou, M. Small, Z. Liu, J. Kurths, Complex network approach to characterize the statistical features of the sunspot series, *New J. Phys.* 16 (1) (2014) 013051. doi:10.1088/1367-2630/16/1/013051.
- [234] B. A. Gonçalves, L. Carpi, O. A. Rosso, M. G. Ravetti, Time series characterization via horizontal visibility graph and Information Theory, *Physica A* 464 (2016) 93–102. doi:10.1016/j.physa.2016.07.063.
- [235] S. Chen, Y. Hu, S. Mahadevan, Y. Deng, A visibility graph averaging aggregation operator, *Physica A* 403 (2014) 1–12. arXiv:1311.4166, doi:10.1016/j.physa.2014.02.015.
- [236] W. Jiang, B. Wei, J. Zhan, C. Xie, D. Zhou, A visibility graph power averaging aggregation operator: A methodology based on network analysis, *Comput. Ind. Eng.* 101 (2016) 260–268. doi:10.1016/j.cie.2016.09.009.
- [237] M. G. Ravetti, L. C. Carpi, B. A. Gonçalves, A. C. Frery, O. A. Rosso, Distinguishing Noise from Chaos: Objective versus Subjective Criteria Using Horizontal Visibility Graph, *PLoS One* 9 (9) (2014) e108004. arXiv:1401.2139, doi:10.1371/journal.pone.0108004.
- [238] R. Zhang, Y. Zou, J. Zhou, Z.-K. Gao, S.-G. Guan, Visibility graph analysis for re-sampled time series from auto-regressive stochastic processes, *Commun. Nonlinear Sci. Numer. Simul.* 42 (2017) 396–403. doi:10.1016/j.cnsns.2016.04.031.
- [239] P. Manshour, Complex network approach to fractional time series, *Chaos* 25 (10) (2015) 103105. arXiv:1512.08205, doi:10.1063/1.4930839.
- [240] Z.-G. Shao, Network analysis of human heartbeat dynamics, *Appl. Phys. Lett.* 96 (7) (2010) 073703. doi:10.1063/1.3308505.
- [241] C. Song, S. Havlin, H. A. Makse, Origins of fractality in the growth of complex networks, *Nat. Phys.* 2 (4) (2006) 275–281. doi:10.1038/nphys266.
- [242] X. Li, Z. Dong, Detection and prediction of the onset of human ventricular fibrillation: An approach based on complex network theory, *Phys. Rev. E* 84 (6) (2011) 062901. doi:10.1103/PhysRevE.84.062901.
- [243] J. Iacovacci, L. Lacasa, Sequential visibility-graph motifs, *Phys. Rev. E* 93 (4) (2016) 042309. arXiv:1512.00297, doi:10.1103/PhysRevE.93.042309.
- [244] R. J. Telford, E. Heegaard, H. J. B. Birks, All age-depth models are wrong: But how badly?, *Quat. Sci. Rev.* 23 (1-2) (2004) 1–5. doi:10.1016/j.quascirev.2003.11.003.
- [245] L. Telesca, M. Lovallo, Analysis of seismic sequences by using the method of visibility graph, *EPL (Europhysics Lett.)* 97 (5) (2012) 50002. doi:10.1209/0295-5075/97/50002.
- [246] J. Theiler, Spurious dimension from correlation algorithms applied to limited time-series data, *Phys. Rev. A* 34 (1986) 2427–2432. doi:10.1103/PhysRevA.34.2427.
- [247] M. Ahmadlou, H. Adeli, Visibility graph similarity: A new measure of generalized synchronization in coupled dynamic systems, *Phys. D* 241 (4) (2012) 326–332. doi:10.1016/j.physd.2011.09.008.
- [248] A. Mitra, B. Pyne, Measuring Synchronization for coupled systems using Visibility Graph Similarity, *Int. J. Recent Technol. Eng.* 1 (3) (2012) 163–168.
- [249] S. Mehraban, a. H. Shirazi, M. Zamani, G. R. Jafari, Coupling between time series: A network view, *EPL (Europhysics Lett.)* 103 (5) (2013) 50011. doi:10.1209/0295-5075/103/50011.
- [250] G. Weiss, Time-reversibility of linear stochastic processes, *Journal of Applied Probability* 12 (4) (1975) 831?836. doi:10.2307/3212735.
- [251] A. J. Lawrance, Directionality and Reversibility in Time Series, *Int. Stat. Rev. / Rev. Int. Stat.* 59 (1) (1991) 67. doi:10.2307/1403575.
- [252] C. Diks, J. C. van Houwelingen, F. Takens, J. DeGoede, Reversibility as a criterion for discriminating time series, *Phys. Lett. A* 201 (2-3) (1995) 221–228. doi:10.1016/0375-9601(95)00239-Y.
- [253] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J. Doyne Farmer, Testing for nonlinearity in time series: the method of surrogate data, *Phys. D* 58 (1-4) (1992) 77–94. arXiv:/doi.org/10.1016/0167-2789(92)90102-S, doi:10.1016/0167-2789(92)90102-S.
- [254] H. Voss, J. Kurths, Test for nonlinear dynamical behavior in symbol sequences, *Phys. Rev. E* 58 (1) (1998) 1155–1158. doi:10.1103/PhysRevE.58.1155.
- [255] C. S. Daw, C. E. A. Finney, M. B. Kennel, Symbolic approach for measuring temporal irreversibility, *Phys. Rev. E* 62 (2) (2000) 1912–1921. doi:10.1103/PhysRevE.62.1912.

- [256] M. B. Kennel, Testing time symmetry in time series using data compression dictionaries, *Phys. Rev. E* 69 (5) (2004) 056208. doi:[10.1103/PhysRevE.69.056208](https://doi.org/10.1103/PhysRevE.69.056208).
- [257] C. Cammarota, E. Rogora, Time reversal, symbolic series and irreversibility of human heartbeat, *Chaos, Solitons and Fractals* 32 (5) (2007) 1649–1654. doi:[10.1016/j.chaos.2006.03.126](https://doi.org/10.1016/j.chaos.2006.03.126).
- [258] M. Costa, A. L. Goldberger, C.-K. Peng, Broken Asymmetry of the Human Heartbeat: Loss of Time Irreversibility in Aging and Disease, *Phys. Rev. Lett.* 95 (19) (2005) 198102. doi:[10.1103/PhysRevLett.95.198102](https://doi.org/10.1103/PhysRevLett.95.198102).
- [259] A. Porporato, J. R. Rigby, E. Daly, Irreversibility and Fluctuation Theorem in Stationary Time Series, *Phys. Rev. Lett.* 98 (9) (2007) 094101. arXiv:0703213, doi:[10.1103/PhysRevLett.98.094101](https://doi.org/10.1103/PhysRevLett.98.094101).
- [260] É. Roldán, J. M. R. Parrondo, Estimating Dissipation from Single Stationary Trajectories, *Phys. Rev. Lett.* 105 (15) (2010) 150607. arXiv:1004.2831, doi:[10.1103/PhysRevLett.105.150607](https://doi.org/10.1103/PhysRevLett.105.150607).
- [261] L. Lacasa, R. Flanagan, Time reversibility from visibility graphs of nonstationary processes, *Phys. Rev. E* 92 (2) (2015) 022817. doi:[10.1103/PhysRevE.92.022817](https://doi.org/10.1103/PhysRevE.92.022817).
- [262] H. Xiong, P. Shang, J. Xia, J. Wang, Time irreversibility and intrinsics revealing of series with complex network approach, *Physica A* 499 (2018) 241–249. doi:[10.1016/j.physa.2018.02.041](https://doi.org/10.1016/j.physa.2018.02.041).
- [263] H. Tong, *Non-Linear Time Series: A Dynamical System Approach*, Oxford University Press, Oxford, 1990.
- [264] P. Yang, P. Shang, Relative asynchronous index: a new measure for time series irreversibility, *Nonlinear Dyn.* doi:[10.1007/s11071-018-4275-1](https://doi.org/10.1007/s11071-018-4275-1).
- [265] F. Xie, Z. Fu, L. Piao, J. Mao, Time irreversibility of mean temperature anomaly variations over China, *Theor. Appl. Climatol.* 123 (1-2) (2016) 161–170. doi:[10.1007/s00704-014-1347-0](https://doi.org/10.1007/s00704-014-1347-0).
- [266] R. Flanagan, L. Lacasa, Irreversibility of financial time series: A graph-theoretical approach, *Phys. Lett. A* 380 (20) (2016) 1689–1697. arXiv:1601.01980, doi:[10.1016/j.physleta.2016.03.011](https://doi.org/10.1016/j.physleta.2016.03.011).
- [267] D. Meng, C. Xiao-yan, L. Hong-ying, H. Qing, B. Rui-xiang, L. Weixin, L. Zewei, Time Irreversibility from Time Series for Analyzing Oil-in-Water Flow Transition, *Math. Probl. Eng.* 2016 (2016) 1–10. doi:[10.1155/2016/2879524](https://doi.org/10.1155/2016/2879524).
- [268] F. Serinaldi, C. G. Kilsby, Irreversibility and complex network behavior of stream flow fluctuations, *Physica A* 450 (2016) 585–600. doi:[10.1016/j.physa.2016.01.043](https://doi.org/10.1016/j.physa.2016.01.043).
- [269] L. Telesca, E. L. Flores-Márquez, A. Ramírez-Rojas, Time-reversibility in seismic sequences: Application to the seismicity of Mexican subduction zone, *Physica A* 492 (2018) 1373–1381. doi:[10.1016/j.physa.2017.11.064](https://doi.org/10.1016/j.physa.2017.11.064).
- [270] J. P. Crutchfield, K. Young, Inferring statistical complexity, *Phys. Rev. Lett.* 63 (1989) 105–108. doi:[10.1103/PhysRevLett.63.105](https://doi.org/10.1103/PhysRevLett.63.105).
- [271] C.-B. Li, H. Yang, T. Komatsuzaki, Multiscale complex network of protein conformational fluctuations in single-molecule time series, *Proc. Natl. Acad. Sci.* 105 (2) (2008) 536–541. doi:[10.1073/pnas.0707378105](https://doi.org/10.1073/pnas.0707378105).
- [272] L. Liu, Z. Wang, Encoding Temporal Markov Dynamics in Graph for Time Series Visualization, ArXiv arXiv:1610.07273.
- [273] T. Weng, J. Zhang, M. Small, R. Zheng, P. Hui, Memory and betweenness preference in temporal networks induced from time series, *Sci. Rep.* 7 (February) (2017) 41951. doi:[10.1038/srep41951](https://doi.org/10.1038/srep41951).
- [274] M. Small, Complex Networks from Time Series: Capturing Dynamics, *IEEE Int. Symp. Circuits Syst. Proc.* (2013) 2509–2512.
- [275] M. McCullough, M. Small, H. H. C. Iu, T. Stemler, Multiscale ordinal network analysis of human cardiac dynamics, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 375 (2017) 20160292. doi:[10.1098/rsta.2016.0292](https://doi.org/10.1098/rsta.2016.0292).
- [276] M. Small, M. McCullough, K. Sakellariou, Ordinal Network Measures Quantifying Determinism in Data, in: 2018 IEEE Int. Symp. Circuits Syst., IEEE, 2018, pp. 1–5. doi:[10.1109/ISCAS.2018.8351743](https://doi.org/10.1109/ISCAS.2018.8351743).
- [277] X. Sun, M. Small, Y. Zhao, X. Xue, Characterizing system dynamics with a weighted and directed network constructed from time series data, *Chaos* 24 (2) (2014) 024402. doi:[10.1063/1.4868261](https://doi.org/10.1063/1.4868261).
- [278] H. Guo, J. Zhang, Y. Zou, S. Guan, Cross and joint ordinal partition transition networks for multivariate time series analysis, *Frontiers of Physics* 13 (5) (2018) 130508. doi:[10.1080/02693799408902004](https://doi.org/10.1080/02693799408902004).
- [279] F. Emmert-Streib, Universal construction mechanism for networks from one-dimensional symbol sequences, *Appl. Math. Comput.* 219 (3) (2012) 1020–1030. doi:[10.1016/j.amc.2012.07.006](https://doi.org/10.1016/j.amc.2012.07.006).
- [280] D. Makowiec, Z. Struzik, B. Graff, J. Wdowczyk-Szulc, M. Żarczyńska-Buchowiecka, S. Kryszewski, Community Structure in Network Representation of Increments in Beat-to-beat Time Intervals of the Heart in Patients After Heart Transplantation, *Acta Phys. Pol. B* 44 (5) (2013) 1219. doi:[10.5506/APhysPo1B.44.1219](https://doi.org/10.5506/APhysPo1B.44.1219).
- [281] D. Makowiec, Z. Struzik, B. Graff, J. Wdowczyk-Szulc, M. Zarczynska- Buchowiecka, M. Gruchala, A. Rynkiewicz, Complexity of the heart rhythm after heart transplantation by entropy of transition network for RR-increments of RR time intervals between heartbeats, *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* (2013) 6127–6130 doi:[10.1109/EMBC.2013.6610951](https://doi.org/10.1109/EMBC.2013.6610951).
- [282] D. Makowiec, Z. Struzik, B. Graff, M. Żarczyńska-Buchowiecka, J. Wdowczyk, Transition Network Entropy in Characterization of Complexity of Heart Rhythm After Heart Transplantation, *Acta Phys. Pol. B* 45 (8) (2014) 1771. doi:[10.5506/APhysPo1B.45.1771](https://doi.org/10.5506/APhysPo1B.45.1771).
- [283] D. Makowiec, A. Kaczkowska, D. Wejer, M. Zarczyńska-Buchowiecka, Z. R. Struzik, Entropic measures of complexity of short-term dynamics of nocturnal heartbeats in an aging population, *Entropy* 17 (3) (2015) 1253–1272. doi:[10.3390/e17031253](https://doi.org/10.3390/e17031253).
- [284] D. Makowiec, D. Wejer, A. Kaczkowska, M. Żarczyńska-Buchowiecka, Z. R. Struzik, Chronographic Imprint of Age-Induced Alterations in Heart Rate Dynamical Organization, *Front. Physiol.* 6 (JUL) (2015) 1–10. doi:[10.3389/fphys.2015.00201](https://doi.org/10.3389/fphys.2015.00201).
- [285] D. Makowiec, J. Wdowczyk, M. Gruchala, Z. R. Struzik, Network tools for tracing the dynamics of heart rate after cardiac transplantation, *Chaos, Solitons and Fractals* 90 (2016) 101–110. doi:[10.1016/j.chaos.2016.03.024](https://doi.org/10.1016/j.chaos.2016.03.024).
- [286] I. Laut, C. Räth, Surrogate-assisted network analysis of nonlinear time series, *Chaos* 26 (10) (2016) 103108.

- doi:10.1063/1.4964646.
- [287] P. Li, B. Wang, An approach to Hang Seng Index in Hong Kong stock market based on network topological statistics, *Chinese Sci. Bull.* 51 (5) (2006) 624–629. doi:10.1007/s11434-006-0624-4.
- [288] P. Li, B.-H. Wang, Extracting hidden fluctuation patterns of Hang Seng stock index from network topologies, *Physica A* 378 (2) (2007) 519–526. doi:10.1016/j.physa.2006.10.089.
- [289] X. Gao, H. An, W. Fang, X. Huang, H. Li, W. Zhong, Y. Ding, Transmission of linear regression patterns between time series: From relationship in time series to complex networks, *Phys. Rev. E* 90 (1) (2014) 012818. doi:10.1103/PhysRevE.90.012818.
- [290] X. Gao, H. An, W. Fang, H. Li, X. Sun, The transmission of fluctuant patterns of the forex burden based on international crude oil prices, *Energy* 73 (2014) 380–386. doi:10.1016/j.energy.2014.06.028.
- [291] X. Gao, H. An, W. Fang, X. Huang, H. Li, W. Zhong, Characteristics of the transmission of autoregressive sub-patterns in financial time series, *Sci. Rep.* 4 (1) (2015) 6290. doi:10.1038/srep06290.
- [292] X. Huang, H. An, X. Gao, X. Hao, P. Liu, Multiresolution transmission of the correlation modes between bivariate time series based on complex network theory, *Physica A* 428 (2015) 493–506. doi:10.1016/j.physa.2015.02.028.
- [293] T. Nakamura, T. Tanizawa, Networks with time structure from time series, *Physica A* 391 (20) (2012) 4704–4710. doi:10.1016/j.physa.2012.05.039.
- [294] T. Nakamura, T. Tanizawa, M. Small, Constructing networks from a dynamical system perspective for multivariate nonlinear time series, *Phys. Rev. E* 93 (3) (2016) 032323. doi:10.1103/PhysRevE.93.032323.
- [295] A. Lichtenberg, M. Lieberman, *Regular and Chaotic Dynamics*, 2nd Edition, Springer, Berlin, 1992.
- [296] J. D. Meiss, Symplectic maps, variational principles, and transport, *Rev. Mod. Phys.* 64 (3) (1992) 795–848. doi:10.1103/RevModPhys.64.795.
- [297] C. F. Karney, Long-time correlations in the stochastic regime, *Phys. D* 8 (3) (1983) 360–380. doi:10.1016/0167-2789(83)90232-4.
- [298] M. H. Olyaei, A. Yaghoubi, M. Yaghoubi, Predicting protein structural classes based on complex networks and recurrence analysis, *J. Theor. Biol.* 404 (2016) 375–382. doi:10.1016/j.jtbi.2016.06.018.
- [299] K. Rehfeld, N. Marwan, S. F. M. Breitenbach, J. Kurths, Late Holocene Asian summer monsoon dynamics from small but complex networks of paleoclimate data, *Clim. Dyn.* doi:10.1007/s00382-012-1448-3.
- [300] S. Böse, Recurrence network analysis of remote sensing data, Master's thesis, University of Bayreuth (2012).
- [301] H. Lange, S. Boese, Recurrence quantification and recurrence network analysis of global photosynthetic activity, in: C. L. Webber, Jr., N. Marwan (Eds.), *Recurrence Quantification Analysis: Theory and Best Practices*, Springer International Publishing, Cham, 2015, pp. 349–374. doi:10.1007/978-3-319-07155-8\_12.
- [302] G. Górski, G. Litak, R. Mosdorff, A. Rysak, Two phase flow bifurcation due to turbulence: transition from slugs to bubbles, *Eur. Phys. J. B* 88 (9) (2015) 239. doi:10.1140/epjb/e2015-60245-8.
- [303] Z.-K. Gao, N.-D. Jin, Complex network from time series based on phase space reconstruction, *Chaos* 19 (3) (2009) 033137. doi:10.1063/1.3227736.
- [304] Z.-K. Gao, N.-D. Jin, Erratum: Complex network from time series based on phase space reconstruction [Chaos 19, 033137 (2009)], *Chaos* 20 (1) (2010) 019902. doi:10.1063/1.3332246.
- [305] Z.-K. Gao, N.-D. Jin, Characterization of chaotic dynamic behavior in the gasliquid slug flow using directed weighted complex network analysis, *Physica A* 391 (10) (2012) 3005–3016. doi:10.1016/j.physa.2012.01.025.
- [306] Z.-K. Gao, N.-D. Jin, A directed weighted complex network for characterizing chaotic dynamics from time series, *Nonlinear Anal. Real World Appl.* 13 (2) (2012) 947–952. doi:10.1016/j.nonrwa.2011.08.029.
- [307] Z.-K. Gao, L.-D. Hu, N.-D. Jin, Markov transition probability-based network from time series for characterizing experimental two-phase flow, *Chinese Phys. B* 22 (5) (2013) 050507. doi:10.1088/1674-1056/22/5/050507.
- [308] Z.-K. Gao, X.-W. Zhang, N.-D. Jin, R. V. Donner, N. Marwan, J. Kurths, Recurrence networks from multivariate signals for uncovering dynamic transitions of horizontal oil-water stratified flows, *EPL (Europhysics Lett.)* 103 (5) (2013) 50004. doi:10.1209/0295-5075/103/50004.
- [309] X.-W. Zhang, N.-D. Jin, Z.-K. Gao, L.-S. Zhai, Local Property of Recurrence Network for Investigating Gas-Liquid Two-Phase Flow Characteristics, *Chinese Phys. Lett.* 30 (5) (2013) 050501. doi:10.1088/0256-307X/30/5/050501.
- [310] Z.-K. Gao, X.-W. Zhang, M. Du, N.-D. Jin, Recurrence network analysis of experimental signals from bubbly oil-in-water flows, *Phys. Lett. A* 377 (6) (2013) 457–462. doi:10.1016/j.physleta.2012.12.017.
- [311] Z.-K. Gao, N.-D. Jin, W.-X. Wang, Y.-C. Lai, Motif distributions in phase-space networks for characterizing experimental two-phase flow patterns with chaotic features, *Phys. Rev. E* 82 (1) (2010) 016210. doi:10.1103/PhysRevE.82.016210.
- [312] Z.-K. Gao, Y.-X. Yang, P.-C. Fang, Y. Zou, C.-Y. Xia, M. Du, Multiscale complex network for analyzing experimental multivariate time series, *EPL (Europhysics Lett.)* 109 (3) (2015) 30005. doi:10.1209/0295-5075/109/30005.
- [313] M. Wickramasinghe, I. Z. Kiss, Effect of temperature on precision of chaotic oscillations in nickel electrodissolution, *Chaos* 20 (2) (2010) 023125. doi:10.1063/1.3439209.
- [314] N. Marwan, N. Wessel, H. Stepan, J. Kurths, Recurrence Based Complex Network Analysis of Cardiovascular Variability Data to Predict Pre-Eclampsia, in: Proc. Int. Symp. Nonlinear Theory its Appl. (NOLTA2010), Krakow, 2010, pp. 585–588.
- [315] G. Ramirez Avila, A. Gapelyuk, N. Marwan, H. Stepan, J. Kurths, T. Walther, N. Wessel, Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, in: Proc. 7th ESGCO, Vol. 178, 2012, pp. 103–110.
- [316] G. M. Ramírez Ávila, A. Gapelyuk, N. Marwan, H. Stepan, J. Kurths, T. Walther, N. Wessel, Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, *Auton. Neurosci. Basic Clin.* 178 (1-2) (2013) 103–110. doi:10.1016/j.autneu.2013.05.003.

- [317] N. P. Subramaniyam, J. Hyttinen, Analysis of nonlinear dynamics of healthy and epileptic EEG signals using recurrence based complex network approach, in: 2013 6th Int. IEEE/EMBS Conf. Neural Eng., IEEE, 2013, pp. 605–608. doi:10.1109/NER.2013.6696007.
- [318] C. Liu, W.-X. Zhou, W.-K. Yuan, Statistical properties of visibility graph of energy dissipation rates in three-dimensional fully developed turbulence, *Physica A* 389 (13) (2010) 2675–2681. doi:10.1016/j.physa.2010.02.043.
- [319] P. Manshour, M. R. Rahimi Tabar, J. Peinke, Fully developed turbulence in the view of horizontal visibility graphs, *J. Stat. Mech. Theory Exp.* 2015 (8) (2015) P08031. doi:10.1088/1742-5468/2015/08/P08031.
- [320] Y. Yang, J. Wang, H. Yang, J. Mang, Visibility graph approach to exchange rate series, *Physica A* 388 (20) (2009) 4431–4437. doi:10.1016/j.physa.2009.07.016.
- [321] M.-C. Qian, Z.-Q. Jiang, W.-X. Zhou, Universal and nonuniversal allometric scaling behaviors in the visibility graphs of world stock market indices, *J. Phys. A Math. Theor.* 43 (33) (2010) 335002. arXiv:0910.2524, doi:10.1088/1751-8113/43/33/335002.
- [322] Z. Dong, X. Li, Comment on Network analysis of human heartbeat dynamics [Appl. Phys. Lett. 96, 073703 (2010)], *Appl. Phys. Lett.* 96 (26) (2010) 266101. doi:10.1063/1.3458811.
- [323] M. Ahmadlou, H. Adeli, A. Adeli, New diagnostic EEG markers of the Alzheimer’s disease using visibility graph, *J. Neural Transm.* 117 (9) (2010) 1099–1109. doi:10.1007/s00702-010-0450-3.
- [324] S. Jiang, C. Bian, X. Ning, Q. D. Y. Ma, Visibility graph analysis on heartbeat dynamics of meditation training, *Appl. Phys. Lett.* 102 (25) (2013) 253702. doi:10.1063/1.4812645.
- [325] S. Bhaduri, D. Ghosh, Electroencephalographic Data Analysis With Visibility Graph Technique for Quantitative Assessment of Brain Dysfunction, *Clin. EEG Neurosci.* 46 (3) (2015) 218–223. doi:10.1177/1550059414526186.
- [326] H. Liu, Q. Meng, Q. Zhang, Z. Zhang, D. Wang, A Novel Feature Extraction Method for Epileptic Seizure Detection Based on the Degree Centrality of Complex Network and SVM, in: D.-S. Huang, K.-H. Jo (Eds.), *Intell. Comput. Theor. Appl.*, Springer International Publishing, Cham, 2016, pp. 170–180. doi:10.1007/978-3-319-42294-7\_14.
- [327] H. Zhang, Q. Meng, M. Liu, Y. Li, A New Epileptic Seizure Detection Method Based on Fusion Feature of Weighted Complex Network, in: T. Huang, J. Lv, C. Sun, A. V. Tuzikov (Eds.), *Adv. Neural Networks – ISNN 2018*, Springer International Publishing, Cham, 2018, pp. 834–841.
- [328] X. Long, P. Fonseca, R. M. Aarts, R. Haakma, J. Foussier, Modeling cardiorespiratory interaction during human sleep with complex networks, *Appl. Phys. Lett.* 105 (20) (2014) 203701. doi:10.1063/1.4902026.
- [329] G. Zhu, Y. Li, P. Wen, S. Wang, Analysis of alcoholic EEG signals based on horizontal visibility graph entropy, *Brain Informatics* 1 (1-4) (2014) 19–25. doi:10.1007/s40708-014-0003-x.
- [330] J. B. Elsner, T. H. Jagger, E. a. Fogarty, Visibility network of United States hurricanes, *Geophys. Res. Lett.* 36 (16) (2009) L16702. doi:10.1029/2009GL039129.
- [331] Q. Tang, J. Liu, H. Liu, Comparison of different daily streamflow series in US and China, under a viewpoint of complex networks, *Mod. Phys. Lett. B* 24 (14) (2010) 1541–1547. doi:10.1142/S0217984910023335.
- [332] J. Wang, H. Yang, Complex network-based analysis of air temperature data in china, *Mod. Phys. Lett. B* 23 (14) (2009) 1781–1789.
- [333] N. Wang, D. Li, Q. Wang, Visibility graph analysis on quarterly macroeconomic series of China based on complex network theory, *Physica A* 391 (24) (2012) 6543–6555. doi:10.1016/j.physa.2012.07.054.
- [334] G. Hloupis, Temporal pattern in Corinth rift seismicity revealed by visibility graph analysis, *Commun. Nonlinear Sci. Numer. Simul.* 51 (2017) 13–22. doi:10.1016/j.cnsns.2017.03.015.
- [335] L. Telesca, 4 - Fractal, Informational and Topological Methods for the Analysis of Discrete and Continuous Seismic Time Series: An Overview, in: T. Chelidze, F. Vallianatos, L. Telesca (Eds.), *Complex. Seism. Time Ser.*, Elsevier, 2018, pp. 95–139. doi:<https://doi.org/10.1016/B978-0-12-813138-1.00004-3>.
- [336] A. Charakopoulos, T. E. Karakasidis, P. N. Papanicolaou, A. Liakopoulos, The application of complex network time series analysis in turbulent heated jets, *Chaos* 24 (2) (2014) 024408. doi:10.1063/1.4875040.
- [337] X. Tang, L. Xia, Y. Liao, W. Liu, Y. Peng, T. Gao, Y. Zeng, New approach to epileptic diagnosis using visibility graph of high-frequency signal, *Clin. EEG Neurosci.* 44 (2) (2013) 150–156. doi:10.1177/1550059412464449.
- [338] H. Liu, X. Zhang, X. Zhang, Exploring dynamic evolution and fluctuation characteristics of air traffic flow volume time series: A single waypoint case, *Physica A* 503 (2018) 560–571. doi:10.1016/j.physa.2018.02.030.
- [339] A. Charakopoulos, G. Katsoulis, T. Karakasidis, Dynamics and causalities of atmospheric and oceanic data identified by complex networks and Granger causality analysis, *Physica A* 495 (2018) 436–453. doi:10.1016/j.physa.2017.12.027.
- [340] H. Zhang, D. Xu, Y. Wu, Predicting catastrophes of non-autonomous networks with visibility graphs and horizontal visibility, *Mech. Syst. Signal Process.* 104 (2018) 494–502. doi:10.1016/j.ymssp.2017.11.020.
- [341] P. Mali, S. Manna, A. Mukhopadhyay, P. Haldar, G. Singh, Multifractal analysis of multiparticle emission data in the framework of visibility graph and sandbox algorithm, *Physica A* 493 (2018) 253–266. doi:10.1016/j.physa.2017.10.015.
- [342] Á. M. Núñez, B. Luque, L. Lacasa, J. P. Gómez, A. Robledo, Horizontal visibility graphs generated by type-I intermittency, *Phys. Rev. E* 87 (5) (2013) 052801. arXiv:1310.0690, doi:10.1103/PhysRevE.87.052801.
- [343] Á. M. Núñez, L. Lacasa, J. P. Gómez, Horizontal Visibility graphs generated by type-II intermittency, *J. Phys. A Math. Theor.* 47 (3). arXiv:1310.0690, doi:10.1088/1751-8113/47/3/035102.
- [344] SIDC-team, The International Sunspot Number & Sunspot Area Data, Monthly Report on the International Sunspot Number, <http://www.sidc.be/sunspot-data/>, Royal Observatory Greenwich, <http://solarscience.msfc.nasa.gov/greenwch.shtml/>.
- [345] H. Newton, A. Milsom, Note on the observed differences in spottedness of the Sun’s northern and southern hemispheres, *Mon. Not. R. Astron. Soc.* 115 (1955) 398–404.
- [346] M. Carbonell, R. Oliver, J. L. Ballester, On the asymmetry of solar activity, *Astron. Astrophys.* 274 (1993) 497.

- [347] N. V. Zolotova, D. I. Ponyavin, Phase asynchrony of the north-south sunspot activity, *Astron. Astrophys.* 449 (1) (2006) L1–L4. doi:[10.1051/0004-6361:200600013](https://doi.org/10.1051/0004-6361:200600013).
- [348] R. V. Donner, M. Thiel, Scale-resolved phase coherence analysis of hemispheric sunspot activity: a new look at the north-south asymmetry, *Astron. Astrophys.* 475 (3) (2007) L33–L36. doi:[10.1051/0004-6361:20078672](https://doi.org/10.1051/0004-6361:20078672).
- [349] R. V. Donner, Phase Coherence Analysis of Decadal-Scale Sunspot Activity on Both Solar Hemispheres, in: *Non-linear Time Ser. Anal. Geosci.*, Vol. 112, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 355–385. doi:[10.1007/978-3-540-78938-3\\_16](https://doi.org/10.1007/978-3-540-78938-3_16).
- [350] K. J. Li, P. X. Gao, L. S. Zhan, X. J. Shi, W. W. Zhu, On the synchronization of hemispheric high-latitude solar activity, *Mon. Not. R. Astron. Soc.* 391 (1) (2008) L34–L38. doi:[10.1111/j.1745-3933.2008.00551.x](https://doi.org/10.1111/j.1745-3933.2008.00551.x).
- [351] Q. Li, Periodicity and Hemispheric Phase Relationship in High-Latitude Solar Activity, *Sol. Phys.* 249 (1) (2008) 135–145. doi:[10.1007/s11207-008-9147-2](https://doi.org/10.1007/s11207-008-9147-2).
- [352] N. V. Zolotova, D. I. Ponyavin, N. Marwan, J. Kurths, Long-term asymmetry in the wings of the butterfly diagram, *Astron. Astrophys.* 503 (1) (2009) 197–201. doi:[10.1051/0004-6361/200811430](https://doi.org/10.1051/0004-6361/200811430).
- [353] N. P. Subramanyam, J. F. Donges, J. Hyttinen, Signatures of chaotic and stochastic dynamics uncovered with  $\epsilon$ -recurrence networks, *Proc. R. Soc. A Math. Phys. Eng. Sci.* 471 (2183) (2015) 20150349. doi:[10.1098/rspa.2015.0349](https://doi.org/10.1098/rspa.2015.0349).
- [354] P. Holme, J. Saramäki, Temporal networks, *Phys. Rep.* 519 (3) (2012) 97–125. arXiv:[1108.1780](https://arxiv.org/abs/1108.1780), doi:[10.1016/j.physrep.2012.03.001](https://doi.org/10.1016/j.physrep.2012.03.001).