

## Related content

### Recurrence networks—a novel paradigm for nonlinear time series analysis

To cite this article: Reik V Donner *et al* 2010 *New J. Phys.* **12** 033025

- [Power-laws in recurrence networks from dynamical systems](#)  
Y. Zou, J. Heitzig, R. V. Donner *et al.*

- ['Clumpiness' mixing in complex networks](#)  
Ernesto Estrada, Naomichi Hatano and Amauri Gutierrez

- [The backbone of the climate network](#)  
J. F. Donges, Y. Zou, N. Marwan *et al.*

View the [article online](#) for updates and enhancements.

## Recent citations

- [Sensitivity and Nonlinearity of Thermoacoustic Oscillations](#)  
Matthew P. Juniper and R.I. Sujith

- [Tidal Analysis Using Time–Frequency Signal Processing and Information Clustering](#)  
Antonio M. Lopes and Jose Tenreiro Machado

- [Dynamics and causalities of atmospheric and oceanic data identified by complex networks and Granger causality analysis](#)  
A.K. Charakopoulos *et al.*

## Recurrence networks—a novel paradigm for nonlinear time series analysis

**Reik V Donner<sup>1,2,3,4,6</sup>, Yong Zou<sup>3</sup>, Jonathan F Donges<sup>3,5</sup>,  
Norbert Marwan<sup>3</sup> and Jürgen Kurths<sup>3,5</sup>**

<sup>1</sup> Max Planck Institute for Physics of Complex Systems, Dresden, Germany

<sup>2</sup> Institute for Transport and Economics, Dresden University of Technology,  
Germany

<sup>3</sup> Potsdam Institute for Climate Impact Research, Potsdam, Germany

<sup>4</sup> Graduate School of Science, Osaka Prefecture University, Sakai, Japan

<sup>5</sup> Institute for Physics, Humboldt University of Berlin, Berlin, Germany

E-mail: [donner@vwi.tu-dresden.de](mailto:donner@vwi.tu-dresden.de)

*New Journal of Physics* **12** (2010) 033025 (40pp)

Received 24 August 2009

Published 15 March 2010

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/12/3/033025

**Abstract.** This paper presents a new approach for analysing the structural properties of time series from complex systems. Starting from the concept of recurrences in phase space, the recurrence matrix of a time series is interpreted as the adjacency matrix of an associated complex network, which links different points in time if the considered states are closely neighboured in phase space. In comparison with similar network-based techniques the new approach has important conceptual advantages, and can be considered as a unifying framework for transforming time series into complex networks that also includes other existing methods as special cases. It has been demonstrated here that there are fundamental relationships between many topological properties of recurrence networks and different nontrivial statistical properties of the phase space density of the underlying dynamical system. Hence, this novel interpretation of the recurrence matrix yields new quantitative characteristics (such as average path length, clustering coefficient, or centrality measures of the recurrence network) related to the dynamical complexity of a time series, most of which are not yet provided by other existing methods of nonlinear time series analysis.

<sup>6</sup> Author to whom any correspondence should be addressed.

## Contents

<b>1. Introduction</b>	<b>2</b>
<b>2. Approaches for transforming time series into complex networks— a comparative review</b>	<b>5</b>
2.1. Partitioning of phase space: transition networks . . . . .	5
2.2. Cycle networks for pseudo-periodic time series . . . . .	6
2.3. Correlation networks of embedded state vectors . . . . .	7
2.4. Visibility graphs . . . . .	8
2.5. Complex networks based on neighbourhood relations in phase space . . . . .	8
2.6. Recurrence networks as a unifying framework for complex network-based time series analysis . . . . .	9
<b>3. Quantitative characteristics of recurrence networks</b>	<b>10</b>
3.1. Local network properties . . . . .	11
3.2. Intermediate scale network properties . . . . .	13
3.3. Global network properties . . . . .	15
<b>4. Examples</b>	<b>20</b>
4.1. Model systems . . . . .	20
4.2. $\epsilon$ -dependence of network properties . . . . .	20
4.3. Spatial distributions of vertex properties . . . . .	23
4.4. Spatial distributions of edge properties . . . . .	28
<b>5. Conclusions</b>	<b>29</b>
<b>Acknowledgments</b>	<b>32</b>
<b>Appendix A. Clustering coefficient of recurrence networks for one-dimensional maps</b>	<b>32</b>
<b>References</b>	<b>37</b>

### 1. Introduction

Since the early stages of quantitative nonlinear sciences, numerous conceptual approaches have been introduced for studying the characteristic features of dynamical systems based on observational time series [1]–[4]. Popular methods that are increasingly used in a variety of applications (see, for example, [5]) include (among others) Lyapunov exponents, fractal dimensions, symbolic discretization and measures of complexity such as entropies and quantities derived from them. 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.

As a particular concept, the basic ideas of which originated in the pioneering work of Poincaré in the late 19th century [6], the quantification of recurrence properties in phase space has recently attracted considerable interest [7]. One particular reason for this is that these recurrences can be easily visualized (and subsequently quantified in a natural way) by means of the so-called recurrence plots obtained from a single trajectory of the dynamical system under study [8, 9]. When observing this trajectory as a scalar time series  $x(t)$  ( $t = 1, \dots, N$ ), one may use a suitable  $m$ -dimensional time delay embedding of  $x(t)$  with delay  $\tau$  [10],

$\mathbf{x}^{(m)}(t) = (x(t), x(t + \tau), \dots, x(t + (m - 1)\tau))$ , for obtaining a recurrence plot as a graphical representation of the binary recurrence matrix

$$R_{i,j}(\epsilon) = \Theta(\epsilon - \|\mathbf{x}_i - \mathbf{x}_j\|), \quad (1)$$

where  $\Theta(\cdot)$  is the Heaviside function,  $\|\cdot\|$  denotes a suitable norm in the considered phase space and  $\epsilon$  is a threshold distance that should be reasonably small (in particular, much smaller than the attractor diameter [9, 11, 12]). To simplify our notation, we have used the abbreviation  $\mathbf{x}_i = \mathbf{x}^{(m)}(t = t_i)$  (with  $t_i$  being the point in time associated with the  $i$ th observation recorded in the time series<sup>7</sup>) wherever appropriate.

Experimental time series often yield a recurrence plot displaying complex structures, which are particularly visible in sets of recurrence points (i.e.  $R_{i,j}(\epsilon) = 1$ ) forming diagonal or vertical ‘line’ structures. A variety of statistical characteristics of the length distributions of these lines (such as maximum, mean, or Shannon entropy) can be used for defining additional quantitative measures that characterize different aspects of dynamic complexity of the studied system in more detail. This conceptual framework is known as recurrence quantification analysis (RQA) [13]–[15] and is nowadays frequently applied to a variety of real-world applications of time series analysis in diverse fields of research [16]. However, most of these RQA measures are sensitive to the choice of embedding parameters, which are found to sometimes induce spurious correlations in a recurrence plot [17].

Recent studies have revealed that the fundamental invariant properties of a dynamical system (i.e. its correlation dimension  $D_2$  and correlation entropy  $K_2$ ) are conserved in the recurrence matrix [18, 19], the estimation of which is independent of the particular embedding parameters. What is more, recurrence plots preserve all the topologically relevant phase space information of the system, such that we can completely reconstruct a time series from its recurrence matrix (modulo some rescaling of its probability distribution function) [20, 21].

A further appealing paradigm for analysing the structural features of complex systems is based on their representation as complex networks of passive or active (i.e. mutually interacting) subsystems. For this purpose, classical graph theory has been systematically extended by a large variety of different statistical descriptors of the topological features of such networks on local, intermediate and global scales [22]–[24]. These measures have been successfully applied for studying real-world networks in various scientific disciplines, including the structural properties of infrastructures [25], biological [26], ecological [27] and climate networks [28, 29], to give some prominent examples. The corresponding results have triggered substantial progress in our understanding of the interplay between the structure and dynamics of such complex networks [30]–[32].

The great success of network theory in various fields of research has recently motivated first attempts to generalize this concept for a direct application to time series [33]–[43]. By means of complex network analysis, important complementary features of dynamical systems (i.e. properties that are not captured by existing methods of time series analysis) can be resolved, which are based on spatial dependences between individual observations instead of temporal correlations. For example, a corresponding idea has been successfully applied for describing causal signatures in seismic activity by means of networks of recurrent events [44]–[46]. However, concerning applications to the field of time series analysis, a substantial number of the recently suggested techniques have certain conceptual limitations, which make them suitable

<sup>7</sup> Note that unlike many other methods of time series analysis, the concept of recurrence plots does not require observations that are equally spaced in time.

only for dealing with distinct types of problem. As an alternative that may provide a unifying conceptual and practical framework for nonlinear time series analysis using complex networks, we reconsider the concept of recurrences in phase space for defining complex network structures directly based on time series. For this purpose, it is straightforward to interpret the recurrence matrix  $\mathbf{R}(\epsilon)$  as the adjacency matrix  $\mathbf{A}(\epsilon)$  of an unweighted and undirected complex network, which we suggest to call the *recurrence network* associated with a given time series. To be more specific, the associated adjacency matrix is given by

$$A_{i,j}(\epsilon) = R_{i,j}(\epsilon) - \delta_{i,j}, \quad (2)$$

where  $\delta_{i,j}$  is the Kronecker delta introduced here in order to avoid artificial self-loops, corresponding to a specific choice of the Theiler window in RQA [9]. A related conceptual idea has recently been independently suggested by different authors [35, 36, 39, 40, 42], but not yet systematically studied. In this work, however, we aim to give a rigorous and detailed interpretation of a variety of quantitative characteristics of recurrence networks, yielding novel concepts for statistically evaluating distinct phase space structures captured in recurrence plots. In this sense, our presented work continues and complements the developments in the field of RQA within the last two decades [9, 43]. It shall be noted that a generalization to weighted networks (as partially studied in [33, 34]) is straightforward if the recurrence matrix is replaced by the associated distance matrix between pairs of states. In any case, recurrence networks based on the mutual phase space distances of observational points on a single trajectory are spatial networks, i.e. fully embedded in an  $m$ -dimensional space, which has important implications for their specific topological features. We will raise this point in detail within this paper.

The consideration of recurrence plots as graphical representations of complex networks allows a reinterpretation of many network-theoretic measures in terms of characteristic phase space properties of a dynamical system. According to the ergodicity hypothesis, we suppose that we can gain full information about these properties by either ensembles of trajectories, or sufficiently long observations of a single trajectory. Following this line of ideas, we approximate the (usually unknown) invariant density  $p(\mathbf{x})$  (which is related to the associated invariant measure  $\mu$  by  $d\mu = p(\mathbf{x})d\mathbf{x}$ ) of the studied system by some empirical estimate  $\hat{p}^{(\epsilon)}(\mathbf{x})$  obtained from a time series, where  $\epsilon$  defines the level of coarse-graining of phase space involved in this procedure. Transforming the time series into a recurrence network then allows us to quantitatively characterize the higher-order statistical properties (i.e. properties that are based on joint features of *different* states) of the invariant density  $p(\mathbf{x})$  by means of network-theoretic measures.

According to the above argumentation, quantitative descriptors of the topological features of recurrence networks can be considered as novel and complementary measures. Specifically, in contrast to the estimation of both dynamical invariants (in particular,  $K_2$ ,  $D_2$  and related nonlinear characteristics), as well as traditional RQA measures that are exclusively based on line structures in recurrence plots (i.e. the presence of temporal correlations between different parts of a trajectory), the network-theoretic approach to quantitatively characterizing recurrence plots is distinctively different as it relates to spatial dependences between state vectors in some appropriate phase space. As we will demonstrate in this paper, our technique exhibits additional deep insights into some phase space properties of dynamical systems such as the local attractor fragmentation or the presence of dynamically invariant objects, which are directly related to their complex dynamics, but are not provided by other existing methods. For this purpose, we will take seriously the duality of adjacency matrices of complex networks, on the one hand,

**Table 1.** Summary of definitions of vertices, criteria for the existence of edges, and the directionality of edges in existing complex network approaches to time series analysis (given in chronological order).

Method	Vertex	Edge	Directed
Transition networks (2005)	Discrete state $s(t)$	Transition probability between states	Yes
Cycle networks (2006)	Cycle	Correlation between cycles	No
Correlation networks (2008)	State $\mathbf{x}(t)$	Correlation between state vectors	No
Visibility graphs (2008)	Scalar state $x(t)$	Mutual visibility of states	No
$k$ -nearest neighbour networks (2008)	State $\mathbf{x}(t)$	Recurrence of states (mass)	Yes
Recurrence networks (2008)	State $\mathbf{x}(t)$	Recurrence of states (volume)	No

and recurrence matrices of dynamical systems, on the other hand, which would in turn also allow transferring concepts from dynamical systems theory (given that the corresponding recurrence plot-based estimates are invariant under temporal reordering) to complex networks. In this work, however, we will concentrate on a detailed discussion of how phase space properties can be further quantified in terms of network theory.

The remainder of this paper is organized as follows. Section 2 presents a review of existing approaches for extracting complex networks from time series, including a comprehensive discussion of their potentials and possible conceptual problems (with a special emphasis on how to interpret the resulting networks' topological properties). The concept of recurrence networks as a natural alternative is further discussed in section 3. In particular, it is demonstrated that many network-theoretic measures yield sophisticated quantitative characteristics corresponding to certain phase space properties of a dynamical system that have not yet been explicitly studied in terms of other dynamical invariants or measures of complexity based on RQA. In order to support our theoretical considerations, section 4 provides some examples of how network-theoretic measures reveal phase space properties of various low-dimensional dynamical systems. Finally, we summarize our main results and outline future directions of further methodological developments based on our proposed technique.

## 2. Approaches for transforming time series into complex networks— a comparative review

In this section, a brief review and classification of existing approaches for studying the properties of time series by means of complex network methods is presented in chronological order (see table 1). In particular, the strengths and possible limitations of the existing approaches will be discussed.

### 2.1. Partitioning of phase space: transition networks

The concept of symbolic dynamics [47] allows characterizing the properties of a dynamical system based on a partition  $\{S_1, \dots, S_K\}$  of its phase space into  $K$  mutually disjoint sets, yielding a transformation of every possible trajectory into an (in principle, infinite) sequence of abstract symbols. Formally, the proper application of concepts of symbolic time series analysis (such as mutual information or other entropic quantities) requires the existence of a generating

partition that corresponds to a unique assignment of symbolic sequences (i.e. sequences of class identifiers) to every trajectory of the system. Note that this prerequisite is usually violated in real-world applications due to the presence of noise; however, even in the ideal noise-free case, generating partitions either do not exist or can hardly be estimated (see [48] and references therein). Nevertheless, applications of symbolic time series analysis have recently attracted considerable interest in numerous applications [49]–[52].

Partitioning of the phase space of a dynamical system can also be used for transforming a time series into a complex network representation. In the simplest possible case, we identify the different sets  $S_i$  with the vertices of a network and consider the  $n$ -step transition probabilities

$$p_{i,j}^{(n)} = P(\mathbf{x}_{t+n} \in S_j | \mathbf{x}_t \in S_i) \quad (3)$$

for characterizing the edges of a fully connected weighted and directed graph [53]. By setting a suitable threshold  $p_{\min}$  to these transition probabilities, we obtain the adjacency matrix of an associated unweighted, directed network, which we suggest referring to as the  $n$ -step transition network, by setting

$$A_{i,j}(p_{\min}; n) = \Theta(p_{i,j}^{(n)} - p_{\min}). \quad (4)$$

We note that related graph-theoretic methods have been recently used for the set-oriented approximation of almost invariant sets in dynamical systems [54, 55]. Recent applications of transition networks in the field of time series analysis include the characterization of transitions in cellular automaton models of vehicular traffic by means of the associated degree distributions [56]–[59] (which has been recently generalized to the consideration of weighted networks [60]), as well as the analysis of stock exchange time series [61, 62] in terms of identifying vertices with the highest relevance for information transfer using betweenness centrality and inverse participation ratio.

From the time series analysis point of view, the main disadvantage of the transition network approach is that it induces a significant loss of information on small amplitude variations. In particular, two observations with even very similar values are not considered to belong to the same class if they are just separated by a class boundary. This might influence the quantitative features of a corresponding network, as it is not exclusively determined by the widths of the individual classes, but also by their specific definition. Corresponding results are known for other methods of symbolic time series analysis such as estimates of mutual information [63, 64] or block entropies [52, 65]. In this respect, the recurrence network approach introduced in this work is more objective as it only depends on a single parameter  $\epsilon$ . Note, however, that coarse-graining might be a valid approach in the case of noisy real-world time series, where extraction of dynamically relevant information hidden by the noise can be supported by grouping the data.

## 2.2. Cycle networks for pseudo-periodic time series

In 2006, Zhang and Small [33]–[35] suggested to study the topological features of pseudo-periodic time series (representing, for example, the dynamics of chaotic oscillators like the Rössler system) by means of complex networks. For this purpose, individual cycles (defined by minima or maxima of the studied time series) have been considered as vertices of an undirected network, and the connectivity of pairs of vertices has been established by considering a generalization of the correlation coefficient to cycles of possibly different length or, alternatively, their phase space distance.

A potential point of criticism concerning this method is that the definition of a cycle is not necessarily straightforward in complex oscillatory systems. In [33]–[35], the authors mainly considered nonlinear oscillators in their phase-coherent regimes; however, it is not clear how a cycle could be defined for non-phase-coherent oscillations, for example, in the Funnel regime of the Rössler system. The same problem arises for systems with multiple timescales, which are hence hard to treat this way. Furthermore, it is not intuitively clear how to interpret correlations of cycles, since the values of the corresponding measures are not exclusively determined by the proximity of the corresponding parts of the trajectory in phase space, but depend also on the specific choice of sampling. This could yield rather different estimates of the cycle correlation coefficient (or, alternatively, the phase space distance) between two cycles, even if the two parts of the trajectory are very close to each other.

### 2.3. Correlation networks of embedded state vectors

A generalization of the cycle networks method used by Zhang and Small that can also be applied to time series without obvious oscillatory components has been suggested by Yang and Yang [37] using a simple embedding of an arbitrary time series. In their formalism, individual state vectors  $\mathbf{x}_i^{(m)}$  in the  $m$ -dimensional phase space of the embedded variables are considered as vertices, from which the Pearson correlation coefficient can be easily computed as

$$r_{i,j} = r(\mathbf{x}_i^{(m)}, \mathbf{x}_j^{(m)}) = \frac{\sum_{k=1}^m \tilde{x}(t_i + (k-1)\tau) \tilde{x}(t_j + (k-1)\tau)}{\sqrt{\sum_{k=1}^m \tilde{x}^2(t_i + (k-1)\tau)} \sqrt{\sum_{k=1}^m \tilde{x}^2(t_j + (k-1)\tau)}}, \quad (5)$$

with

$$\tilde{x}(t_i + (k-1)\tau) = x(t_i + (k-1)\tau) - \frac{1}{m} \sum_{k=1}^m x(t_i + (k-1)\tau) \quad (6)$$

being the zero-mean components of the respective embedding vectors (for scalar time series, the generalization to vector-valued time series is straightforward). If  $r_{i,j}$  is larger than a given threshold, the vertices  $i$  and  $j$  are considered to be connected, resulting in an undirected network representation. An equivalent approach called fluid-dynamic complex networks (FDCN) has recently been proposed by Gao and Jin [42, 66] and successfully applied for characterizing the nonlinear dynamics of conductance fluctuating signals in a gas–liquid two-phase flow.

One potential conceptual problem of this particular technique is that the consideration of correlation coefficients between two phase space vectors usually requires a sufficiently large embedding dimension  $m$  for a proper estimation with low uncertainty (more specifically, the standard error of the correlation coefficient is approximately proportional to  $1/\sqrt{m-1}$ ). Hence, local information about the short-term dynamics captured in a time series might get lost when following this approach. Even more, since embedding is known to induce spurious correlations to a system under study, the results of the correlation method of network construction may suffer from related effects.

With respect to the interpretation of the resulting network patterns, we note that for vertices corresponding to mutually overlapping time series segments, the consideration of correlation coefficients, as applied in the papers cited above, corresponds to studying the local autocorrelation function of the signal. Hence, the presence of edges between these vertices is exclusively determined by linear correlations within the signals. In principle, we might think of replacing the correlation coefficient by other measures of interrelationships such as the mutual

information, which are also sensitive to general statistical dependences [28, 29, 67]; however, the appropriate estimation of such nonlinear quantities would require a considerably larger amount of data, i.e. a very large embedding dimension  $m$ .

Finally, when studying time series with pronounced cycles (like trajectories of the Rössler system), different cycles could be included in one embedding vector (depending on the sampling rate), which casts additional doubts with respect to the direct interpretability of the resulting network properties.

#### 2.4. Visibility graphs

An alternative to the latter two threshold-based concepts has been suggested by Lacasa *et al* [38] in terms of the so-called visibility graph. In this formalism, individual observations are considered as vertices, and edges are introduced whenever a partial convexity constraint is fulfilled, i.e.  $x(t_a)$  and  $x(t_b)$  are connected if for all states  $x(t_c)$  with  $t_a < t_c < t_b$ ,

$$\frac{x_a - x_c}{t_c - t_a} > \frac{x_a - x_b}{t_b - t_a} \quad (7)$$

holds. Since the visibility condition is symmetric with respect to  $a$  and  $b$ , visibility graphs are undirected. Visibility graphs have been used to study the behaviour of certain fractal as well as multifractal stochastic processes [68, 69], energy dissipation in three-dimensional turbulence [70] and the nonlinear properties of financial [71, 72] and environmental time series [73]. Recently, a slightly modified approach of horizontal visibility graphs has been proposed and applied to certain random time series [74].

Although visibility graphs are easily established and allow an alternative estimation of the Hurst exponent of fractal time series, an interpretation of the convexity constraint in terms of other phase space properties of the considered system has not yet been provided. Moreover, in its present formulation, the application of this approach is restricted to univariate time series.

#### 2.5. Complex networks based on neighbourhood relations in phase space

As it has already been mentioned, the transformation of time series into complex networks by means of neighbourhood relationships has already been discussed by different authors. In particular, there are two possible approaches that can directly be related to slightly different definitions of recurrence plots [9].

On the one hand, a neighbourhood can be defined by a fixed number  $k$  of nearest neighbours of a single observation, i.e. a constant ‘mass’ of the considered environments [35, 36]. We refer to this method as a  $k$ -nearest neighbour network in phase space. This setting implies that the degrees  $k_v$  of all vertices in the network are kept fixed at the same value  $k$ . Hence, information about the local geometry of the phase space, which is mainly determined by the invariant density  $p(\mathbf{x})$ , cannot be directly obtained by most traditional complex network measures (see section 3)<sup>8</sup>. Note that the adjacency matrix of a  $k$ -nearest neighbour network is in general not symmetric, i.e. the fact that a vertex  $j$  is among the  $k$ -nearest neighbours of a vertex  $i$  does not imply that  $i$  is also among the  $k$ -nearest neighbours of  $j$ . Hence,  $k$ -nearest neighbour networks are directed networks.

<sup>8</sup> As an alternative measure, we could consider the inverse distance of the  $k$ th nearest neighbour as a measure for phase space density.

On the other hand, we can define the neighbourhood of a single point in phase space by a fixed phase space distance, i.e. considering a constant ‘volume’ [39, 42, 43, 75]. This approach has the advantage that the degree centrality  $k_v$  gives direct information about the local phase space density (see section 3). The consideration of neighbourhood relationships within a fixed phase space volume corresponds to the standard definition of a recurrence plot as mentioned in the introduction section. The resulting networks will therefore be referred to as recurrence networks in the following. Unlike the  $k$ -nearest neighbour networks, these recurrence networks are undirected graphs by definition. We note that Gao and Jin [42] termed an equivalent approach as fluid–structure complex networks (FSCN) and used it for analysing gas–liquid two-phase flow and the Lorenz system as a toy model in terms of link density. In addition, they related their observations to the presence of unstable periodic orbits (UPOs) in a dynamical system. We will come back to this point in section 4.3.

We note that since all complex network approaches based on the proximity of different parts of the trajectory (i.e. cycle, correlation,  $k$ -nearest neighbour and recurrence networks) do not preserve information about the temporal order of the respective state vectors (i.e. are invariant with respect to random permutations of these vectors), the resulting characteristics of network topology represent dynamically invariant properties related to the state density in the corresponding (abstract) phase space. As a consequence, considering a recurrence network based directly on a univariate time series (or a one-dimensional dynamical system), we cannot distinguish between stochastic and deterministic systems with the same phase space density (for a simple example, see appendix A.2). However, in such cases, embedding or consideration of phase space dimensions larger than one provides a feasible solution for this identification problem. Furthermore, as we will show in the following sections, the statistical characterization of recurrence networks provides useful and complementary insights into the phase space structures that cannot be obtained by other existing approaches of nonlinear time series analysis.

## 2.6. Recurrence networks as a unifying framework for complex network-based time series analysis

Following the considerations in the previous subsections, many existing methods for transforming time series into complex networks suffer from specific conceptual limitations. In particular, the concepts used for defining both vertices and edges of the networks, which differ across the various techniques, are in some cases rather artificial from a dynamical systems point of view (table 1).

In contrast to the other recently suggested approaches, the identification of a recurrence matrix with the adjacency matrix of a complex network is a straightforward and natural idea that conserves many local properties of the invariant density of the studied dynamical system captured in single observational time series. In particular, individual values of the respective observable can be directly considered as vertices of the recurrence network (similar to the visibility graph concept), while the existence of an edge serves as an indicator of a recurrence, i.e. pairs of states whose values do not differ by more than a small value  $\epsilon$  in terms of a suitable norm in phase space.

It should be pointed out that not only recurrence networks but also other existing methods of complex network-based time series analysis are based on the concept of recurrences in phase space. Apart from the  $k$ -nearest neighbour networks originating from the idea of a fixed local recurrence rate (i.e. a fixed mass of the neighbourhoods), the idea of considering a threshold

**Table 2.** Relationships between recurrence network entities and corresponding geometrical objects and their properties in phase space.

Recurrence network	Phase space
Vertex	State $\mathbf{x}(t)$
Edge	Recurrence of states
Path	Overlapping sequence of $\epsilon$ -balls

value to the proximity of two vertices can also be found in other previously suggested methods. In particular, correlation networks [37] (see section 2.3) can be regarded as a special case of recurrence networks where the usual metric distance has been replaced by the correlation distance [76]

$$d_C(\mathbf{x}_i, \mathbf{x}_j) = 1 - r_{i,j}, \quad (8)$$

where  $r_{i,j}$  is the correlation coefficient between two embedded state vectors defined in equation (5). Similar considerations apply to the case of cycle networks. Note, however, that the advantage of using the concept of recurrences of individual states in phase space defined by metric distances instead of correlations or cycle properties associated with state vectors of dynamical systems is that it allows for creating networks without any embedding or consideration of groups of states. On the one hand, this independence from a particular embedding is beneficial when dynamical invariants of the studied system are of interest. On the other hand, the statistical properties of the resulting recurrence networks reflect exclusively the invariant density of states in phase space (in terms of certain higher-order statistics), because time-ordering information is lost in this framework.

Following the above considerations, it can be argued that the concept of recurrence networks yields a unifying framework for transferring time series into complex networks in a dynamically meaningful way. In particular, this approach can be applied (i) to both univariate and multivariate time series (phase space trajectories), (ii) with and without pronounced oscillatory components and (iii) with as well as without embedding<sup>9</sup>. Moreover, similar to traditional RQA, studying network properties for sliding windows in time also allows for coping with non-stationary time series [43]. Consequently, unlike for most of the existing techniques, there are no fundamental restrictions with respect to its practical applicability to arbitrary time series.

### 3. Quantitative characteristics of recurrence networks

While the definitions of edges and vertices in our approach have already been given above (table 1), we now provide a geometrical interpretation of a third important network entity, the path, within the framework of recurrence networks (table 2). A path between two vertices  $i$  to  $j$  in a simple graph without multiple edges can be written as an ordered sequence of the vertices it contains, i.e.  $(i, k_1, \dots, k_{l_{i,j}-1}, j)$ , where the associated number of edges  $l_{i,j}$  measures the length of the path. In phase space, a path in the recurrence network is hence defined

<sup>9</sup> Possible embedding effects on recurrence plots have already been exhaustively treated in the literature [43], whereas a corresponding detailed investigation of recurrence networks remains a subject for future studies.

**Table 3.** Correspondence between recurrence network measures and phase space properties. Specific terms are discussed in the text.

Scale	Recurrence network	Phase space
Local	Edge density $\rho$ Degree centrality $k_v$	Global recurrence rate $RR$ Local recurrence rate $RR_v$
Intermediate	Clustering coefficient $C$ Local degree anomaly $\Delta k_v$ Assortativity $\mathcal{R}$ Matching index $\mu_{i,j}$	Invariant objects Local heterogeneity of phase space density Continuity of phase space density Twinness of $i, j$
Global	Average path length $\mathcal{L}$ Network diameter $\mathcal{D}$ Closeness centrality $c_v$ Betweenness centrality $b_v$	Mean phase space separation $\langle d_{i,j} \rangle_{i,j}$ Phase space diameter $\Delta$ Local centredness in phase space Local attractor fractionation

as a sequence of mutually overlapping  $\epsilon$ -balls  $B_\epsilon(\mathbf{x}_i), B_\epsilon(\mathbf{x}_{k_1}), \dots, B_\epsilon(\mathbf{x}_{k_{l_{i,j}-1}}), B_\epsilon(\mathbf{x}_j)$ , where  $B_\epsilon(\mathbf{x}_i) \cap B_\epsilon(\mathbf{x}_{k_1}) \neq \emptyset, \dots, B_\epsilon(\mathbf{x}_{k_{l_{i,j}-1}}) \cap B_\epsilon(\mathbf{x}_j) \neq \emptyset$ .<sup>10</sup>

Due to the natural interpretation of vertices, edges and paths, the topological characteristics of a recurrence network closely capture the fundamental phase space properties of the dynamical system that has generated the considered time series. In the following, we will present a detailed analysis of the corresponding analogies for different network properties that are defined on a local (i.e. considering only the direct neighbourhood of a vertex), intermediate (i.e. considering the individual neighbourhoods of the neighbours of a vertex) and global (i.e. considering all vertices) scale (table 3)<sup>11</sup>. It has to be emphasized that these quantities can be considered as (partly novel and complementary) measures within the framework of RQA.

### 3.1. Local network properties

**3.1.1. Degree centrality (local recurrence rate).** As a first measure that allows us to quantify the importance of a vertex in a complex network, the degree centrality [77] of a vertex  $v$ ,  $k_v$ , is defined as the number of neighbours, i.e. the number of vertices  $i \neq v$  that are directly connected with  $v$ :

$$k_v = \sum_{i=1}^N A_{v,i}. \quad (9)$$

Note that, in general, the sum is taken over all  $i \neq v$ . However, according to our definition (2), we skip the corresponding condition in the following. Normalizing this measure by the

<sup>10</sup> An  $\epsilon$ -ball centred at state vector  $\mathbf{x}$  is defined as the open set  $B_\epsilon(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^m : \|\mathbf{x} - \mathbf{y}\| < \epsilon\}$ .

<sup>11</sup> Alternatively, we could classify the corresponding phase space properties according to the fact whether they refer to individual points, small regions, or the entire phase space. In this respect, measures related to a single vertex  $v$  (centralities, local clustering coefficient and local degree anomaly) give local, those related to a specific edge  $(i, j)$  (shortest path length, matching index and edge betweenness) give intermediate, and all others give global information about the phase space properties.

maximum number of possible connections,  $N - 1$ , we obtain the local connectivity

$$\rho_v = \frac{1}{N-1} \sum_{i=1}^N A_{v,i} = RR_v, \quad (10)$$

which, from the recurrence plot point of view, corresponds to the local recurrence rate  $RR_v$  of the state  $v$ . Thus, degree centrality and local connectivity yield an estimator for the local phase space density, since for a vertex  $v$  located at position  $\mathbf{x}_v$  in phase space,

$$\frac{1}{N} (k_v(\epsilon) + 1) \approx \int_{B_\epsilon(\mathbf{x}_v)} d\mathbf{x} p(\mathbf{x}) \approx (2\epsilon)^m p(\mathbf{x}_v) \quad (11)$$

(when using the maximum norm) and, hence,

$$\hat{p}(\mathbf{x}_v) = \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{k_v + 1}{(2\epsilon)^m N}. \quad (12)$$

In complex network studies, one is often interested in the frequency distribution of degree centralities,  $P(k)$ , in particular, the presence of an algebraic scaling behaviour, which is a characteristic of scale-free networks [22]. In the case of recurrence networks, however, we note that the emergence of a scale-free property in the degree distribution requires distinct structures of the phase space density, which are probably only present in rather specific cases. Moreover, although several authors have recently focused their attention on this characteristic obtained from different types of complex networks derived from time series [33], [37]–[39], [42], [56]–[59], [68]–[70], we would like to underline that for a complete characterization of the phase space properties of a dynamical system, not only degree centralities but also other higher-order statistical measures have to be studied.

**3.1.2. Edge density (global recurrence rate).** In some situations, it is useful not to consider the full distribution of degree centralities in a network, but to focus on the mean degree of all vertices

$$\langle k \rangle = \frac{1}{N} \sum_{v=1}^N k_v = \frac{2L}{N}, \quad (13)$$

as a simple characteristic quantity of this distribution, where

$$L = \sum_{i < j} A_{i,j} = \rho \frac{N(N-1)}{2} \quad (14)$$

is the total number of edges in the recurrence network. The mean degree centrality  $\langle k \rangle$  is directly proportional to the edge density  $\rho$  of the network or, alternatively, its recurrence plot equivalent, the global recurrence rate  $RR$ ,

$$\begin{aligned} \rho(\epsilon) &= \frac{1}{N} \sum_{v=1}^N \rho_v(\epsilon) = \frac{2}{N(N-1)} \sum_{v < i} A_{v,i}(\epsilon) \\ &= \frac{2}{N(N-1)} \sum_{v < i} \Theta(\epsilon - \|\mathbf{x}_v - \mathbf{x}_i\|) \\ &= RR(\epsilon) = C_2(\epsilon) \sim \epsilon^{-D_2}. \end{aligned} \quad (15)$$

Note that the recurrence rate coincides with the definition of the correlation integral  $C_2(\epsilon)$ , which is commonly used to estimate the correlation dimension  $D_2$ , for example, using the Grassberger–Procaccia algorithm [78].

The connection between the edge density and the correlation dimension can be understood by the fact that the local recurrence rate  $RR_v$  of a vertex  $v$  corresponds to the measure of an  $m$ -dimensional ball  $B_\epsilon(\mathbf{x}_v)$  of radius  $\epsilon$  centred at the point  $\mathbf{x}_v$  in the  $m$ -dimensional phase space in the limit that time goes to infinity ( $N \rightarrow \infty$ ). When considering the Euclidean norm as a distance measure in phase space, these balls are defined as hyperspheres, for the maximum norm as hypercubes, etc. Then, the pointwise (information) dimension of the probability measure  $\mu$  at  $\mathbf{x}_v$  is defined as  $D_p(\mathbf{x}_v) = -\lim_{\epsilon \rightarrow 0} (\ln \mu(B_\epsilon(\mathbf{x}_v)))/\ln \epsilon$  [4]. Owing to the heterogeneity of the phase space visited by the trajectory (i.e. the non-uniform phase space density that results in different degree centralities  $k_v$  in different parts of this space), the proper estimation of  $D_p$  is a nontrivial task and often requires expensive computational power and a high data quality and quantity. Thus, we expect a better statistics for  $D_2$ , since it more heavily weights regions of the phase space that have a higher probability measure  $\mu$ . Although the correlation integral has been well established in the literature for estimating the correlation dimension, we point out the improvement in estimating  $D_2$  based on the diagonal lines in  $R_{i,j}(\epsilon)$ , which yields an algorithm that is independent of the embedding parameters [19]. Consequently, the recurrence network representation  $A_{i,j}$  of a time series fully conserves the geometric properties of the phase space of the underlying dynamical system.

### 3.2. Intermediate scale network properties

**3.2.1. Local clustering coefficient.** The clustering coefficient of a vertex  $v$ ,  $\mathcal{C}_v$ , characterizes the density of connections in the direct neighbourhood of this vertex in terms of the density of connections between all vertices that are incident with  $v$ . In many networks, such loop structures formed by three vertices occur more often than one would expect for a completely random network. Hence, high clustering coefficients reveal a specific type of structure in a network, which is related to the cliquishness of a vertex [24].

In this work, we consider the definition of clustering coefficient proposed by Watts and Strogatz [24],

$$\mathcal{C}_v = \frac{2}{k_v(k_v - 1)} N_v^\Delta, \quad (16)$$

where  $N_v^\Delta$  is the total number of closed triangles including vertex  $v$ , which is bound by the maximum possible value of  $k_v(k_v - 1)/2$ . For vertices of degree  $k_v = 0$  or 1 (isolated or tree-like points, respectively), the clustering coefficient is defined as  $\mathcal{C}_v = 0$ , as such vertices cannot participate in triangles by definition.

Equation (16) can be rewritten in terms of conditional probabilities as

$$\mathcal{C}_v = P(A_{i,j} = 1 | A_{v,i} = 1, A_{v,j} = 1) = \frac{P(A_{i,j} = 1, A_{v,i} = 1, A_{v,j} = 1)}{P(A_{v,i} = 1, A_{v,j} = 1)} \quad (17)$$

using Bayes' theorem with

$$P(A_{v,i} = 1, A_{v,j} = 1) = \frac{1}{(N-1)(N-2)} \sum_{i=1}^N \sum_{j=1, j \neq i}^N A_{v,i} A_{v,j} \quad (18)$$

and a similar expression for  $P(A_{i,j} = 1, A_{v,i} = 1, A_{v,j} = 1)$ . As for a recurrence network, the value of  $A_{i,j}$  depends only on the phase space distance and the choice of  $\epsilon$ ; the latter relationship can be used to derive analytical results at least for one-dimensional systems based on their invariant density. Corresponding details can be found in appendix A, including the corresponding treatment of the Bernoulli and logistic maps as specific examples.

**3.2.2. Global clustering coefficient.** As for the degree centrality, we consider the average value of the clustering coefficient taken over all vertices of a network, the so-called global clustering coefficient

$$\mathcal{C} = \frac{1}{N} \sum_{v=1}^N \mathcal{C}_v, \quad (19)$$

as a global characteristic parameter of the topology of a network. We expect that the value of  $\mathcal{C}$  is—for a given dynamical system with a phase space density  $p(\mathbf{x})$ —in the asymptotic limit  $N \rightarrow \infty$  exclusively determined by the choice of  $\epsilon$ , which defines the scale of resolution. A more detailed discussion of the corresponding effects and their implications for certain model systems will be given in section 4.

**3.2.3. Mean nearest neighbour degree.** The mean nearest neighbour degree  $k_v^{\text{nn}}$  of vertex  $v$  gives the average degree in the neighbourhood of  $v$ ,

$$k_v^{\text{nn}} = \frac{1}{k_v} \sum_{i=1}^N A_{v,i} k_i. \quad (20)$$

The degree centrality  $k_v$  is a measure of the density of states in the immediate neighbourhood of state  $v$ , whereas  $k_v^{\text{nn}}$  can be interpreted to indicate the mean density of states in the next neighbourhood (next topological shell of neighbours) of state  $v$ . Hence, both measures taken together contain information about the local density anomaly in the vicinity of  $v$ , which we propose to measure by the local degree anomaly

$$\Delta k_v = k_v - k_v^{\text{nn}}. \quad (21)$$

Vertices with a positive degree anomaly ( $\Delta k_v > 0$ ) hence indicate local maxima of phase space density, while those with  $\Delta k_v < 0$  correspond to local density minima. Hence, the local degree anomaly may be considered as a proxy for the local heterogeneity of the phase space density. In a similar way, the average absolute value of the local degree anomaly,  $\langle |\Delta k_v| \rangle_v$ , serves as a measure of the overall spatial heterogeneity of the phase space density profile.

**3.2.4. Assortativity.** A network is called assortative if vertices tend to connect preferentially to vertices of a similar degree  $k$ . On the other hand, it is called disassortative if vertices of high degree prefer to connect to vertices of low degree, and vice versa. Hence, assortativity can be quantified by the Pearson correlation coefficient of the vertex degrees on both ends of all edges [24, 79],

$$\mathcal{R} = \frac{\frac{1}{L} \sum_{j>i} k_i k_j A_{i,j} - \left[ \frac{1}{L} \sum_{j>i} \frac{1}{2}(k_i + k_j) A_{i,j} \right]^2}{\frac{1}{L} \sum_{j>i} \frac{1}{2}(k_i^2 + k_j^2) A_{i,j} - \left[ \frac{1}{L} \sum_{j>i} \frac{1}{2}(k_i + k_j) A_{i,j} \right]^2}. \quad (22)$$

If the density of states in phase space hardly varies within an  $\epsilon$ -ball, the degrees on either ends of an edge will tend to be similar and hence the assortativity coefficient  $\mathcal{R}$  will be positive. This means that the more continuous and slowly changing the density of states is, the closer  $\mathcal{R}$  will be to its maximum value one. Within the framework of recurrence networks,  $\mathcal{R}$  can hence be interpreted as a measure of the continuity of the density of states or put differently of the fragmentation of the attractor. Note that this aspect has not yet been specifically addressed by other nonlinear measures, in particular, within the RQA framework.

**3.2.5. Matching index (*twinness*).** The overlap of the neighbourhood spaces of two vertices  $i, j$  is measured by the matching index

$$\mu_{i,j} = \frac{\sum_{l=1}^N A_{i,l} A_{j,l}}{k_i + k_j - \sum_{l=1}^N A_{i,l} A_{j,l}}, \quad (23)$$

where  $\mu_{i,j} = 0$  if there are no common neighbours, and  $\mu_{i,j} = 1$  if the neighbourhoods coincide [24]. Using the notion of  $\epsilon$ -balls around points in phase space, we can alternatively write

$$\mu_{i,j} \propto \frac{\mu(B_\epsilon(\mathbf{x}_i) \cap B_\epsilon(\mathbf{x}_j))}{\mu(B_\epsilon(\mathbf{x}_i)) + \mu(B_\epsilon(\mathbf{x}_j)) - \mu(B_\epsilon(\mathbf{x}_i) \cap B_\epsilon(\mathbf{x}_j))}. \quad (24)$$

Owing to the spatial constraints of the recurrence network, the neighbourhood spaces of  $i, j$  can only overlap if

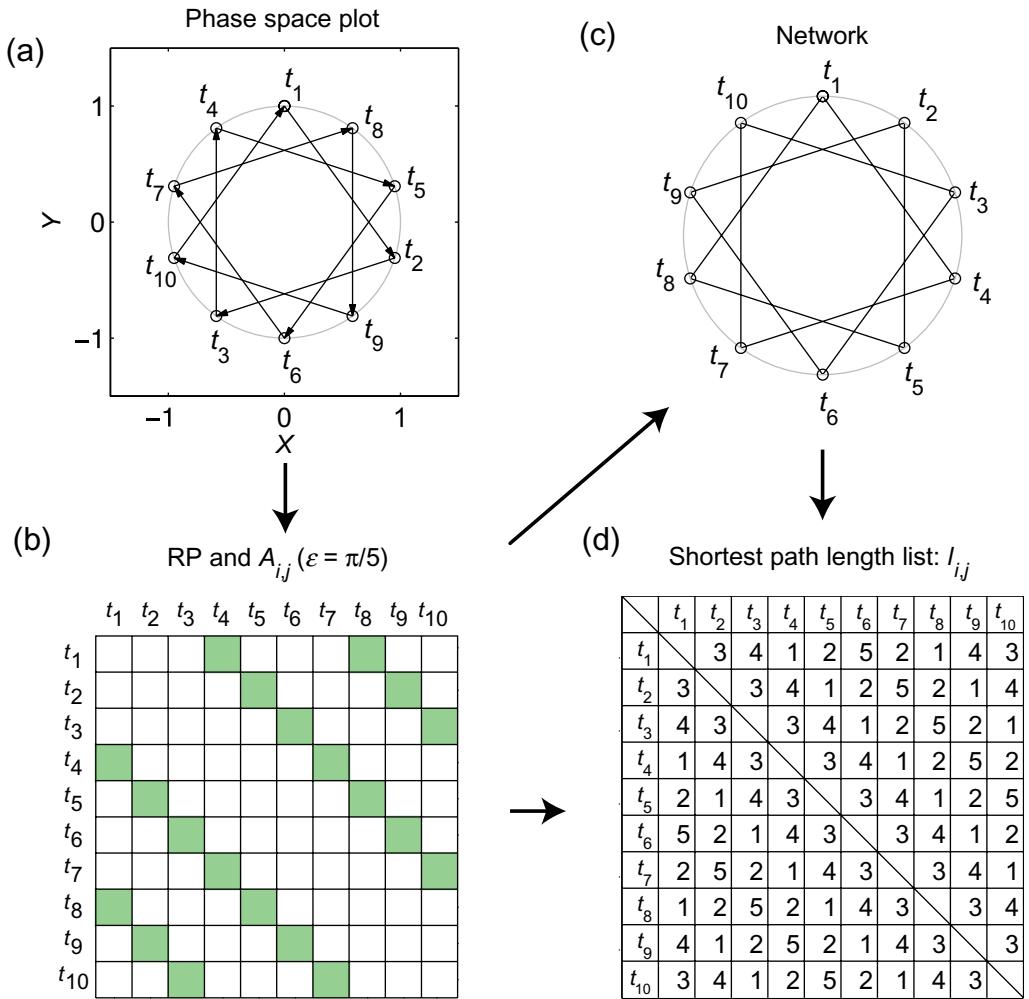
$$d_{i,j} = \|\mathbf{x}_i - \mathbf{x}_j\| \leq 2\epsilon, \quad (25)$$

i.e.  $\mu_{i,j} = 0$  for all pairs of vertices  $(i, j)$  with  $d_{i,j} > 2\epsilon$ . Moreover, the matching index  $\mu_{i,j}$  decreases on average with increasing spatial distance  $d_{i,j}$  between the two considered states. Note that since  $A_{i,j} = 0$  already for  $d_{i,j} > \epsilon$ , unconnected points with a matching index  $\mu_{i,j} > 0$  may exist.

The matching index of pairs of vertices in a recurrence network is closely related to the concept of *twins* [80], which has recently been successfully applied for constructing surrogate data (twin surrogates) in the context of statistical hypothesis testing for the presence of complex synchronization [81, 82]. Twins are defined as two states of a complex system that share the same neighbourhood in phase space, i.e. the two vertices of the recurrence network representing these states have a matching index  $\mu_{i,j} = 1$ . Hence, the matching index can be used for identifying candidates for twins. Note that pairs of vertices  $i$  and  $j$  in a recurrence network with  $\mu_{i,j} \lesssim 1$  can still be considered as *potential* twins, since  $\mu_{i,j} = 1$  is in some cases approached by only slight changes of the threshold  $\epsilon$ . Consequently, we suggest interpreting the matching index as a measure of the *twinness* of  $i$  and  $j$ . Furthermore, it should be noted that adjacent pairs of edges  $(i, j)$  ( $A_{i,j} = 1$ ) with a low matching index  $\mu_{i,j} \simeq 0$  connect two distinct regions of the attractor and may therefore be indicative of geometrical bottlenecks in the dynamics (cf our discussion of the betweenness centrality in section 3.3.5).

### 3.3. Global network properties

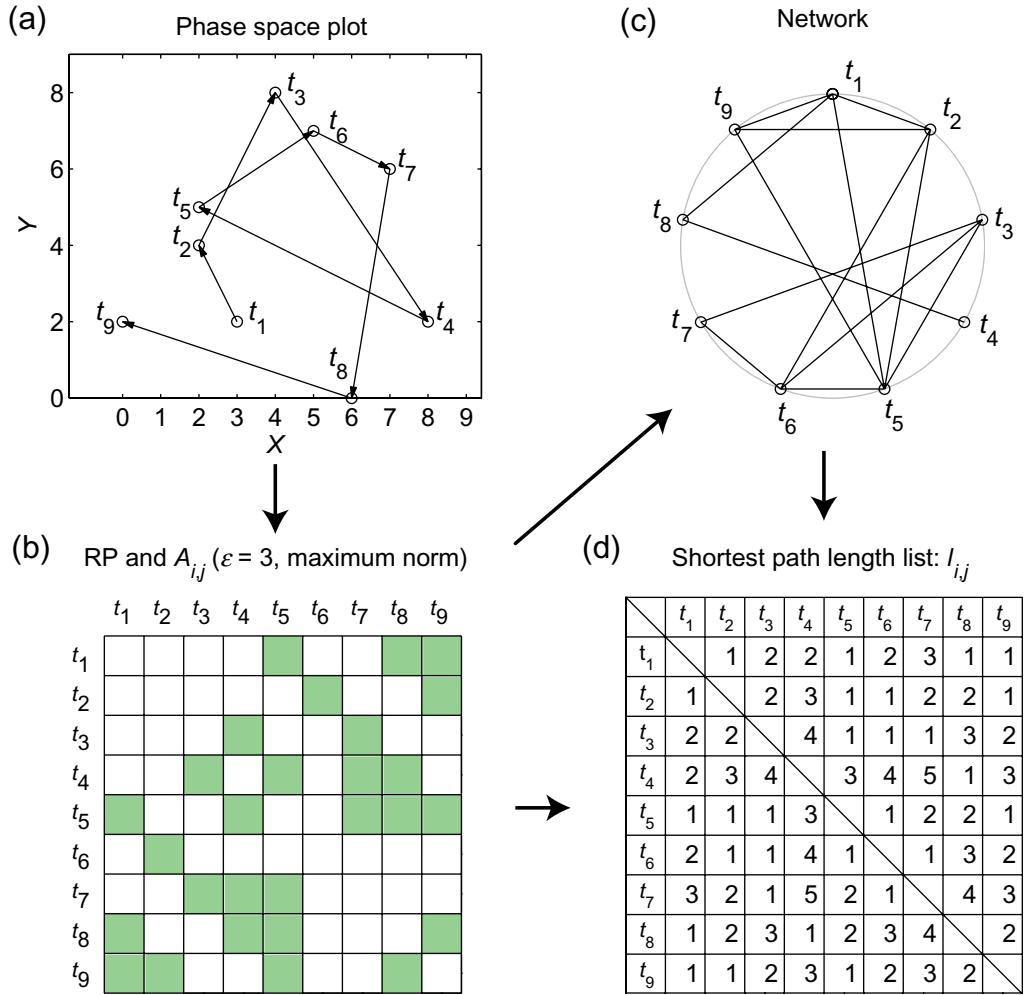
**3.3.1. Shortest path length.** As we consider recurrence networks as undirected and unweighted, we assume all edges to be of unit length in terms of graph (geodesic) distance. Consequently, the distance between any two vertices of the network is defined as the length of the shortest path between them. Note that time information is lost after transforming the



**Figure 1.** Schematic representation of the transformation of (a) a periodic trajectory in phase space (arrows indicate the temporal order of observations) into (c) the recurrence network (lines illustrate the mutual neighbourhood relations). (b) The associated recurrence plot representation. (d) A table listing the resulting shortest path lengths between any two vertices.

trajectory into a network presentation. Therefore, the terminology of the shortest path length  $l_{i,j}$  in the recurrence network reflects the minimum number of edges that have to be passed on a graph from a vertex  $i$  to a vertex  $j$ . In the same spirit,  $l_{i,j}$  is related to the distance of states  $i$  and  $j$  in phase space.

In order to better understand the meaning of shortest path lengths, let us study their calculation for two toy model series: first, we consider a periodic trajectory  $x(t) = \sin(6\pi \cdot 0.1t)$ ,  $y(t) = \cos(6\pi \cdot 0.1t)$ , with  $t = 0, 1, \dots, 10$ , i.e. there are  $N = 10$  points in the phase space (figure 1(a)). The corresponding recurrence plot is shown for  $\epsilon = \frac{\pi}{5}$  (figure 1(b)). As it has already been mentioned above, the recurrence matrix  $R_{i,j}$  and the adjacency matrix  $A_{i,j}$  of the associated recurrence network are basically equivalent. Adopting a common visualization of connectivity patterns from the literature on complex networks, we illustrate the recurrences



**Figure 2.** The same as in figure 1 for a general non-periodic trajectory (3, 2, 4, 8, 2, 5, 7, 6, 0, 2) embedded in a two-dimensional phase space ( $X, Y$ ) = ( $x_t, x_{t+1}$ ).

of the considered model time series by placing the individual observations (vertices) on a circle with equal common distances (figure 1(c)). In this representation, the shortest path length (in the network sense) between two vertices  $i$  and  $j$  corresponds to the smallest number of ‘jumps’ in phase space via pairs of neighbours (i.e. recurrences) in phase space. Obviously, the number of such jumps is determined by the prescribed value of  $\epsilon$  and the spatial distance between  $i$  and  $j$ . For instance, the shortest path from vertex  $i = 1$  to  $j = 10$  is  $l_{1,10} = 3$  as indicated by the matrix of mutual shortest path lengths (figure 1(d)). Note that this list is symmetric by definition, i.e.  $l_{i,j} = l_{j,i}$ . The same heuristic analysis can also be performed for a general non-periodic trajectory in phase space as shown in figures 2(a)–(d).

We wish to underline that the terminology of shortest path lengths in networks does not have a direct relevance for the dynamical evolution of the observed system. In contrast,  $l_{i,j}$  measures distances among a discrete set of points on the attractor in units of the neighbourhood size  $\epsilon$ . For example, in the periodic case displayed in figure 1, it takes nine iterations (time

points) from vertex 1 to vertex 10 in the time domain, while the shortest path to cover the phase space distance has only a length of  $l_{1,10} = 3$ . Hence, shortest paths do not allow to infer the temporal evolution of the system. Even more, for the path concept in a recurrence network, no information about the temporal order of the individual observations is considered (for example, the shortest path between vertices 1 and 2 in figure 1 is given by the sequence (1, 8, 5, 2), which is not ordered in time).

One should note that if the phase space is strongly fragmented (for instance, in the period-3 window of the logistic map, which has been discussed elsewhere [43], the phase space consists of three discrete points), the resulting recurrence networks can be composed of different disconnected clusters. Furthermore, there might be more than one shortest path connecting two nodes. For example, in the aperiodic example in figure 2, the shortest path from node 1 to node 7 ( $l_{1,7} = 3$  as shown in figure 2(d)) can be obtained by three different choices: (1, 2, 6, 7), (1, 5, 3, 7) and (1, 5, 6, 7).

**3.3.2. Average path length.** The average path length  $\mathcal{L}$  is defined as the mean value of the shortest path lengths  $l_{i,j}$  taken over all pairs of vertices  $(i, j)$ ,

$$\mathcal{L} = \langle l_{i,j} \rangle = \frac{2}{N(N-1)} \sum_{i < j} l_{i,j}. \quad (26)$$

Here, for disconnected pairs of vertices, the shortest path length is set to zero by definition. Note that in most practical applications, this has no major impact on the corresponding statistics.

The average phase space separation of states  $\langle d_{i,j} \rangle_{i,j}$  serves as an  $\epsilon$ -lower bound to  $\mathcal{L}$ , since

$$d_{i,j} \leq \epsilon l_{i,j} \quad (27)$$

due to the triangular inequality, and hence

$$\langle d_{i,j} \rangle \leq \epsilon \mathcal{L}. \quad (28)$$

Interpreted geometrically, this inequality holds because  $\mathcal{L}$  approximates the average distance of states along geodesics on the recurrence network graph (which can be considered as the geometric backbone of the attractor) in multiples of  $\epsilon$ , while  $\langle d_{i,j} \rangle_{i,j}$  gives the mean distance of states in  $\mathbb{R}^m$  as measured by the norm  $\|\cdot\|$ .

**3.3.3. Network diameter.** By a similar argument as used in equation (27) for the average path length, the diameter

$$\mathcal{D} = \max_{i,j} l_{i,j} \quad (29)$$

of the recurrence network (i.e. the maximum path length) serves as an  $\epsilon$ -upper bound to the estimated diameter

$$\Delta = \max_{i,j} d_{i,j} \quad (30)$$

of the attractor in phase space:

$$\Delta \leq \epsilon \mathcal{D}. \quad (31)$$

**3.3.4. Closeness centrality.** The inverse average shortest path length of vertex  $v$  to all others in the recurrence network is measured by the closeness centrality [77]

$$c_v = \frac{N-1}{\sum_{i=1}^N l_{v,i}}. \quad (32)$$

If there is no path connecting two vertices  $i$  and  $j$ , the maximum shortest path length in the graph,  $N-1$ , is used in the sum by definition. In a recurrence network,  $c_v$  can be geometrically interpreted as measuring the closeness of  $v$  to all other states with respect to the average length (in units of  $\epsilon$ ) of geodesic connections on the recurrence network graph. In other words,  $c_v$  is large if most of the other vertices are reachable in a small number of  $\epsilon$ -jumps from state to state.

From equations (27) and (32), we can see that the inverse closeness  $c_v^{-1}$  is bounded from below by the average phase space distance of vertex (state)  $v$  to all other vertices (states)  $\langle d_{v,i} \rangle_i$  in units of  $\epsilon$  (geometrical closeness), as measured by the norm  $\|\cdot\|$ ,

$$\frac{1}{\epsilon} \frac{N}{N-1} \langle d_{v,i} \rangle_i \leq c_v^{-1}. \quad (33)$$

Put differently, geometrical closeness provides an upper bound for topological closeness,

$$c_v \leq \epsilon \frac{N-1}{N} \langle d_{v,i} \rangle_i^{-1}. \quad (34)$$

**3.3.5. Betweenness centrality.** Let  $\sigma_{i,j}$  be the number of shortest paths between two vertices  $i$  and  $j$ , and  $\sigma_{i,j}(v)$  the number of such paths that include a specific vertex  $v$  of the network. Then, the betweenness centrality of  $v$  is defined as [83]

$$b_v = \sum_{i,j \neq v}^N \frac{\sigma_{i,j}(v)}{\sigma_{i,j}}. \quad (35)$$

In addition to degree and closeness centralities, betweenness centrality yields another possibility to identify especially relevant vertices. Note that unlike the degree centrality  $k_v$ , it is defined locally but depends on global adjacency information.

For general complex networks, the meaning of  $b_v$  can be understood in terms of the importance of individual vertices for the transport of information or matter, assuming that both typically travel through the network on shortest paths. In general, there are  $\sigma_{i,j}$  different shortest paths connecting two vertices  $i$  and  $j$ . We then regard a vertex  $v$  to be an important mediator for the transport on the network, if it is traversed by a large number of all existing shortest paths. According to these conceptual ideas, in equation (35), the contribution of shortest paths is weighted by their respective multiplicity  $\sigma_{i,j}$ , the physical rationale for this normalization being that the total volume of information flow between two vertices, when summed over all shortest paths connecting them, should be the same for all pairs in the network.

For a recurrence network, the notion of information transfer is not useful anymore. However, we can still argue in a geometric way that high betweenness states are typical for regions of sparse phase space density that separate different high-density clusters (referring to the information flow analogy mentioned above, we consider the corresponding vertices as *geometric bottlenecks*). Thus, the occurrence of high betweenness values can be a sign of highly fractionated attractors (on the scale resolved by the considered threshold  $\epsilon$ ). A more detailed discussion of the corresponding implications for some simple model systems will be given in section 4.

**3.3.6. Edge betweenness.** While betweenness centrality refers to vertex properties of a network, an equivalent measure can also be defined based on the number of shortest paths on the network that include a specific edge  $(i, j)$ . We refer to the corresponding property as the edge betweenness  $b_{i,j}$ . Note that although there is a conceptual difference between vertex-related and edge-related betweenness, both quantities are indicators for regions of low phase space density that separate regions with higher density (or, to phrase it differently, of regions of high attractor fractionation) and thus have practically the same dynamical meaning.

#### 4. Examples

In the following, we will show the potentials of the network-theoretic measures discussed in the previous section for recurrence networks obtained from three paradigmatic chaotic model systems.

##### 4.1. Model systems

Basic results for one-dimensional maps have already been described for the logistic map (see [43]) based on numerical calculations and are supplemented by some further computations in the appendix. At this point, we prefer to discuss in some more detail the properties of systems that are defined in somewhat higher dimensions. In particular, we consider the Hénon map

$$x_{i+1} = y_i + 1 - 1.4x_i^2, \quad y_{i+1} = 0.3x_i \quad (36)$$

as an example for a chaotic two-dimensional map, and the Rössler system

$$\frac{d}{dt}(x, y, z) = (-y - z, x + 0.2y, 0.2 + z(x - 5.7)) \quad (37)$$

as well as the Lorenz system

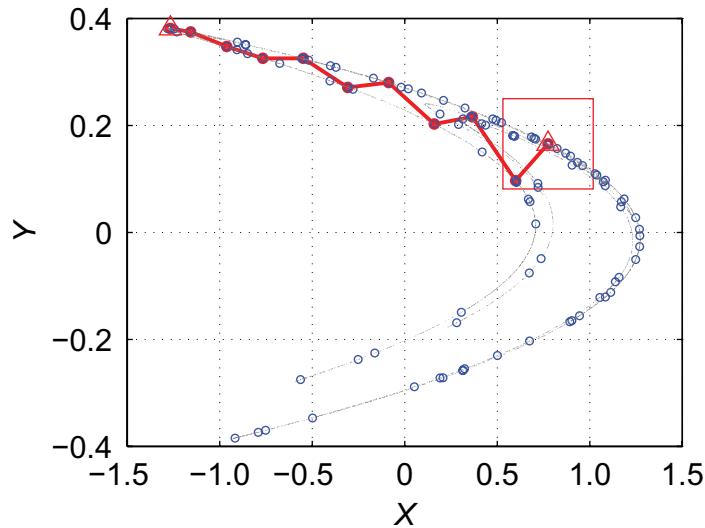
$$\frac{d}{dt}(x, y, z) = \left( 10(y - x), x(28 - z) - y, xy - \frac{8}{3}z \right) \quad (38)$$

as two examples for three-dimensional chaotic oscillators. In all the following considerations, no additional embedding will be used. Note, however, that for the continuous systems, temporal correlations between subsequent observations have been excluded by removing all sojourn points from the recurrence matrix [9, 84].

Figures 3 and 4 show examples of typical trajectories of these three model systems. In addition, the shortest paths between the first and last points of the individual realization are indicated, underlining the deep conceptual differences between the concepts of the trajectory (in phase space) and path (in a recurrence network, see section 3).

##### 4.2. $\epsilon$ -dependence of network properties

In order to evaluate the robustness of the topological properties of recurrence networks, their dependence on the free parameter of the method,  $\epsilon$ , has to be explicitly considered [12]. In the following, we will briefly summarize the main findings for the three model systems.

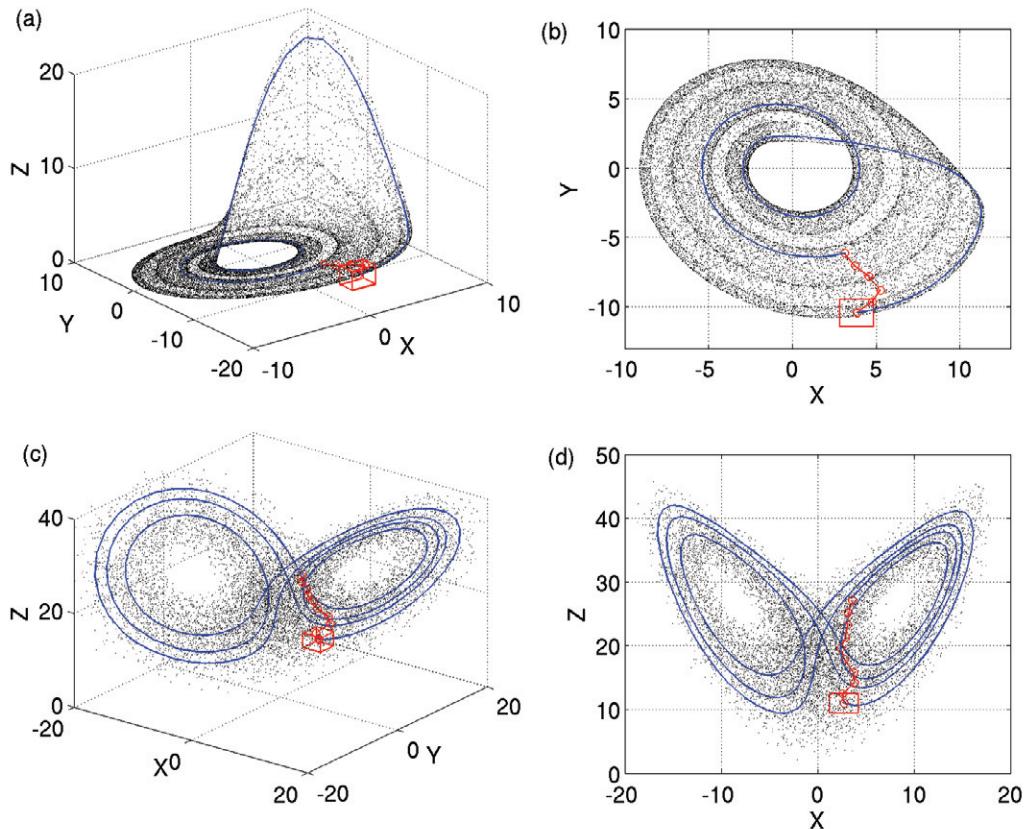


**Figure 3.** One trajectory of the Hénon map (attractor indicated by grey dots) formed by 100 iterations (i.e.  $N = 101$ ), indicated by circles. The initial condition is marked by a square, the size of which corresponds to the neighbourhood threshold  $\epsilon = 0.25$  (maximum norm) considered in the derivation of the corresponding recurrence network. The shortest path (with  $l_{1,101} = 10$ ) between the first and the last point is indicated by a continuous red line.

**4.2.1. Global network properties.** The variations of  $\mathcal{L}$  with the threshold  $\epsilon$  are shown in figure 5 and verify the existence of an inverse relationship of a corresponding lower bound postulated in equation (28).

For the global clustering coefficient, the dependence on  $\epsilon$  is more complicated and depends on the specific properties of the considered system (figure 5). In particular, while for too small  $\epsilon$ , problems may occur, since the recurrence network will generally decompose into different disconnected clusters for a length  $N$  of the considered time series, for intermediate threshold values, an approximately linear increase of  $\mathcal{C}$  with  $\epsilon$  seems to be a common feature of all three examples. Following the discussion of the behaviour of one-dimensional maps in appendix A, we argue that this increase is most likely related to the effect of the attractor boundaries.

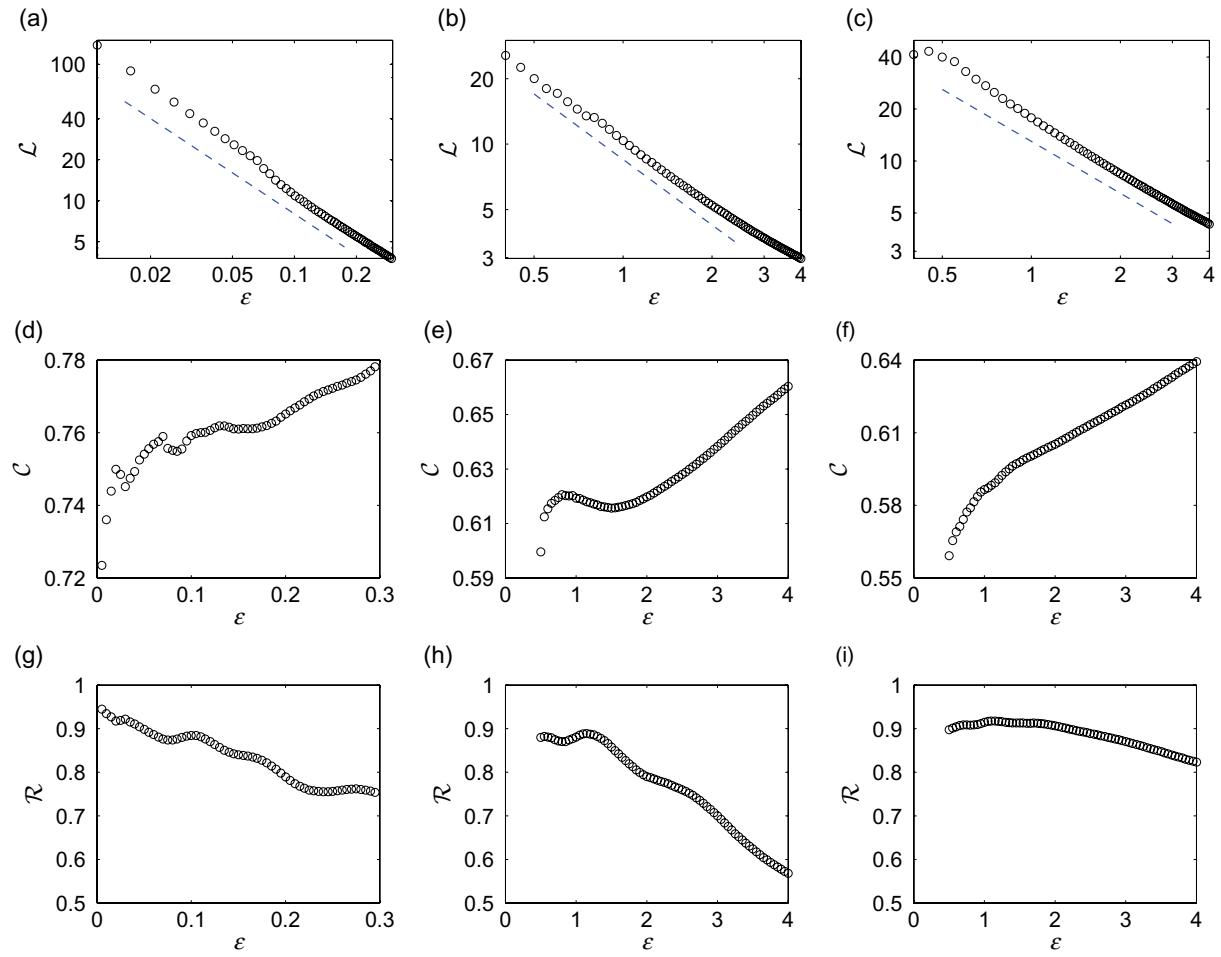
Finally, concerning the assortativity coefficient  $\mathcal{R}$ , we observe that for small  $\epsilon$ , the recurrence networks are highly assortative (e.g.  $\mathcal{R}$  is close to 1). This behaviour can be related to the fact that in the case of small neighbourhoods, these phase space regions are usually characterized by only weak variations of the phase space density, so that neighbouring vertices have a tendency to obey a similar degree. As  $\epsilon$  becomes larger, larger regions of the phase space are covered, where the density may vary much stronger, which implies that the degrees of neighbouring vertices become less similar. Note, however, that since in this case, the mutual overlap of the different neighbourhoods becomes successively larger, there is still a significantly positive correlation between the degrees of neighbouring vertices. We further observe that the decrease of  $\mathcal{R}$  with  $\epsilon$  is interrupted by intermediate increases, which are probably related to some preferred spatial scale of the separation of certain dynamically invariant objects such as UPOs. We will come back to this point in section 4.3.3.



**Figure 4.** (a, b) Chaotic attractor of the Rössler system (dark dots) and realization of one particular non-periodic trajectory (blue line), corresponding to  $T = N\Delta t = 11.4$  time steps. The size of the considered neighbourhood  $\epsilon = 1$  (maximum norm) is indicated by a red square around the initial condition. (Note that the square gives an idealized representation of the considered neighbourhood.) The red line displays the shortest path between the initial condition and the final value on this trajectory ( $l = 5$ ). (c, d) One example trajectory of the Lorenz system with  $T = 5$  time steps, and the resulting shortest path between the initial and the final state ( $\epsilon = 1.5$ ,  $l = 9$ ).

**4.2.2. Vertex properties.** Similar to the global network properties discussed above, the properties of the individual vertices of a recurrence network show a corresponding dependence on the recurrence threshold  $\epsilon$  as well. This observation is rather trivial for the degree centrality, where not only the global mean value (proportional to the recurrence rate  $RR$ ) but also the local value at an individual vertex must be a monotonously increasing function of  $\epsilon$  [9]. We note that for comparing the properties of recurrence networks between different dynamical systems (or for different values of the control parameters of the same system) that generally have attractors with different phase space diameters, it is indeed favourable to fix the recurrence rate  $RR$  instead of the metric distance  $\epsilon$ . For a detailed discussion on this question see, e.g., [85].

Concerning other vertex properties such as  $b_v$  or  $C_v$ , we note that the spatial distribution of the individual values becomes successively smoother as  $\epsilon$  increases, since the small-scale features of the phase space density captured by these measures cannot be resolved with large recurrence thresholds anymore. However, for sufficiently small  $\epsilon$ , we find that the main

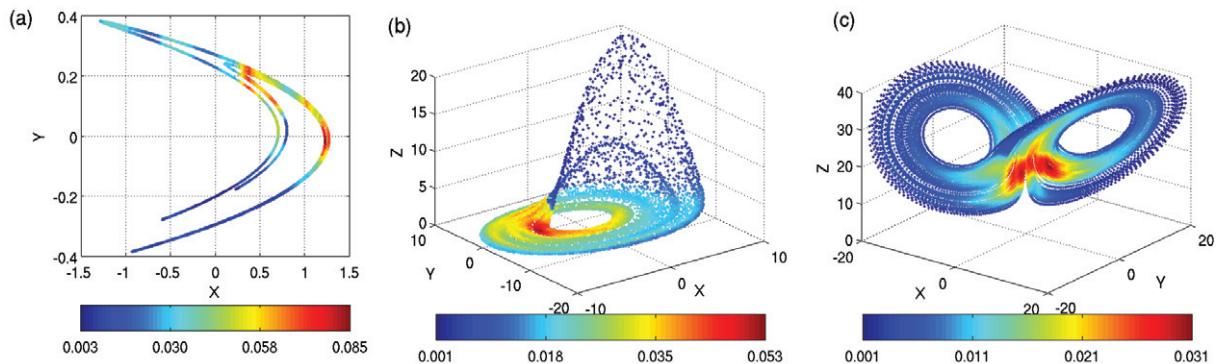


**Figure 5.** Dependence of (a–c) the average path length  $L$ , (d–f), the global clustering coefficient  $C$  and (g–i) the assortativity coefficient  $R$  for the Hénon map, Rössler and Lorenz system (from left to right) using the maximum norm. Dashed lines in the plots of  $L(\epsilon)$  indicate the approximate presence of the theoretically expected  $1/\epsilon$  dependence of the average path length. Note that although a normalized threshold  $\epsilon$  might yield a better comparability of the results for the different systems, we prefer using the absolute values here since the typical normalization factors—either an empirical estimate of the standard deviation of the phase space density (which may be crucially influenced by asymmetric densities) or the attractor diameter (which is itself not *a priori* known in advance) have certain conceptual problems.

structures resolved by the local vertex properties are preserved [12]. In the following, we will operate with different values of  $RR$  for different systems and network-theoretic measures, depending on the computational requirements and the spatial scales of interest.

#### 4.3. Spatial distributions of vertex properties

In the following, we will study the interrelationships between local network properties and structural features of the phase space for the three considered chaotic model systems.

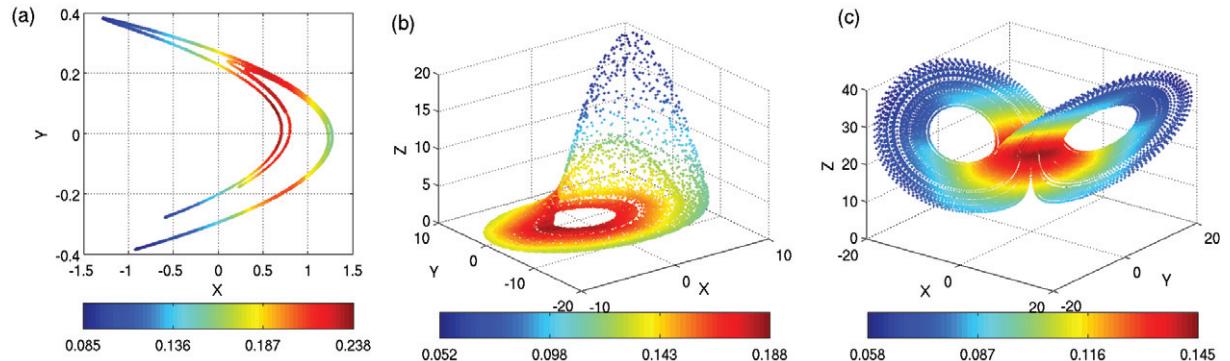


**Figure 6.** Colour-coded representation of the local recurrence rate  $RR_v$  (proportional to the degree centrality  $k_v$ ) in phase space (a) Hénon map ( $N = 10\,000$ ), (b) Rössler ( $N = 10\,000$ ) and (c) Lorenz systems ( $N = 20\,000$ ). The value of  $\epsilon$  for each case is chosen in such a way that the global recurrence rate  $RR = \rho \approx 0.03$ .

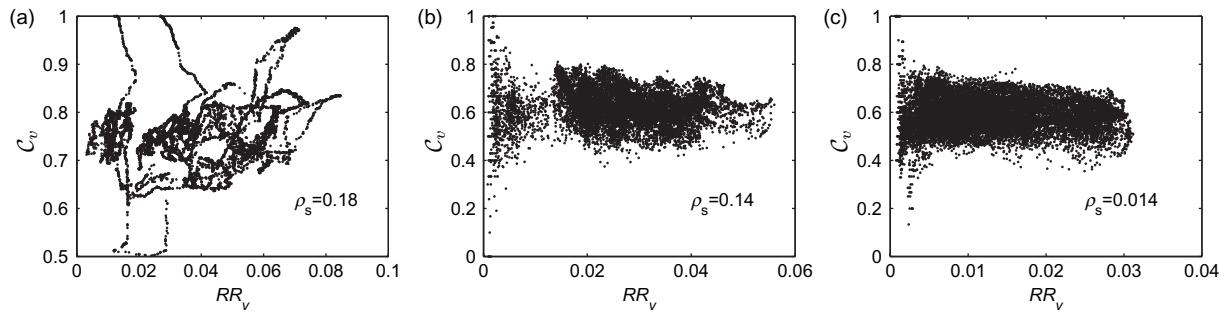
**4.3.1. Degree centrality.** When considering the degree centrality  $k_v$  or, equivalently, the local density  $\rho_v$  for all vertices of the network, a broad range of variability is found (figure 6). In particular, the behaviour follows the expectation that regions with a high phase space density (for example, the merger of the two scrolls of the Lorenz oscillator) also reveal a high density of vertices and, hence, high degree centralities. Note that the calculation of a recurrence plot depends on the parameter  $\epsilon$ , which should be tailored to the considered system under study and the specific questions one wishes to address. Several ‘rules of thumb’ for the choice of the threshold  $\epsilon$  have been advocated in the literature [9, 11]. It has been suggested that the choice of  $\epsilon$  to achieve a fixed recurrence rate  $RR$  is helpful for the estimation of dynamical invariants in many systems [9]. Therefore, this procedure will be adopted here to obtain an overall visualization of the degree centrality  $k_v$  in phase space, with  $RR = \rho \approx 0.03$  (which lies within the typical scaling region of the correlation integral). However, as we will see later, for the local clustering coefficients (section 4.3.3) disclosing local fine structures of the phase space density, it is necessary to choose smaller  $\epsilon$ .

**4.3.2. Closeness centrality.** Figure 7 reflects the spatial distribution of the closeness centrality  $c_v$ . In good agreement with our previous theoretical considerations on the geometric meaning of this measure (section 3.3.4), we find high values of  $c_v$  near the centre of gravity of the attractor in phase space, and low values at phase space regions that have large distances from this centre.

**4.3.3. Clustering coefficient.** There is no distinct relationship between the local clustering coefficient  $C_v$  and the degree centrality  $k_v$  (figure 8), which implies that the spatial patterns of  $C_v$  are not primarily related to variations of the phase space density itself. Instead, we argue that  $C_v$  characterizes higher-order properties related to the heterogeneity of the phase space density in the vicinity of a vertex  $v$ . For example, in a two-dimensional system, an alignment of vertices along a one-dimensional subspace will produce clearly higher clustering coefficients than a homogeneous filling of the neighbourhood, since a larger number of neighbours of a specific vertex are also mutually connected in this case (see figure 9).



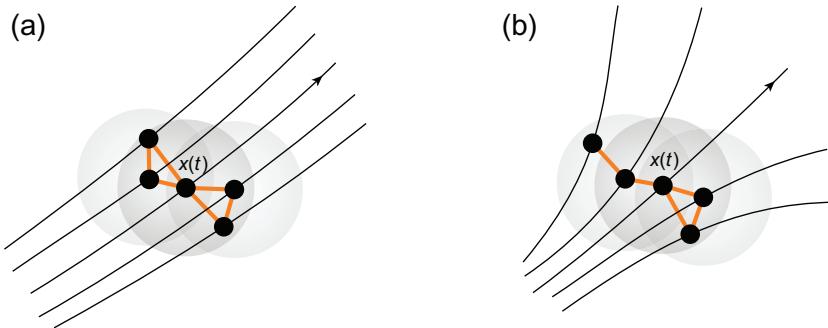
**Figure 7.** Colour-coded representation of the closeness centrality  $c_v$  (equation (32)) in phase space of (a) Hénon map, (b) Rössler system and (c) Lorenz system ( $N$  and  $\epsilon$  are the same as in figure 6).



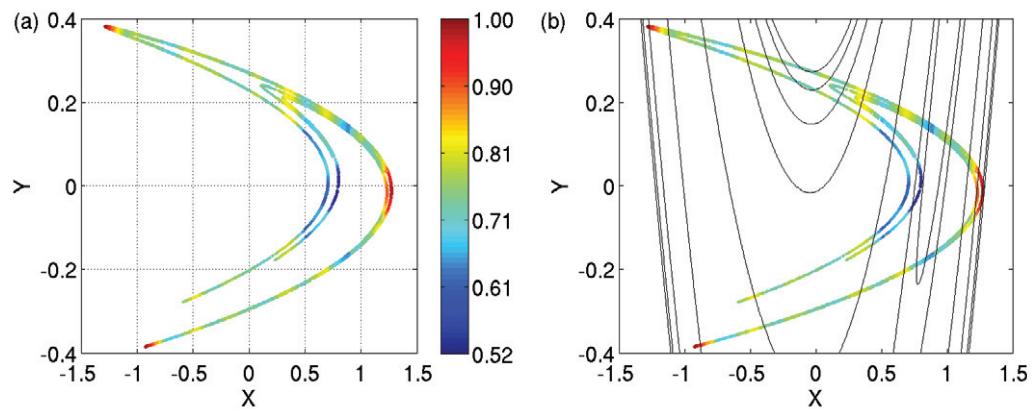
**Figure 8.** Scatter diagrams between the local recurrence rate  $RR_v$  and the local clustering coefficient  $C_v$  for typical realizations of (a) Hénon map, (b) Rössler system and (c) Lorenz system. The inserted values  $\rho_s$  give the corresponding rank-order correlation coefficient (Spearman's rho).

The corresponding behaviour is underlined in figure 10 for the Hénon map, where maximum values of  $C_v = 1$  can be particularly found at the two tips of the attractor. Generalizing these considerations to arbitrary dynamical systems, we note that  $C_v$  for a recurrence network quantifies how randomly the vertices are distributed in a specific part of the phase space. In this sense,  $C_v$  can be considered as an entropy-related quantity. As a more detailed interpretation of this finding, we emphasize that from the theory of spatial random graphs [88] (which can be assumed to yield the lowest possible clustering coefficients among all networks that are embedded in a space with a prescribed dimension  $m$ ), it is known that in the asymptotic limit  $N \rightarrow \infty$  and  $\epsilon \rightarrow 0$ , the possible values of  $C_v$  are bound between a theoretical lower limit and 1. Moreover, this lower bound systematically decreases with increasing  $m$ , which appears reasonable if for a spatial network, we interpret  $C_v$  as an entropy-related quantity. To be more specific, according to Dall and Christensen [88], this decay follows an exponential function for sufficiently large spatial dimensions.

The presence of distinct structures in the spatial profile of the local clustering coefficient is related to the emergence of specific dynamically invariant objects in the considered model systems. In the case of the Hénon map (figure 10), there is a clear tendency that points that

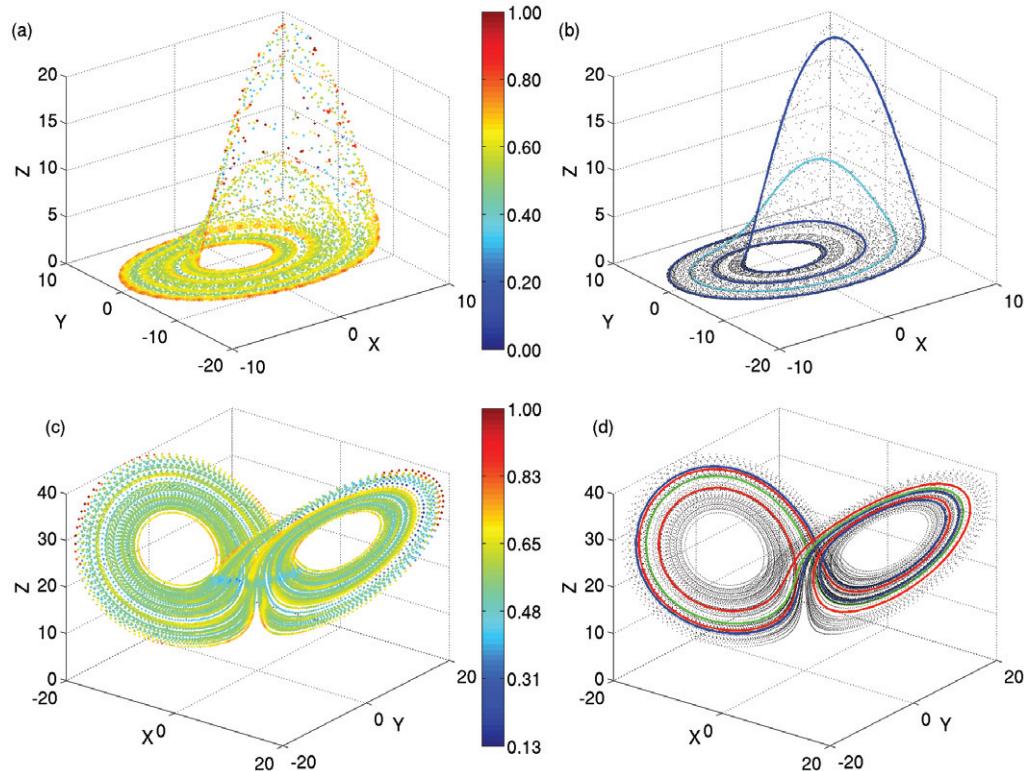


**Figure 9.** Schematic representation of the alignment of trajectories and the resulting recurrence network structures (a) close to a UPO and (b) within a phase space region with larger divergence of neighbouring trajectories and, hence, more homogeneous spatial filling with observational points.



**Figure 10.** (a) Colour-coded representation of the local clustering coefficients  $C_v$  for the Hénon map in its phase space. (b) Relationship between the local clustering coefficients and the stable manifold of the hyperbolic fixed point of the Hénon map. The finite length segment of the stable manifold is calculated by the method described in [86, 87] (with 20 000 iterations).

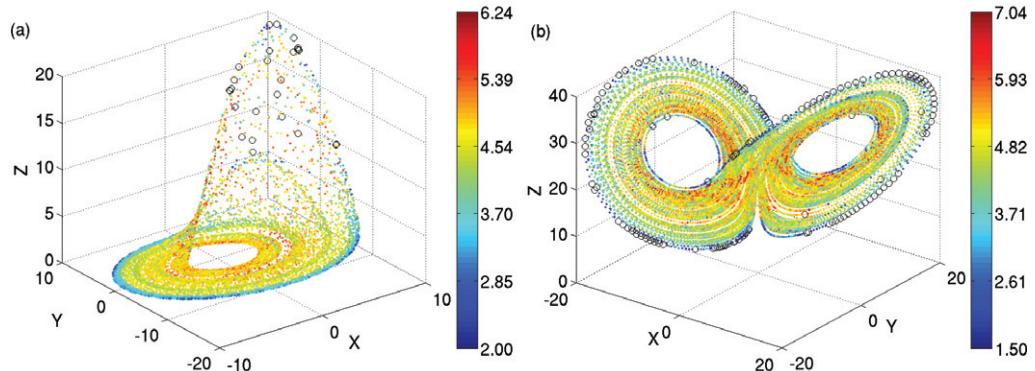
are close to the stable manifold associated with the hyperbolic fixed point of the system have remarkably higher values of  $C_v$ . Note, however, that because of finite size effects, this coincidence cannot be found for all corresponding regions of the phase space. For the two continuous systems (figure 11), points close to the trapping regimes of UPOs have higher clustering coefficients. It is, in some sense, trivial to understand the role of UPOs in forming such regimes of higher clustering coefficients. Whenever a trajectory visits the vicinity of a UPO, it is captured in this neighbourhood for a certain finite time, during which the probability of recurrences is increased. Furthermore, once the trajectory is trapped, the local divergence rate becomes smaller (see figure 9). This smaller local divergence rate is captured by the clustering coefficient in terms of higher-order correlations between neighbours of a vertex. Similar to the finite- $\epsilon$  effect in the Hénon map, the regions with increased clustering coefficients in most cases only coincide with UPOs of lower periods. Therefore, in figure 11, only a few UPOs of low order are shown for comparison. Note that if two UPOs are separated by a distance smaller



**Figure 11.** (a) Colour-coded representation of the local clustering coefficients  $C_v$  for the Rössler system ( $N = 10\,000$ ) in its phase space. (b) Locations of several UPOs with low periods, obtained using a method based on the windows of parallel lines in the corresponding recurrence plot as described in [9]. The chosen value of  $\epsilon$  corresponds to a global recurrence rate of  $RR = 0.01$ . (c, d) The same as (a, b) for the Lorenz system ( $N = 20\,000$ ).

than  $\epsilon$  in phase space, the clustering coefficient is not able to distinguish between these two structures and, hence, shows a broad band with increased values. In the limit of  $\epsilon \rightarrow 0$  (and, hence,  $N \rightarrow \infty$ ), also UPOs of higher periods can be detected by  $C_v$ . Following this argument,  $C_v$  is a useful measure for detecting phase space regions with a high density of low-order UPOs, which supports corresponding considerations in [42], where the influence of a finite value of  $\epsilon$  has, however, not been properly considered [12].

**4.3.4. Betweenness centrality.** Our interpretation of the betweenness centrality in section 3.3.5 implies that  $b_v$  is a rather sensitive measure of the local fragmentation of the attractor and thus gives complementary information, especially on very small scales. Unfortunately, numerical limitations on the calculation of this measure did not allow us to explore the limit of small neighbourhoods ( $\epsilon \rightarrow 0$ ). However, from our computations with somewhat larger thresholds (see figure 12), we can already derive some general statements about the behaviour of betweenness centrality for the considered model systems. Firstly, note that regions close to the outer boundaries of the attractor (in contrast to those in the vicinity of the inner boundaries, e.g. of the Rössler oscillator) are not important for many shortest path connections on the



**Figure 12.** Logarithm of the betweenness centrality  $b_v$  for (a) the Rössler system and (b) the Lorenz system ( $N$  and  $\epsilon$  as in figure 11). Points shown as circles have betweenness values below the lower limit of the displayed colour scale.

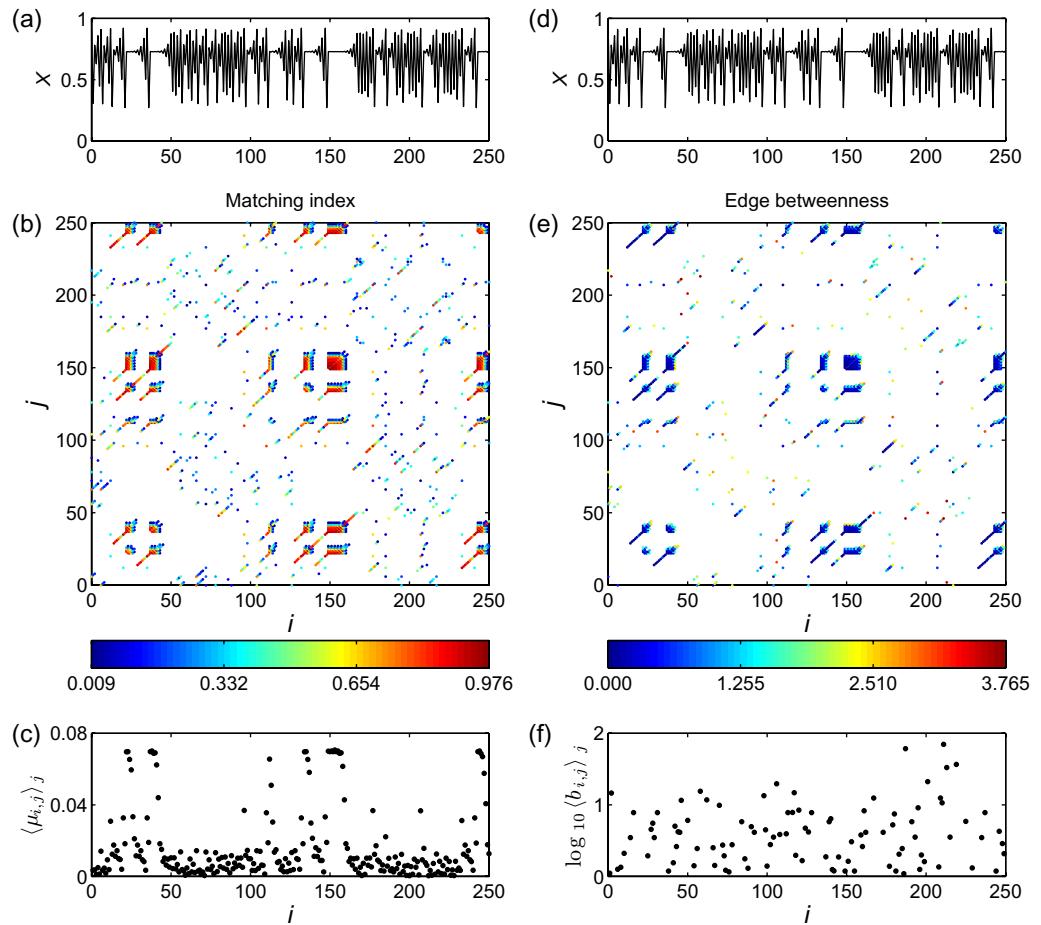
recurrence network. Hence, vertices settled in the corresponding parts of the phase space are characterized by low betweenness values. Secondly, if there are pronounced regions with rather few isolated points in between high-density regions (for example, between two UPOs in the Rössler or Lorenz systems), there is an increasing number of shortest paths crossing these vertices, which leads to higher values of  $b_v$ . In turn, vertices in the vicinity of UPOs (i.e. high-density regions) typically show lower betweenness values. Therefore, betweenness centrality provides a complementary view on the attractor geometry in comparison with the local clustering coefficient  $C_v$  (figure 11). However, we have to emphasize that  $b_v$  alone does not provide a feasible measure for detecting UPOs in dynamical systems [12].

#### 4.4. Spatial distributions of edge properties

Similarly to the local vertex properties, we also illustrate the characteristics of different edges in the recurrence networks. Since the resulting structures are more pronounced than for the three model systems considered so far, figure 13 shows the matching index and edge betweenness for one realization of the logistic map  $x_{i+1} = ax_i(1 - x_i)$  in the intermittent chaotic regime (see [43]). The presence of intermittent dynamics can be clearly seen from the recurrence plots in terms of extended square recurrence patterns, which hence lead to mutually connected vertices of the associated recurrence networks that correspond to subsequent points in time.

Figures 13 and 14 show the complex dependence between phase space distance  $d_{i,j}$ , matching index  $\mu_{i,j}$  and edge betweenness  $b_{i,j}$ . For the matching index, the results are consistent with our theoretical considerations presented in section 3.2.5. In particular, for  $d_{i,j} \rightarrow 0$ , we have  $\mu_{i,j} \rightarrow 1$ , while for  $d_{i,j} \rightarrow 2\epsilon$ ,  $\mu_{i,j} \rightarrow 0$ . Concerning the temporal evolution during the laminar (intermittent) phase, we recognize that at the beginning, there is hardly any change in the state of the system; hence,  $d_{i,j}$  is very small for subsequent points in time (vertices of the recurrence network), which relates to large values of the matching index near 1. As the laminar phases are close to their termination, chaotic variations emerge and rise in amplitude, which leads to a subsequent increase of  $d_{i,j}$  and, hence, decrease of  $\mu_{i,j}$ .

Concerning the edge betweenness  $b_{i,j}$  (the spatial pattern of which is very similar to that of the vertex-based betweenness centrality  $b_v$  due to the spatial proximity of edge and corresponding vertices), the overall behaviour is opposite to that of  $\mu_{i,j}$ . During laminar phases,

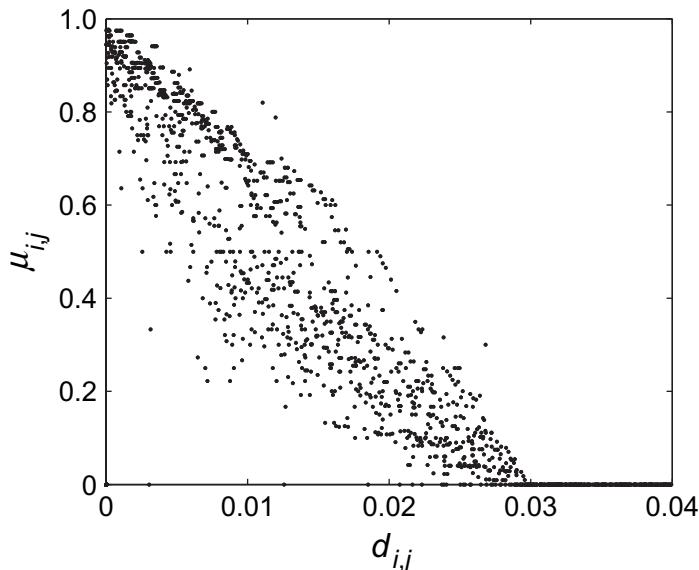


**Figure 13.** (a, d) One example trajectory of the logistic map at  $a = 3.679$  (chaotic regime with intermittency),  $\epsilon = 0.015\sigma_x$  ( $\sigma_x$  being the empirical standard deviation of the time series),  $N = 1000$  (only a part of the trajectory with 250 points is shown). (b) Matching index  $\mu_{i,j}$  between all pairs of vertices (colour-coded). (c) Average matching index  $\langle \mu_{i,j} \rangle_j$  of all vertices of the considered recurrence network. (e, f) As in (b, c) for the logarithm of the edge betweenness  $b_{i,j}$ .

we find that since all states are very close to each other, possible shortest connections may alternatively pass through a variety of different edges, leading to low values of the edge betweenness. Close to the termination, there is in turn an increase of this measure. However, the most interesting feature of the edge betweenness is presented by isolated edges with very high values of  $b_{i,j}$ , which correspond to rarely visited phase space regions between intervals of higher phase space density. More specifically, the average edge betweenness of vertices in such low-density regions can exceed that of high-density regions by orders of magnitude (figure 13).

## 5. Conclusions

This paper has reconsidered the analysis of time series from complex systems by means of complex network theory. We have argued that most existing approaches for such an analysis suffer from certain methodological limitations or a lack of generality in their applicability.



**Figure 14.** Scatter diagram of the matching index  $\mu_{i,j}$  against the phase space distance  $d_{i,j}$  for the logistic map at  $a = 3.679$  (parameters as in figure 13).

As an appealing solution, we have suggested recurrence networks as a unifying framework for studying time series as complex networks, which is based on a novel approach for the quantitative assessment of recurrence plots in terms of complex network measures. As we have argued, this specific approach is applicable to univariate as well as multivariate time series with or without embedding. In addition, recurrence networks can be applied for studying time series with non-equidistant timescales as well as temporal variations in non-stationary data (by using sliding windows in time) and allow the construction of simple significance tests with respect to the associated network-theoretic measures [43].

Our main achievement is to have provided a thorough reinterpretation of a variety of statistical measures from network theory computed for recurrence networks in terms of phase space properties of dynamical systems. Since all time ordering information is lost in this approach, all complex network characteristics are dynamically invariant, i.e. they are only sensitive to certain properties of the invariant density of the considered dynamical system. From this invariance, it follows that specific measures such as the local clustering coefficient might be used for detecting dynamically invariant objects like UPOs or chaotic saddles. However, this feature also implies that the proposed method cannot be used to distinguish between deterministic (chaotic) and stochastic systems, which is exemplified by our comparison between the Bernoulli map and uniform noise in appendix A. As a possible means to overcome this potential point of criticism, we emphasize that an additional embedding should change the properties of deterministic systems in a different way than for a stochastic system; hence, studying complex network properties dependent on the embedding dimension might help to solve this interpretation problem. We will further elaborate this idea in future research.

Using widespread statistical characteristics of complex networks such as the ‘trinity’ of centrality measures (degree, closeness and betweenness) and the clustering coefficient, we were able to provide a detailed interpretation of the corresponding results for recurrence networks in

terms of higher-order properties of the phase space density of a dynamical system. In particular, degree centrality relates to the local density, closeness centrality to the average geometrical proximity of an observation to all other observations, and betweenness centrality to the local fragmentation of points in phase space. For the local clustering coefficient  $C_v$ , the analysis of different model systems has revealed that the resulting values are related to the homogeneity of the spatial filling, which becomes important close to the attractor boundaries and certain dynamically invariant objects such as invariant manifolds or UPOs. The resulting local network properties are qualitatively robust under changes of the recurrence threshold  $\epsilon$  given that the associated recurrence rate  $RR$  remains sufficiently small [12]. Additionally, we have presented a rigorous analytical treatment of the local and global clustering coefficients of the recurrence networks of one-dimensional chaotic maps, which perfectly matches our numerical results (see appendix A). A possible further relationship between the local clustering coefficient of a recurrence network and other nonlinear characteristics of the underlying dynamical system such as the local Lyapunov exponent or the point-wise dimension remains a topic for future work.

With respect to existing recurrence plot-based methods of time series analysis, e.g. RQA, we would like to emphasize that our approach yields a complementary view on the phase space properties of the underlying dynamical system. In particular, we note that nearly all of the considered network-theoretical measures have no direct equivalent in traditional RQA, and vice versa. Hence, there may be situations when either one of the two frameworks (i.e. RQA or recurrence networks) provides better results than the other. Examples include the detection of dynamical transitions between periodic and chaotic behaviour in maps [43], continuous dynamical systems [85] or the analysis of protein folding [89, 90]. In [85], an additional comparison with the standard approach based on numerical estimates of the maximum Lyapunov exponents has also been presented. In turn, there have been some recent approaches to applying methods of time series analysis to general complex networks (e.g. [91, 92]). We would like to underline that due to the duality between the recurrence matrix of a time series and the adjacency matrix of the associated recurrence network, RQA might be another promising candidate for this purpose (as long as we restrict ourselves to measures that are invariant under random permutations of observations), yielding interesting complementary insights on complex networks in a variety of different situations. A more detailed investigation of the potential of a corresponding approach will be the subject of our future research.

Finally, we would like to point out that although our studies in this work have only been supported by results for low-dimensional systems, the proposed methods can also be applied to higher-dimensional systems, e.g. for detecting dynamically relevant phenomena like coherent structures or laminar phases. However, in such cases, we expect some technical problems related to computational demands (associated with certain basic requirements to the number of observations) for calculating different network properties (especially when additional embedding is necessary), the removal of sojourn points, and the visualization of the resulting vertex properties, which could present serious practical challenges for applications of recurrence networks. Nevertheless, an application of recurrence networks to carefully chosen low-dimensional subsets of observables from a high-dimensional system can still yield insightful results, as recently exemplified by a study of a palaeoclimate time series [43].

## Acknowledgments

This work was partly supported by the German Research Foundation (DFG project no. He 2789/8-2, SFB 555 (project C1) and DFG graduate school GK 1364), the Japanese Ministry for Science and Education, the Max Planck Society, the Leibniz Society (project ECONS) and the BMBF (research cluster PROGRESS). Most of the complex network measures have been calculated using the `igraph` package [93]. We acknowledge the helpful comments of two anonymous referees and P Grassberger on a previous version of this paper, as well as the discussions with M Small and K Padberg-Gehle.

## Appendix A. Clustering coefficient of recurrence networks for one-dimensional maps

In order to compute the clustering coefficient of the recurrence network of a dynamical system, certain system-specific integrals have to be solved. For the case of one-dimensional maps that are defined on the interval  $[0, 1]$ , these integrals can be explicitly expressed and eventually also evaluated analytically.

### A.1. General treatment

The computation of the local clustering coefficient involves the evaluation of certain integrals to estimate the required probability terms discussed in section 3. For a simplified theoretical analysis, let us assume that for a given observation  $x_v$  at time  $t_v$ , the probabilities that two other observations  $x_i$  and  $x_j$  made at *arbitrary* times  $t_i$  and  $t_j$  are closer to  $x_v$  than  $\epsilon$  are mutually independent. We have to emphasize that in the presence of temporal correlations, this assumption is violated. As long as these correlations decay sufficiently fast, we argue, however, that deviations from statistical independence do not lead to significant errors in the following theoretical considerations, which is supported by the excellent agreement between approximate analytical theory and numerical results (see figures A.2 and A.3)<sup>12</sup>.

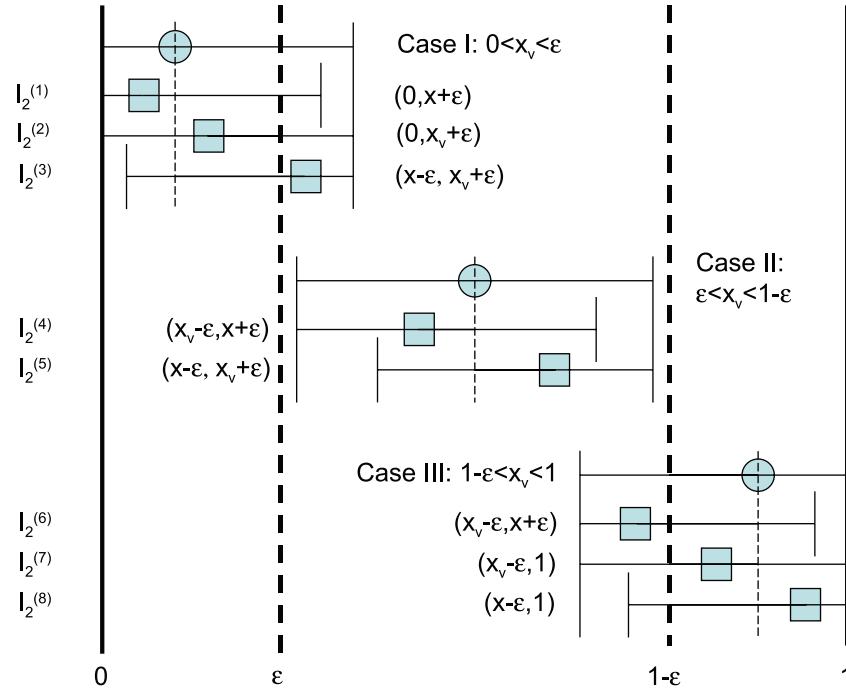
Making use of the independence assumption discussed above, we may write for the condition in the denominator of equation (17):

$$\begin{aligned} P(A_{v,i} = 1, A_{v,j} = 1) &= P(||x_v - x_i|| < \epsilon, ||x_v - x_j|| < \epsilon) \\ &\approx P(||x_v - x_i|| < \epsilon)P(||x_v - x_j|| < \epsilon) \\ &= P(||x_v - x_i|| < \epsilon)^2 = \left[ \int_{x_v-\epsilon}^{x_v+\epsilon} dx p(x) \right]^2. \end{aligned} \quad (\text{A.1})$$

In contrast, we cannot justify a similar assumption in evaluating the corresponding numerator  $P(A_{ij} = 1, A_{vi} = 1, A_{vj} = 1)$ , which therefore requires a subtle choice of the integration boundaries to assure a correct treatment of the three-point relationships

$$P(A_{i,j} = 1, A_{v,i} = 1, A_{v,j} = 1) = \int_{x_v-\epsilon}^{x_v+\epsilon} dx p(x) \int_{\max(x-\epsilon, x_v-\epsilon)}^{\min(x+\epsilon, x_v+\epsilon)} dy p(y). \quad (\text{A.2})$$

<sup>12</sup> The assumption of statistical independence is consistent with the fact that we do not make use of any time information in the construction of recurrence networks unless additional embedding is involved. In contrast, our approach relies exclusively on spatial information. This implies that independent of a specific temporal sampling (which can therefore be chosen to exclude temporal correlations), feasible results may be obtained as long as the invariant density of the system under study is sufficiently well approximated by the state density of the sample time series.



**Figure A.1.** Schematic representation of the integration boundaries entering the three-point relationships in equation (A.2) for  $\epsilon \leq 0.5$ .

As shown in figure A.1, the correct integration is not trivial. In particular, if we use the abbreviations

$$I_1(a, b) = \int_a^b dx p(x), \quad (\text{A.3})$$

$$I_2(a, b; c, d) = \int_a^b dx \left[ p(x) \int_c^d dy p(y) \right], \quad (\text{A.4})$$

we obtain the following expressions for the local clustering coefficient  $C_v = C(x_v)$  if  $\epsilon \leq 0.5$ :  
 $0 \leq x_v \leq \epsilon$ :

$$\begin{aligned} C_v &= \frac{I_2(0, x_v; 0, x_v + \epsilon) + I_2(x_v, \epsilon; 0, x_v + \epsilon) + I_2(\epsilon, x_v + \epsilon; x_v - \epsilon, x_v + \epsilon)}{I_1(0, x_v + \epsilon)^2} \\ &= \frac{I_2^{(1)} + I_2^{(2)} + I_2^{(3)}}{I_1(0, x_v + \epsilon)^2}, \end{aligned} \quad (\text{A.5})$$

$\epsilon \leq x_v \leq 1 - \epsilon$ :

$$\begin{aligned} C_v &= \frac{I_2(x_v - \epsilon, x_v; x_v - \epsilon, x_v + \epsilon) + I_2(x_v, x_v + \epsilon; x_v - \epsilon, x_v + \epsilon)}{I_1(x_v - \epsilon, x_v + \epsilon)^2} \\ &= \frac{I_2^{(4)} + I_2^{(5)}}{I_1(x_v - \epsilon, x_v + \epsilon)^2}, \end{aligned} \quad (\text{A.6})$$

$1 - \epsilon \leqslant x_v \leqslant 1$ :

$$\begin{aligned} \mathcal{C}_v &= \frac{I_2(x_v - \epsilon, 1 - \epsilon; x_v - \epsilon, x + \epsilon) + I_2(1 - \epsilon, x_v; x_v - \epsilon, 1) + I_2(x_v, 1; x - \epsilon, 1)}{I_1(x_v - \epsilon, 1)^2} \\ &= \frac{I_2^{(6)} + I_2^{(7)} + I_2^{(8)}}{I_1(x_v - \epsilon, 1)^2}. \end{aligned} \quad (\text{A.7})$$

Understanding the global clustering coefficient as the expectation value of the local one (taken over the whole possible range in  $x$ ), we use the following expression for deriving its value:

$$\mathcal{C}(\epsilon) = \int_0^1 dx_v p(x_v) \mathcal{C}_v(x_v; \epsilon) = \int_0^1 dx_v p(x_v) \frac{P(A_{ij} = 1, A_{vi} = 1, A_{vj} = 1)}{P(A_{vi} = 1, A_{vj} = 1)}. \quad (\text{A.8})$$

### A.2. The Bernoulli map

Among all nonlinear maps defined on the unit interval  $[0, 1]$ , the Bernoulli map  $x_{i+1} = 2x_i \bmod 1$  has the simplest possible invariant density  $p(x) \equiv 1$ .<sup>13</sup> This allows an easy evaluation of the integrals for an analytic computation of the local clustering coefficient, being aware of the restricted integration range. As a result, we find:

$$\mathcal{C}_v(x_v; \epsilon) = \begin{cases} 1 - \left(\frac{x}{x + \epsilon}\right)^2, & 0 \leqslant x \leqslant \epsilon, \\ \frac{3}{4}, & \epsilon \leqslant x \leqslant 1 - \epsilon, \\ 1 - \left(\frac{1-x}{1-x+\epsilon}\right)^2, & 1 - \epsilon \leqslant x \leqslant 1. \end{cases} \quad (\text{A.9})$$

In particular, for  $\epsilon \rightarrow 0$  (and  $N \rightarrow \infty$ ), we have  $\mathcal{C}_v \rightarrow \frac{3}{4} \forall x \in [0, 1]$ , which corresponds to the value of random geometric graphs in one dimension [88]. Hence, one may speculate about this value being a universal limit for the recurrence networks of one-dimensional chaotic maps. Moreover, for  $x \rightarrow 0$  and  $x \rightarrow 1$ , we have  $\mathcal{C}_v \rightarrow 1$  independent of  $\epsilon$ . This behaviour is consistent with our previous observations concerning the effect of sharp attractor boundaries on the local clustering coefficient, for example, in the case of the tips of the Hénon attractor (section 4.3.3).

For the global clustering coefficient, the computation of equation (A.8) leads to

$$\mathcal{C}(\epsilon) = \frac{3}{4} + \epsilon \left( 4 \ln 2 - \frac{5}{2} \right), \quad (\text{A.10})$$

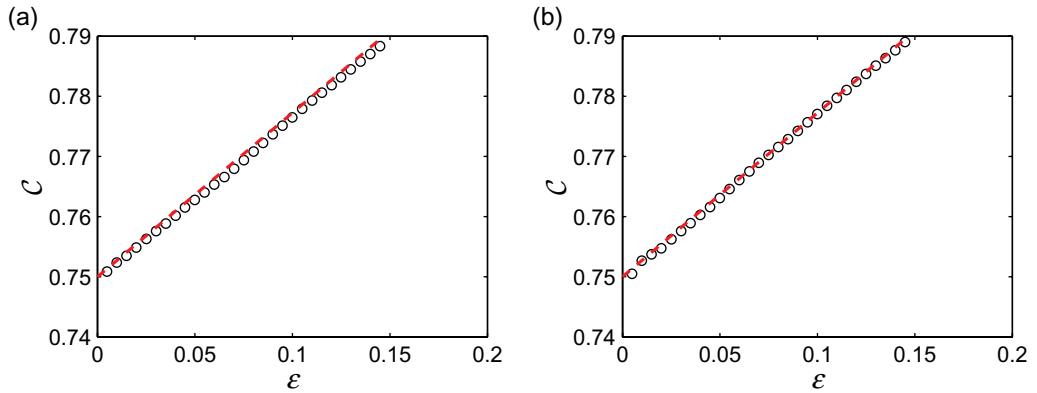
which is in excellent agreement with numerical results (see figure A.2). Hence, we argue that deviations from the theoretical value  $\frac{3}{4}$  occur exclusively due to boundary effects, and can even lead to  $\mathcal{C} = 1$  for very large thresholds  $\epsilon$ .

### A.3. Logistic map for $a = 4$

The logistic map at  $a = 4$  is known to have the invariant density

$$p(x) = \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}}. \quad (\text{A.11})$$

<sup>13</sup> Note that from the recurrence network properties, this map cannot be distinguished from a stochastic process with a uniform density in the same interval (see figure A.2).



**Figure A.2.** Dependence of the global clustering coefficient  $C$  on the threshold  $\epsilon$  for (a) the Bernoulli map and (b) uniformly distributed noise with values in  $[0, 1]$ . Note that due to the same distribution of the data, i.e. the same density in the one-dimensional phase space, both empirical curves (circles) are equal and match the theoretical expectations (dashed red lines). However, note that already a two-dimensional embedding would lead to significant differences between both systems (not shown here).

With this, we find the following expressions:

$$P(A_{v,i} = 1, A_{v,j} = 1) = \begin{cases} \frac{1}{2} - \frac{1}{\pi} \arcsin(1 - 2x_v - 2\epsilon), & 0 \leq x_v \leq \epsilon, \\ \frac{1}{\pi} \{\arcsin(1 - 2x_v + 2\epsilon) - \arcsin(1 - 2x_v - 2\epsilon)\}, & \epsilon \leq x_v \leq 1 - \epsilon, \\ \frac{1}{2} + \frac{1}{\pi} \arcsin(1 - 2x_v + 2\epsilon), & 1 - \epsilon \leq x_v \leq 1 \end{cases} \quad (\text{A.12})$$

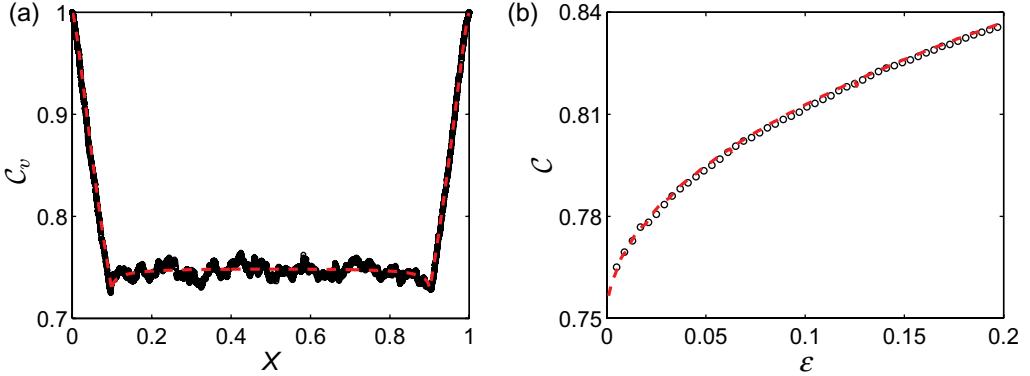
and

$$I_2^{(1)} = \frac{1}{4} - \frac{1}{2\pi} \arcsin(1 - 2x_v) - \frac{1}{\pi^2} \int_0^{x_v} dx \frac{\arcsin(1 - 2x - 2\epsilon)}{\sqrt{x(1-x)}}, \quad (\text{A.13})$$

$$I_2^{(2)} = \left( \frac{1}{2\pi} - \frac{1}{\pi^2} \arcsin(1 - 2x_v - 2\epsilon) \right) (\arcsin(1 - 2x_v) - \arcsin(1 - 2\epsilon)), \quad (\text{A.14})$$

$$\begin{aligned} I_2^{(3)} = & \frac{1}{\pi^2} \arcsin(1 - 2x_v - 2\epsilon) (\arcsin(1 - 2x_v - 2\epsilon) - \arcsin(1 - 2\epsilon)) \\ & + \frac{1}{\pi^2} \int_\epsilon^{x_v+\epsilon} dx \frac{\arcsin(1 - 2x + 2\epsilon)}{\sqrt{x(1-x)}}, \end{aligned} \quad (\text{A.15})$$

$$\begin{aligned} I_2^{(4)} = & \frac{1}{\pi^2} \arcsin(1 - 2x_v + 2\epsilon) (\arcsin(1 - 2x_v + 2\epsilon) - \arcsin(1 - 2x_v)) \\ & - \frac{1}{\pi^2} \int_{x_v-\epsilon}^{x_v} dx \frac{\arcsin(1 - 2x - 2\epsilon)}{\sqrt{x(1-x)}}, \end{aligned} \quad (\text{A.16})$$



**Figure A.3.** (a) Analytical (dashed red line) and simulation results (circles) on the dependence of the local clustering coefficient  $C_v$  on the spatial position  $x$  for the logistic map at  $a = 4$  ( $N = 5000$ ,  $\epsilon = 0.1$ ). (b) As in (a) for the dependence of the global clustering coefficient  $C$  on  $\epsilon$ .

$$\begin{aligned} I_2^{(5)} &= \frac{1}{\pi^2} \arcsin(1 - 2x_v - 2\epsilon) (\arcsin(1 - 2x_v - 2\epsilon) - \arcsin(1 - 2x_v)) \\ &\quad + \frac{1}{\pi^2} \int_{x_v}^{x_v+\epsilon} dx \frac{\arcsin(1 - 2x + 2\epsilon)}{\sqrt{x(1-x)}}, \end{aligned} \quad (\text{A.17})$$

$$\begin{aligned} I_2^{(6)} &= \frac{1}{\pi^2} \arcsin(1 - 2x_v + 2\epsilon) (\arcsin(1 - 2x_v + 2\epsilon) + \arcsin(1 - 2\epsilon)) \\ &\quad - \frac{1}{\pi^2} \int_{x_v-\epsilon}^{1-\epsilon} \frac{\arcsin(1 - 2x - 2\epsilon)}{\sqrt{x(1-x)}}, \end{aligned} \quad (\text{A.18})$$

$$I_2^{(7)} = - \left( \frac{1}{2\pi} + \frac{1}{\pi^2} \arcsin(1 - 2x_v + 2\epsilon) \right) (\arcsin(1 - 2x_v) + \arcsin(1 - 2\epsilon)), \quad (\text{A.19})$$

$$I_2^{(8)} = \frac{1}{4} + \frac{1}{2\pi} \arcsin(1 - 2x_v) + \frac{1}{\pi^2} \int_{x_v}^1 dx \frac{\arcsin(1 - 2x + 2\epsilon)}{\sqrt{x(1-x)}}. \quad (\text{A.20})$$

Note that the integrals  $(I_2^{(1)}, I_2^{(2)}, I_2^{(3)})$  and  $(I_2^{(8)}, I_2^{(7)}, I_2^{(6)})$  can be transformed into each other by the transformation  $x \rightarrow 1 - x$ , and that  $I_2^{(4)} + I_2^{(5)}$  is invariant under the same transformation, reflecting the corresponding symmetry of the phase space density  $p(x)$ .

Since the remaining integrals can only be solved numerically, we are not able to give an explicit equation for the functional dependence of the clustering coefficient on both  $x$  and  $\epsilon$ . However, comparing the numerical solution of our analytical results with the clustering coefficient of the recurrence network for one realization of the logistic map at  $a = 4$ , we find (apart from remaining fluctuations due to the finite length of the considered time series) again excellent agreement (see figure A.3). In particular, for the attractor boundaries at  $x = 0$  and  $x = 1$ , we have  $C_v = 1$  independent of  $\epsilon$  as for the Bernoulli map. Moreover, we observe almost uniform values of the local clustering coefficient close to  $\frac{3}{4}$  within the interval  $[\epsilon, 1 - \epsilon]$ , and a

systematic tendency towards larger values close to the boundaries. Since the measure of the latter intervals systematically increases with increasing  $\epsilon$ , these boundary effects are again responsible for the systematic increase in the global clustering coefficient  $C$  as  $\epsilon$  becomes larger. In turn, the point-wise convergence of  $C_v \rightarrow \frac{3}{4}$  for  $\epsilon \rightarrow 0$  underlines the supposed universality of this value for chaotic one-dimensional maps in agreement with corresponding theoretical predictions for spatial random graphs [88].

## References

- [1] Abarbanel H D I 1996 *Analysis of Observed Chaotic Data* (New York: Springer)
- [2] Kantz H and Schreiber T 1997 *Nonlinear Time Series Analysis* (Cambridge: Cambridge University Press)
- [3] Diks C 1999 *Nonlinear Time Series Analysis—Methods and Applications* (Singapore: World Scientific)
- [4] Sprott J C 2003 *Chaos and Time Series Analysis* (Oxford: Oxford University Press)
- [5] Donner R V and Barbosa S M (ed) 2008 *Nonlinear Time Series Analysis in the Geosciences—Applications in Climatology, Geodynamics and Solar-Terrestrial Physics* (Berlin: Springer)
- [6] Poincaré H 1890 Sur la problème des trois corps et les équations de la dynamique *Acta Math.* **13** A3–A270
- [7] Saussol B, Troubetzkoy S and Vaienti S 2002 Recurrence, dimensions and Lyapunov exponents *J. Stat. Phys.* **106** 623–34
- [8] Eckmann J-P, Oliffson Kamphorst S and Ruelle D 1987 Recurrence plots of dynamical systems *Europhys. Lett.* **5** 973–7
- [9] Marwan N, Romano M C, Thiel M and Kurths J 2007 Recurrence plots for the analysis of complex systems *Phys. Rep.* **438** 237–329
- [10] Takens F 1981 Detecting strange attractors in turbulence *Dynamical Systems and Turbulence (Lecture Notes in Mathematics vol 898)* ed D Rand and L-S Young (Berlin: Springer) pp 366–81
- [11] Schinkel S, Dimigen O and Marwan N 2008 Selection of recurrence threshold for signal detection *Eur. Phys. J. Spec. Top.* **164** 45–53
- [12] Donner R V, Zou Y, Donges J F, Marwan N and Kurths J 2010 Ambiguities in recurrence-based complex network representations of time series *Phys. Rev. E* **81** 015101
- [13] Zbilut J P and Webber C L Jr 1992 Embeddings and delays as derived from quantification of recurrence plots *Phys. Lett. A* **171** 199–203
- [14] Webber C L Jr and Zbilut J P 1994 Dynamical assessment of physiological systems and states using recurrence plot strategies *J. Appl. Physiol.* **76** 965–73
- [15] Marwan N, Wessel N and Kurths J 2002 Recurrence plot based measures of complexity and its applications to heart rate variability data *Phys. Rev. E* **66** 026702
- [16] Marwan N 2008 A historical review of recurrence plots *Eur. Phys. J. Spec. Top.* **164** 3–12
- [17] Thiel M, Romano M C and Kurths J 2006 Spurious structures in recurrence plots induced by embedding *Nonlinear Dyn.* **44** 299–305
- [18] Faure P and Korn H 1998 A new method to estimate the Kolmogorov entropy from recurrence plots: its application to neuronal signals *Physica D* **122** 265–79
- [19] Thiel M, Romano M C, Read P L and Kurths J 2004 Estimation of dynamical invariants without embedding by recurrence plots *Chaos* **14** 234–43
- [20] Thiel M, Romano M C and Kurths J 2004 How much information is contained in a recurrence plot? *Phys. Lett. A* **330** 343–9
- [21] Robinson G and Thiel M 2007 Recurrences determine the dynamics *Chaos* **19** 023104
- [22] Albert R and Barabási A-L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47–97
- [23] Newman M J E 2003 The structure and function of complex networks *SIAM Rev.* **45** 167–256
- [24] da Costa L F, Rodrigues F A, Travieso G and Villas Boas P R 2007 Characterization of complex networks: a survey of measurements *Adv. Phys.* **56** 167–242

- [25] Crucitti P, Latora V and Marchiori M 2004 A topological analysis of the Italian electric power grid *Physica A* **338** 92–7
- [26] Zhou C S, Zemanová L, Zamora G, Hilgetag C C and Kurths J 2006 Hierarchical organization unveiled by functional connectivity in complex brain networks *Phys. Rev. Lett.* **97** 238103
- [27] Dunne J A, Williams R J and Martinez N D 2002 Food-web structure and network theory: the role of connectance and size *Proc. Natl Acad. Sci. USA* **99** 12917–22
- [28] Donges J F, Zou Y, Marwan N and Kurths J 2009 Complex networks in climate dynamics. Comparing linear and nonlinear network construction methods *Eur. Phys. J. Spec. Top.* **174** 157–79
- [29] Donges J F, Zou Y, Marwan N and Kurths J 2009 The backbone of the climate network *Europhys. Lett.* **87** 48007
- [30] Wang X F 2002 Complex networks: topology, dynamics and synchronization *Int. J. Bifurcation Chaos* **12** 885–916
- [31] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D-U 2006 Complex networks: structure and dynamics *Phys. Rep.* **426** 175–308
- [32] Arenas A, Diaz-Guilera A, Kurths J, Moreno Y and Zhou C 2008 Synchronization in complex networks *Phys. Rep.* **469** 93–153
- [33] Zhang J and Small M 2006 Complex network from pseudoperiodic time series: topology versus dynamics *Phys. Rev. Lett.* **96** 238701
- [34] Zhang J, Sun J, Luo X, Zhang K, Nakamura T and Small M 2008 Characterizing pseudoperiodic time series through the complex network approach *Physica D* **237** 2856–65
- [35] Xu X, Zhang J and Small M 2008 Superfamily phenomena and motifs of networks induced from time series *Proc. Natl Acad. Sci. USA* **105** 19601–5
- [36] Shimada Y, Kimura T and Ikeguchi T 2008 Analysis of chaotic dynamics using measures of the complex network theory *ICANN 2008, Part I (Lecture Notes in Computer Science* vol 5163) ed V Kurková *et al* (Berlin: Springer) pp 61–70
- [37] Yang Y and Yang H 2008 Complex network-based time series analysis *Physica A* **387** 1381–6
- [38] Lacasa L, Luque B, Ballesteros F, Luque J and Nuño J C 2008 From time series to complex networks: the visibility graph *Proc. Natl Acad. Sci. USA* **105** 4972–5
- [39] Wu J, Sun H and Gao Z 2008 Mapping to complex networks from chaos time series in the car following model *Proc. 6th Int. Congress on Traffic and Transportation Studies (ASCE, Reston)* pp 397–407
- [40] Small M, Zhang J and Xu X 2009 Transforming time series into complex networks *Complex 2009, Part II (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* vol 5) ed J Zhou (Berlin: Springer) pp 2078–89
- [41] Donner R 2009 Complex network approach to geophysical time series analysis *Geophys. Res. Abstr.* **11** EGU2009-11157
- [42] Gao Z and Jin N 2009 Flow-pattern identification and nonlinear dynamics of gas-liquid two-phase flow in complex networks *Phys. Rev. E* **79** 066303
- [43] Marwan N, Donges J F, Zou Y, Donner R V and Kurths J 2009 Complex network approach for recurrence analysis of time series *Phys. Lett. A* **373** 4246–54
- [44] Davidsen J, Grassberger P and Paczuski M 2006 Earthquake recurrence as a record breaking process *Geophys. Res. Lett.* **33** L11304
- [45] Davidsen J, Grassberger P and Paczuski M 2008 Networks of recurrent events, a theory of records, and an application to finding causal signatures in seismicity *Phys. Rev. E* **77** 066104
- [46] Peixoto T P and Davidsen J 2008 Network of recurrent events for the Olami–Feder–Christensen model *Phys. Rev. E* **77** 066107
- [47] Robinson C 1999 *Dynamical Systems. Stability, Symbolic Dynamics, and Chaos* 2nd edn (Boca Raton, FL: CRC Press)
- [48] Hirata Y, Judd K and Kilminster D 2004 Estimating a generating partition from observed time series: symbolic shadowing *Phys. Rev. E* **70** 016215

- [49] Kurths J, Voss A, Saparin P, Witt A, Kleiner H J and Wessel N 1995 Quantitative analysis of heart rate variability *Chaos* **5** 88–94
- [50] Voss A, Kurths J, Kleiner H J, Witt A, Wessel N, Saparin P, Osterziel K J, Schurath R and Dietz R 1996 The application of methods of nonlinear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death *Cardiovasc. Res.* **31** 419–33
- [51] Daw C S, Finney C E A and Tracy E R 2003 A review of symbolic analysis of experimental data *Rev. Sci. Instrum.* **74** 915–30
- [52] Donner R, Hinrichs U and Scholz-Reiter B 2008 Symbolic recurrence plots: a new quantitative framework for performance analysis of manufacturing networks *Eur. Phys. J. Spec. Top.* **164** 85–104
- [53] Nicolis G, García Cantú A and Nicolis C 2005 Dynamical aspects of interaction networks *Int. J. Bifurcation Chaos* **15** 3467–80
- [54] Dellitz M, Hessel-von Molo M, Metzner P, Preis R and Schütte C 2006 Graph algorithms for dynamical systems *Analysis, Modeling and Simulation of Multiscale Problems* ed A Mielke (Heidelberg: Springer) pp 619–46
- [55] Padberg K, Thiere B, Preis R and Dellnitz M 2009 Local expansion concepts for detecting transport barriers in dynamical systems *Commun. Nonlinear Sci. Numer. Simulat.* **14** 4176–90
- [56] Gao Z-Y and Li K-P 2005 Evolution of traffic flow with scale-free topology *Chin. Phys. Lett.* **22** 2711–4
- [57] Li K-P and Gao Z Y 2006 A topological approach to traffic dynamics *Europhys. Lett.* **74** 369–75
- [58] Li X-G, Gao Z-Y, Li K-P and Zhao X-M 2007 Relationship between microscopic dynamics in traffic flow and complexity in networks *Phys. Rev. E* **76** 016110
- [59] Gao Z-Y, Li K-P, Li X-G, Huang H-J, Mao B-H and Zheng J-F 2007 Scaling laws of the network traffic flow *Physica A* **380** 577–84
- [60] Zheng J-F and Gao Z-Y 2008 A weighted network evolution with traffic flow *Physica A* **387** 6177–82
- [61] Li P and Wang B H 2006 An approach to Hang Seng Index in Hong Kong stock market based on network topological statistics *Chin. Sci. Bull.* **51** 624–9
- [62] Li P and Wang B-H 2007 Extracting hidden fluctuation patterns of Hang Seng stock index from network topologies *Physica A* **378** 519–26
- [63] Schwarz U, Benz A O, Kurths J and Witt A 1993 Analysis of solar spike events by means of symbolic dynamics methods *Astron. Astrophys.* **277** 215–24
- [64] Cellucci C J, Albano A M and Rapp P E 2005 Statistical validation of mutual information calculations: comparison of alternative numerical algorithms *Phys. Rev. E* **71** 066208
- [65] Donner R, Scholz-Reiter B and Hinrichs U 2008 Nonlinear characterization of the performance of production and logistics networks *J. Manuf. Syst.* **27** 84–99
- [66] Gao Z and Jin N 2009 Community structure detection in complex networks with applications to gas–liquid two-phase flow *Complex 2009, Part II (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering vol 5)* ed J Zhou (Berlin: Springer) pp 1917–28
- [67] Donner R Spatial correlations of river runoffs in a catchment *In Extremis: Extremes, Trends and Correlations in Hydrology and Climate* ed J Kropff and H-J Schellnhuber (Berlin: Springer) in press
- [68] Ni X-H, Jiang Z-Q and Zhou W-X 2008 Degree distributions of the visibility graphs mapped from fractional Brownian motions and multifractal random walks *Phys. Lett. A* **373** 3822–6
- [69] Lacasa L, Luque B, Luque J and Nuño J C 2009 The visibility graph. A new method for estimating the Hurst exponent of fractional Brownian motion *Europhys. Lett.* **86** 30001
- [70] Liu C, Zhou W-X and Yuan W-K 2009 Statistical properties of visibility graph of energy dissipation rates in three-dimensional fully developed turbulence arXiv:0905.1831 [physics.flu-dyn]
- [71] Tang Q and Liu J 2009 Network analysis of shanghai stock exchange composite index and Shenzhen stock exchange component index *Eight Wuhan Int. Conf. on E-Business* vols I–III ed F G Duserick (Alfred: Alfred University Press) pp 883–7
- [72] Yang Y, Wang J, Yang H and Mang J 2009 Visibility graph approach to exchange rate series *Physica A* **388** 4431–7

- [73] Elsner J B, Jagger T H and Fogarty E A 2009 Visibility network of United States hurricanes *Geophys. Res. Lett.* **36** L16702
- [74] Luque B, Lacasa L, Ballesteros F and Luque J 2009 Horizontal visibility graphs: exact results for random time series *Phys. Rev. E* **80** 046103
- [75] Gao Z and Jin N 2009 Complex network from time series based on phase space reconstruction *Chaos* **19** 033137
- [76] Donner R, Sakamoto T and Tanizuka N 2008 Complexity of spatio-temporal correlations in Japanese air temperature records *Nonlinear Time Series Analysis in Geosciences* ed R V Donner and S M Barbosa (Berlin: Springer) pp 125–55
- [77] Freeman L C 1979 Centrality in social networks—conceptual clarification *Soc. Netw.* **1** 215–39
- [78] Grassberger P and Procaccia I 1983 Characterization of strange attractors *Phys. Rev. Lett.* **50** 346–9
- [79] Newman M E J 2002 Assortative mixing in networks *Phys. Rev. Lett.* **89** 208701
- [80] Thiel M, Romano M C, Kurths J, Rolfs M and Kliegl R 2006 Twin surrogates to test for complex synchronisation *Europhys. Lett.* **75** 535–41
- [81] Romano M C, Thiel M, Kurths J, Mergenthaler K and Engbert R 2009 Hypothesis test for synchronization: twin surrogates revisited *Chaos* **19** 015108
- [82] van Leeuwen P, Geue D, Thiel M, Cysarz D, Lange S, Romano M C, Wessel N, Kurths J and Grönemeyer D H 2009 Influence of paced maternal breathing of fetal-maternal heart rate coordination *Proc. Natl Acad. Sci. USA* **106** 13661–6
- [83] Freeman L C 1977 A set of measures of centrality based on betweenness *Sociometry* **40** 35–41
- [84] Gao J B 1999 Recurrence time statistics for chaotic systems and their applications *Phys. Rev. Lett.* **83** 3178–81
- [85] Zou Y, Donner R V, Donges J F, Marwan N and Kurths J 2010 Identifying shrimps in continuous dynamical systems using recurrence-based methods *Phys. Rev. E* submitted
- [86] Ott E 1993 *Chaos in Dynamical Systems* (Cambridge: Cambridge University Press)
- [87] Nusse H E and Yorke J A 1994 *Dynamics: Numerical Explorations* (New York: Springer)
- [88] Dall J and Christensen M 2002 Random geometric graphs *Phys. Rev. E* **66** 016121
- [89] Zbilut J P, Mitchell J C, Giuliani A, Marwan N and Webber C L Jr 2004 Singular hydrophobicity patterns and net charge: a mesoscopic principle for protein aggregation/folding *Physica A* **343** 348–58
- [90] Krishnan A, Zbilut J P, Tomita M and Giuliani A 2008 Proteins as networks: usefulness of graph theory in protein science *Curr. Protein Pept. Sci.* **9** 28–38
- [91] Yang H, Zhao F, Qi L and Hu B 2004 Temporal series analysis approach to spectra of complex networks *Phys. Rev. E* **69** 066104
- [92] Li K P, Gao Z Y and Zhao X M 2008 Multiple scale analysis of complex networks using the empirical mode decomposition method *Physica A* **387** 2981–6
- [93] Csárdi G and Nepusz T 2006 The igraph software package for complex network research *InterJournal CX.18* 1695