LUDWIG-MAXIMILIANS-UNIVERSITÄT

MASTER THESIS

Network Analyses of Nonlinear Time Series and Complex Spatial Structures

Author:
Ingo LAUT

Supervisors:
Prof. Dr. Gregor MORFILL
Dr. Christoph RÄTH

Munich, May 21, 2014

Cover art is a three-dimensional plot of the Lorenz system (see Sec. 2.3.1; 25 time units at a time step length of 0.01) to the left and a scatter plot of a rotating three-dimensional complex plasma cluster observed in Wörner et al. [2012] (see Sec. 3.2; time evolution over 4 seconds plotted every 0.1 seconds) to the right. The projection onto the xy plane is shown in grey for both systems.
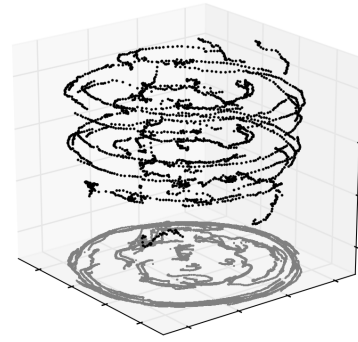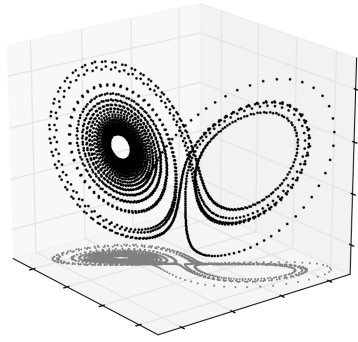
LUDWIG-MAXIMILIANS-UNIVERSITÄT

MASTERARBEIT

Netzwerkanalysen Nichtlinearer Zeitserien und Komplexer Räumlicher Strukturen

Autor:
Ingo LAUT

Betreuer:
Prof. Dr. Gregor MORFILL
Dr. Christoph RÄTH

München, den 21. Mai 2014

Nomenclature and Abbreviations

$\{a_n\}$	autoregressive process
A_{ij}	an adjacency matrix
A_{ij}^{cyl}, A_{ij}^{spy}	adjacency matrix for cylindrical, spherical properties
A_{ijt}	a multislice adjacency matrix
A_ρ, A_z	fitting parameters
$\{A_n\}$	amplitudes of the (complex) Fourier transform of $\{y_n\}$
α	parameter for unweighted network
B_ρ	fitting parameter
β	Lorenz parameter
c	parameter of AR process
$c_{\mathrm{PM}}(\Delta)$	cross correlation in phase map at shift Δ
C	average clustering coefficient
C_ν	clustering coefficient of node ν
$C_{\mathrm{p}}, C_{\mathrm{np}}$	periodic and nonperiodic autocorrelation function
C_{ijt}	a three-dimensional intraslice coupling matrix
$C(\tau)$	autocorrelation function
d	a dimension
$d_{i,j}$	horizontal distance
$\delta_{ij} = \delta(i,j)$	Kroenecker delta
$d_{\min}$	minimal horizontal distance in a time series
e	elementary charge
$\exp(\cdot)$	exponential function
ϵ	threshold for adjacency matrix of unweighted network
$\mathbf{F}(\mathbf{y}_i, T)$	predictor for point $\mathbf{y}_i$ at lead time T
$\mathbf{F}$	a Force
g	number of nearest neighbors for nonlinear prediction error
g_ν	community of node ν
γ	resolution parameter for community finding algorithm
η	distance of strings vs. extent of string
η_n	number drawn from Gaussian distribution
k_ν	number of neighbors of node ν
K	component of a graph
κ	average connectivity
κ_ν	connectivity of node ν
κ_{or}	parameter; ϵ will be chosen such that $\kappa = \kappa_{\mathrm{or}}$ for original data
$l_{\nu,i}$	length of path from note ν to node i
L	average path length
L_ν	path length of node ν
λ	screening length
m	mixing parameter

M	mass of a dust particle
μ	total edge strength of a network
n	number of constituents of a system
n^{cyl}	number of particles of cylindrical structure in a null model
n^{sph}	number of particles of spherical structure in a null model
$N_n^{\Delta}u$	number of links between neighbors of node ν
P	partition of a graph
$\{\phi_n\}$	phases of the (complex) Fourier transform of $\{y_n\}$
Q	quality of a partition
$\Theta(\cdot)$	Heaviside function
$R(t)$	stability of a partition
R	ratio $n^{\mathrm{sph}}/n^{\mathrm{cyl}}$
$\mathbf{r}_i, r_i$	3D vector of position of particle i and its modulus
ρ	Lorenz parameter
ρ_i	(cylindrical) radius of particle i
$\{s_n\}$	random shuffle of $\{y_n\}$
$S(X)$	significance of measure X
σ	Lorenz parameter
τ	a time: argument of correlation function; delay time for embedding
T	lead time for NLPE
$U_g(i)$	set of indices of the g nearest neighbors of point i
$\boldsymbol{\omega}, \omega$	rotation vector and its modulus
Ω_ρ, Ω_z	cylindrical and vertical confinement parameter
$\{y_n\}$	a time series of length n

3D	three-dimensional
AGN	active galactic nucleus
AR	autoregressive
Ark	Ark 564 (galaxy)
ASCA	Advanced Satellite for Cosmology and Astrophysics
MCG	MCG-6-30-15
Mrk	Mrk766 (galaxy)
NLPE	nonlinear prediction error
QPO	quasi periodic oscillation
RE J	RE J1034+396 (galaxy)
RQA	recurrence quantification analysis
XMM	X-ray Multi-Mirror Mission (satellite)
surrogate algorithms:	
FT	Fourier-transformed
AAFT	amplitude-adjusted Fourier-transformed
IAAFT	iterative amplitude-adjusted Fourier-transformed
SA	simulated annealing

Contents

Chapter 1

Networks

1.1   What is a (Complex) Network?

Originating from "classical" networks such as power grids [Watts and Strogatz, 1998] or the World Wide Web [Albert et al., 1999], complex network analysis has been adopted for a wide range of systems [Albert and Barabási, 2002; Boccaletti et al., 2006]. Common to all these applications is the procedure of associating the constituents of a system (current generators or routers) with the nodes of a network, and their interactions (transmission lines between generators or connections between routers) with edges connecting them.

The term *complex* network is used to set it apart from a regular network which was studied at the beginning of graph theory, and will sometimes be omitted in this work. Erdös and Rényi [1960] modeled large, complicated networks with random graphs, where any two nodes are connected with a probability p. Watts and Strogatz [1998], guided by the properties of real complex networks, proposed the small world[1] network, which is created by a "rewiring" process of a ring of n nodes, each connected to its k nearest neighbors. The rewiring is done by reconnecting each node to a random node with probability p. Scale-free networks focus more on the dynamics of a growing network. With a preferred attachment of new nodes to nodes that have already relatively many neighbors, scale-free networks reproduce the power-law degree distribution (see Sec. 1.2 for the definition of degree centrality) observed in many real networks [Barabási and Albert, 1999].

In recent years, complex network analysis has also been applied to systems whose architecture is not obviously network-like. The contributions of network analysis to time series analysis are eminent and lead to insights in both fields [Lacasa et al., 2008; Donner et al., 2010; Donges et al., 2013]. Community detection, which has contributed vastly to the understanding of structural properties of large networks, was used in other fields: Mucha et al. [2010], for example, used it to divide the representatives of the U.S. senates of the last two centuries in groups, depending on

---

[1]The name *small world* stems from Milgram [1967] who found that the chain of acquaintances between any two persons in the US has typically the length six.
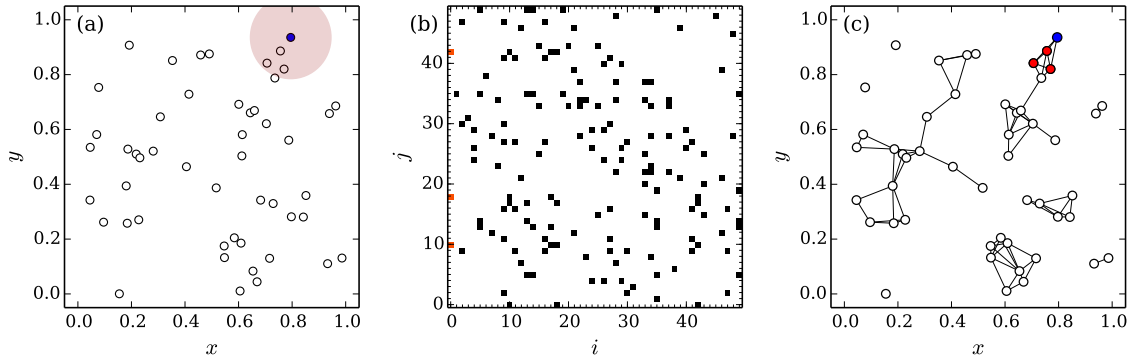
Figure 1.1: Creating a simple network. (a) Scatter plot of 50 points with x and y positions drawn from a uniform distribution. A ball with radius $\epsilon = 0.15$ centered on the zeroth point (in blue) contains the particles closer than ϵ. (b) The corresponding adjacency matrix. Black dots denote unity entries in the matrix and denote the connections in the network. Accordingly, the neighbors of point 0 are shown in the zeroth column (or line) of the matrix. They are highlighted in red. (c) A representation of the network where the nodes have the coordinates of the points they correspond to. The neighbors of the zeroth node are highlighted in red.

the similarity of their voting behavior. As expected for a two-party system, most of the senates consist of two communities. The senates with more than two communities at a time, however, correspond to historically important events like the beginning of the American Civil War.

The information stored in a network can be expressed by a square matrix A of size n, where n is the number of constituents of the system. A is called the adjacency matrix. A_{ij} then takes the value of the edge connecting nodes i and j. For an undirected network, the matrix is symmetric; for an unweighted network, $A_{ij} \in \{0, 1\}$. The components of a network are sets of nodes that are connected by paths of one or more non-zero edges. In Fig. 1.1(a), 50 points with random xy positions are plotted. A network is created by considering points that are closer to each other than a threshold $\epsilon = 0.15$ as neighbors of the corresponding network. The adjacency matrix then reads

$$A_{ij}(\epsilon) = \Theta\left(\epsilon - \|\mathbf{x}_i - \mathbf{x}_j\|\right) - \delta_{ij}, \tag{1.1}$$

where $\Theta(\cdot)$ is the Heaviside function and $\mathbf{x}_i$ is the two-dimensional position vector of point i. The Kronecker δ sets the diagonal terms to zero in order to avoid self-loops in the network: a point is not considered to be its own neighbor. A is shown in Fig. 1.1(b). In Chapters 2 and 3, similar definitions of the adjacency matrix, albeit with different measures of distance, will be used to create networks.

1.2   Measures on Networks

As was stated in Sec. 1.1, the degree centrality reflects important properties of the network. This measure and others will be described below. The network measures reduce the information stored in the adjacency matrix to a value or a distribution of values that characterize the network at a given scale.

As a first network measure, the degree centrality k_ν counts the number of nodes that are connected to node ν [Freeman, 1979]. It is defined as

$$k_\nu = \sum_{i=0}^{n-1} A_{\nu,i}.$$ (1.2)

Dividing by the maximum possible value yields the connectivity κ_ν. The average connectivity κ is defined as the mean of all nodes κ_ν,

$$\kappa = \frac{1}{n} \sum_{\nu=0}^{n-1} \kappa_\nu = \frac{1}{n(n-1)} \sum_{\nu=0}^{n-1} k_\nu.$$ (1.3)

Taking into account only the immediate neighbors of the node[2], this measure may be regarded as a local measure. For a random network, k_ν obeys a Poisson distribution, for scale-free networks on the other hand, it obeys a power-law distribution.

The clustering coefficient C_ν [Watts and Strogatz, 1998] acts on intermediate scales of the network. It evaluates the number N_ν^Δ of links between neighbors of a given node versus the maximum possible number $k_\nu(k_\nu - 1)/2$:

$$C_\nu = \frac{2}{k_\nu(k_\nu - 1)} N_\nu^\Delta = \frac{1}{k_\nu(k_\nu - 1)} \sum_{i,j=0}^{n-1} A_{\nu,i} A_{i,j}.A_{j,\nu}.$$ (1.4)

The average clustering coefficient is the mean value of the C_ν, summed over all nodes of the network,

$$C = \frac{1}{n} \sum_{\nu=0}^{n-1} C_\nu.$$ (1.5)

Finally, the path length L_ν of node ν is considered. For every node ν of the network, the minimum numbers $l_{\nu,i}$ of edges that have to be traversed to get to any other node i of the same component are calculated. The path length L_ν is then calculated by averaging over all nodes i that are in the same component as ν.

$$L_\nu = \frac{1}{n_K(\nu)} \sum_{i=0}^{n_K(\nu)-1} l_{\nu,i},$$ (1.6)

---

[2]Note that depending on the definition of the adjacency matrix, connected nodes (neighbors) of the network do not necessarily represent constituents of the system that have a small spatial separation.
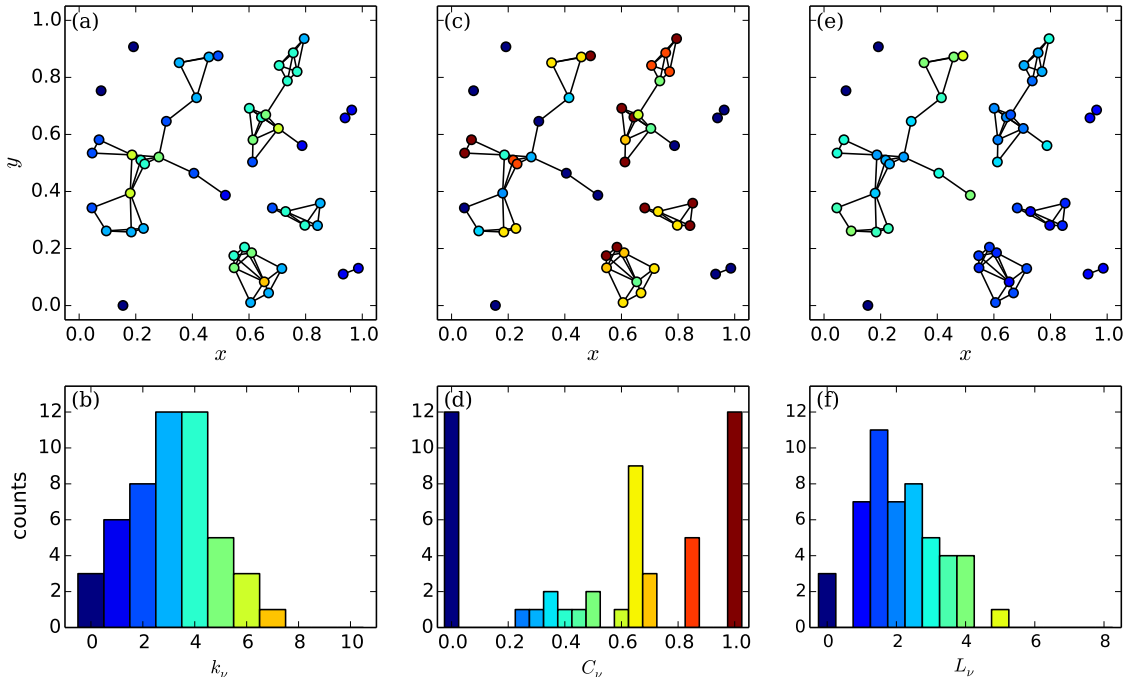
Figure 1.2: (a) The Network of Fig. 1.1 color coded with the network measure degree centrality k_ν. The histogram (b) has the same colors as a function of k_ν. For the zeroth node, highlighted in Fig. 1.1, $k_0 = 3$. (c), (d): The same for the clustering coefficient C_ν. C_ν can take a multitude of non-integer values, the bin size was therefore reduced. $C_0 = 1$, since all neighbors of the zeroth node are connected to each other. (e), (f): The same for path length L_ν. Being at the border of its component, the path length for the zeroth particle is relatively big $L_0 \simeq 3.09$.

where $n_K(\nu)$ is the number of nodes in the same component as node ν. By averaging over all nodes ν, the average path length L is obtained [Watts and Strogatz, 1998]:

$$L = \frac{1}{n} \sum_{\nu=0}^{n-1} L_\nu. \tag{1.7}$$

The small world property mentioned in Sec. 1.1 is characterized by a relatively small average path length L (like a random network) and at the same time a high average clustering C (like a regular network, much higher than a random network). This is modeled by the rewiring step, which dramatically reduces L of the initial ring network already for small $p \ll 1$, while keeping the value of C almost unchanged.

The measures on the network from Sec. 1.1 are plotted in Fig. 1.2. As can be seen in Fig. 1.2(a), the nodes of dense regions have a relatively high degree centrality k_ν. The average path length L_ν, [see Fig. 1.2(e)], on the other hand, yields high values for nodes that are situated at the border of large components. For the clustering coefficient C_ν, which depends on the "connectedness" of the neighbors of a node, it is difficult to find a pattern in the coloring of the nodes. The value for an isolated node is equal to zero for each of the considered network measures.

1.3   Community Finding in Networks

If a network has only one component, it is called connected. Communities, on the other hand, are sets of nodes in a connected network which are "more connected" to each other than to the rest of the network. In a network communities often correspond to the functional units of the system under consideration. For the example of power grids from Sec. 1.1, communities may represent different cities, or, on a larger scale, countries. Having defined the communities, one may then consider a higher-level network, where each community is represented as a node, to examine the interactions of these functional units. In this Section sums without limits are summed over all possible values.

Common to the different methods to partition a network into communities is the definition of the goodness of a partition. A quality function Q quantifies a partition by comparing the number of intracommunity edges of the partition to the expectation value of a null model [Porter et al., 2009]:

$$Q \propto \sum_{ij} \left(A_{ij} - A_{ij}^{\mathrm{null}} \right) \delta(g_i, g_j). \tag{1.8}$$

Here, A_{ij} is the adjacency matrix of the network, A_{ij}^{null} is the expected weight between nodes i and j under a given null model and g_i is the community assignment of node i. A popular choice for the null model is the *modularity* proposed by Newman and Girvan [Newman and Girvan, 2004; Newman, 2006]:

$$A_{ij}^{\mathrm{null}} = \frac{k_i k_j}{2\mu}, \tag{1.9}$$

where k_i is the degree of node i and $\mu = \sum_{i,j} A_{ij}/2$ is the total edge strength in the network.

The quality Q of a partition can also be expressed in terms of its stability R, that is, an autocovariance function of a Markov process on the network [Lambiotte et al., 2008]. With this equivalence of the quality function on one hand, and Laplacian dynamics on the other hand, a resolution parameter can be derived, determining the size of the communities that maximize Q. Considering an ergodic Markov Process for the underlying dynamics, the stability of partition P can be written as [Lambiotte et al., 2008]

$$R(t) = \sum_K P(K, t) - P(K, \infty), \tag{1.10}$$

where $P(K, t)$ is the probability for a walker to be in the same community K initially and at time t. $R(t)$ describes the probability that a random walker remains within the same community after a time t, relative to the expected value under independence. The sum includes all communities of the network. Assuming independent, identical Poisson processes on each node of the graph, Lambiotte et al. [2008] derived with the Kolmogorov equation for the probability density p_i of a random walker i,

$$\dot{p}_i = \sum_j \frac{A_{ij}}{k_j} p_j - p_i \equiv \sum_j L_{ij} p_i \tag{1.11}$$

and its stationary solution $p_i^* = k_i/2\mu$ the stability to be

$$
\begin{aligned}
R(t) &= \sum_K \sum_{i,j \in K} \left[\left(e^{tL} \right)_{ij} \frac{k_j}{2\mu} - \frac{k_i}{2\mu} \frac{k_j}{2\mu} \right] \delta(g_i, g_j) \\
&\approx \sum_K \sum_{i,j \in K} \left[\delta_{ij} \frac{k_j}{2\mu} + t \left(\frac{A_{ij}}{k_j} - \delta_{ij} \right) \frac{k_j}{2\mu} - \frac{k_i}{2\mu} \frac{k_j}{2\mu} \right] \delta(g_i, g_j) \qquad (1.12) \\
&= \frac{1}{2\mu} \sum_K \sum_{i,j \in K} \left[tA_{ij} - \frac{k_i k_j}{2\mu} \right] \delta(g_i, g_j) + \frac{1}{2\mu} \sum_K \sum_{i,j \in K} \left[\delta_{ij} \left(k_i - tk_i \right) \right].
\end{aligned}
$$

Here, the matrix exponential was expanded to first order in t. Omitting the second sum, which makes the same contribution for all partitions due to the Kronecker δ, and multiplying by $\gamma \equiv 1/t$, the stability can be understood as a modularity:

$$
R(\gamma) = \frac{1}{2\mu} \sum_{ij} \left[A_{ij} - \gamma \frac{k_i k_j}{2\mu} \right] \delta(g_i, g_j), \qquad (1.13)
$$

where the resolution parameter γ now determines the scale at which communities will be found in the network. $R(\gamma)$ reduces to the modularity proposed by Newman and Girvan [2004] (Eq. 1.9) for $\gamma = 1$. By varying γ, the community structure at all scales of the systems can be examined.

In the same manner, Mucha et al. [2010] derived a quality function for a multislice network, i.e., a set of adjacency matrices A_{ijt}, one for every time step t, and interslice couplings C_{jtr} connecting node j at time t to itself at time r [Mucha et al., 2010]. It can be derived as

$$
Q_{\text{multislice}}(\gamma) = \frac{1}{2\mu} \sum_{ijtr} \left[\left(A_{ijt} - \gamma \frac{k_{it} k_{jt}}{2\mu_t} \right) \delta_{tr} + \delta_{ij} C_{jtr} \right] \delta(g_{it}, g_{jr}), \qquad (1.14)
$$

where γ is the resolution parameter and $k_{jt} = \sum_i A_{jit}$ the strength of an individual node. The normalization factor is defined as $2\mu = \sum_{jt} (k_{jt} + c_{jt})$ with $c_{jt} = \sum_r C_{jtr}$ the strength of node j across the slices.

Multislice networks are particularly interesting for systems that evolve over time [Bassett et al., 2013]. The interslice coupling C_{jtr} can be chosen to connect every node j to itself for two successive time steps t and r. Then, the nodes tend to stay in the same community over time, except if the structure of the network changes such that a node becomes very connected to a different community. The value of C_{jtr} determines how this coupling to the future time step r compares to the connections A_{ijt} in the system at time t. The dynamics of functional units for a system can thus be examined, at a scale defined by resolution parameter γ.

The algorithm used here is an adaption of the so-called Louvain method that aggregates nodes to small communities which are the nodes of a new network at a later iteration step until Q is maximized [Mucha et al., 2010; Blondel et al., 2008].

Chapter 2

Analysis of Nonlinear Time Series

2.1 Motivation

One main goal of nonlinear time series analysis is to determine whether a system has underlying chaotic dynamics [Kantz and Schreiber, 2004; Strogatz, 2006]. To this end, some measure for nonlinearity is calculated for the time series. In order to make a significant statement, the measure may also be applied to a set of so-called surrogate data sets which mimic the linear properties of the original data (see Sec. 2.2.1).

Most tests for nonlinearity are derived or at least motivated from chaos theory. In his much renowned paper on cellular convection, Lorenz [1963] conjectured that low-dimensional, deterministic equations can show chaotic behavior in such a way that "prediction of the sufficiently distant future is impossible by any method, unless the present conditions are known exactly". The attractor of such systems was later called *strange* [Ruelle and Takens, 1971], as it often exhibits a fractal structure.

The correlation dimension, as one example of a measure for nonlinearity, attempts to estimate the dimensionality of this fractal structure [Grassberger and Procaccia, 1983]. To this end, the correlation sum is calculated, which counts how many points in phase space are nearer to each other than a distance r. The slope of the double-log plot of correlation sum versus r then yields the correlation dimension. If the value of the slope is constant in a range of r values and not equal to an integer, it is probable that the system has chaotic dynamics. Noise can, however, easily blur the fractal geometry of the trajectory [Schreiber and Schmitz, 1997].

The correlation dimension can be attributed to the class of static measures for nonlinearity, since it considers the trajectory in phase space as a whole [Wackerbauer et al., 1994]. The nonlinear prediction error, which takes the dynamics at each time step into account, is an example of a dynamic measure for nonlinearity and will be described in Sec. 2.2.4.

Since a network contains the complete information about interactions of the system, time series may also be subject of analysis. In order to use network analysis for the detection of nonlinearities, first of all a suitable network has to be found. While it is straightforward to derive a network from systems like social or neural

networks [Albert and Barabási, 2002], there is no "natural" way of how to create a network from a time series.

One approach is the recurrence network [Xu et al., 2008; Marwan et al., 2009] motivated from recurrence quantification analysis (RQA) [Eckmann et al., 1987]. In RQA, the time series is embedded in an artificial phase space (see Sec. 2.2.2). The recurrence matrix then contains the information which points are sufficiently close to each other in phase space. The structural properties of the recurrence matrix can be used to characterize different dynamic aspects of the time series. By interpreting the recurrence matrix as an adjacency matrix, a network can be constructed from the time series. This approach may be used to characterize the underlying dynamical system [Donner et al., 2010] or detect dynamic changes by a sliding-window technique [Marwan et al., 2009]. Both RQA and the analysis of recurrence networks are static measures for nonlinearity.

In this chapter, recurrence networks are used to test for weak nonlinearities in time series. To this end, network measures are calculated for the time series under study and their surrogate data sets. Four different surrogate generating algorithms are examined. The performance of the network test is compared to the nonlinear prediction error. The Lorenz system is used as an example of low-dimensional chaotic time series. The data is mixed with a (linear) autoregressive process in order to test both the ability to detect weak nonlinearities and the susceptibility to erroneously consider a linear time series as nonlinear. As real-world data, light curves of active galactic nuclei (AGN) are examined. AGN represent an active research domain and nonlinear time series analysis may help to understand their fundamental physical processes.

2.2   Tools

2.2.1   Surrogate Data

Surrogates are an important tool for the detection of nonlinearities in time series. They are data sets which mimic the linear properties, i.e. the autocorrelation function, of the original data, while possible higher order correlations are randomized. Comparing the result of some measure for nonlinearity for the original time series with those for the surrogate data offers a rigorous statistical test for nonlinearity. The most commonly used methods for generating surrogates are Fourier-transformed (FT) surrogates and their amplitude-adjusted (AAFT) and iterative amplitude-adjusted FT (IAAFT) generalizations, as well as simulated annealing (SA) surrogates. As it is impossible to perfectly reproduce both the autocorrelation function as well as the amplitude distribution, the available surrogate generating algorithms focus on different aspects. They will be described below.

FT surrogates are compatible with the null hypothesis of a linear Gaussian process [Theiler et al., 1992]. They are generated by randomizing the phases $\{\phi_n\}$ of the discrete Fourier transform of the original time series and subsequently performing

the inverse transform. The Wiener-Khinchin theorem guarantees the surrogates to have the same autocorrelation function as the original time series. Being truly linear, the surrogates can unveil higher order correlations, however, this test is limited to time series which themselves obey a Gaussian distribution.

AAFT surrogates extend the null hypothesis to a Gaussian process which was distorted by an instantaneous, time-independent measurement function [Theiler et al., 1992]. Here, the original time series is first rank-order remapped[1] to a set of random numbers drawn from a Gaussian distribution. Then, FT surrogates of this remapped time series are created. Finally, the surrogates are rank-order remapped to the original time series. The surrogates now mimic both the autocorrelation function and the amplitude distribution. The final step, however, leads to a whitening of the power spectrum (or, equivalently, of the autocorrelation function) as compared to the original time series. In Schreiber and Schmitz [1996] it was shown that this may lead to a false detection of nonlinearity in purely linear time series.

IAAFT surrogates were designed to overcome this shortcoming [Schreiber and Schmitz, 1996, 2000]. The IAAFT scheme starts with a random shuffle $\{s_n\}$ of the original time series $\{y_n\}$. The Fourier amplitudes $\{A_n\}$ of the original time series are saved. Now, the following two steps are repeated iteratively. (1) Take the Fourier transform of $\{s_n\}$, replace the corresponding Fourier amplitudes by $\{A_n\}$, and transform back. $\{s_n\}$ now has exactly the same autocorrelation function as $\{y_n\}$, but not the same amplitude distribution. (2) Rank-order remap the resulting $\{s_n\}$ to $\{y_n\}$. As step (2) changes the power spectrum of the surrogate $\{s_n\}$, the two steps are repeated iteratively. The algorithm may be stopped after a given number of iterations, or continued until the rank-order remapping no longer leads to a change in the surrogate. Here, the algorithm is always stopped after step (1) in order to have exactly the same power spectrum as the original time series.

SA surrogates follow a different iteration scheme [Schreiber, 1998], and were promoted to be more flexible, for example for data that are not evenly sampled. Initially, after calculating the autocorrelation function $C_{\text{data}}(\tau)$ of the data, the data are shuffled. Then, at each iteration step, the following scheme is repeated: Calculate the autocorrelation $C_{\text{before}}(\tau)$. Swap[2] two random data positions a and b, and calculate the new autocorrelation $C_{\text{swap}}(\tau)$. Calculate the cost function $E[C(\tau)] = \sum_{\tau=1}^{n}[C(\tau) - C_{\text{data}}(\tau)]/\tau$ for the two autocorrelation functions, as well as their difference $dE = E[C_{\text{swap}}] - E[C_{\text{before}}]$. Now, do a so-called *Metropolis* step [Metropolis et al., 1953]: If $dE \leq 0$, accept the swapping; if $dE > 0$, accept the swapping only with a probability $\exp(-dE/kT)$ and revert it otherwise. The thermal energy term kT controls the acceptance of an increase of the cost function. kT decreases with increasing iteration steps, hence the name of the scheme.

---

[1]The rank-order remapping of a data set $\{x_n\}$ to $\{y_n\}$ is a reordering of $\{y_n\}$ such that if x_i is the ith smallest of the $\{x_n\}$, then y_i is the ith smallest of the $\{y_n\}$. The reordered $\{y_n\}$ thus "follow" the original set $\{x_n\}$ while having an amplitude distribution identical to the one of $\{y_n\}$.

[2]For a and b drawn randomly from zero to n, the swapping is done as follows: tmp $= s_a$, $s_a = s_b$, $s_b =$ tmp.

The surrogates based on the Fourier transform assume a periodic autocorrelation $C_{\mathrm{p}}(\tau) = 1/n \sum_{i=0}^{n-1} x_i x_{(i+\tau)\%n}$ where % is the *modulo* operator. SA surrogates are not, in this perspective, limited, and a more general, nonperiodic autocorrelation can be used[3]: $C_{\mathrm{np}}(\tau) = 1/(n - \tau) \sum_{i=0}^{n-\tau-1} x_i x_{i+\tau}$. This makes sense when a periodic signal cannot be assumed, or, more specifically, start and end values of a time series are significantly different. Here, only the nonperiodic autocorrelation function is considered for the SA surrogates.

2.2.2   Embedding of a Time Series

Often, only a one-dimensional series of observables is available for a system that is expected to depend on more than one parameters. Takens [1981] addressed the question how the phase space of an attractor could be reconstructed from this time series of observables. He proved that there exists an embedding, i. e., a continuous mapping of the time series to a space that is topologically equivalent to the attractor.

To embed a time series in d dimensions, one may use the consecutive time derivatives [Packard et al., 1980]:

$$\mathbf{y}_i = \left(y_i, \frac{dy_i}{dt}, ..., \frac{d^{d-1}y_i}{dt^{d-1}} \right). \tag{2.1}$$

Due to the uncertainties for estimating higher time derivatives for time series, the method of delay coordinates is generally preferred [Packard et al., 1980]:

$$\mathbf{y}_i = (y_i, y_{i+\tau}, \ldots, y_{i+(d-1)\tau}), \tag{2.2}$$

where τ is the delay time. The embedded time series has $(d - 1)\tau$ less data points than the original time series.

For an experimental time series whose phase space is not known *a priori*, suitable delay time τ and embedding dimension d have to be found. One method of determining the embedding dimension is the false neighbor method [Kennel et al., 1992], where the behavior of near neighbors under changes in the embedding dimension are examined. The idea is that if the embedding dimension is too small, points may be near to one another only because the attractor is projected onto a lower dimension. By increasing the embedding dimension, the percentage of these false neighbors can be calculated. The embedding dimension with a small enough percentage of false neighbors is used. Choosing a too high embedding dimension leads to a sparsely filled phase space that complicates the analysis. Also, especially when dealing with time series with a small number of data points or when a large delay time τ has to be chosen, a high dimensional embedding decreases the number of points in the embedding space

For determining an appropriate delay time τ three values can be considered: first zero-crossing or first minimum of the autocorrelation function, or first minimum of mutual information.

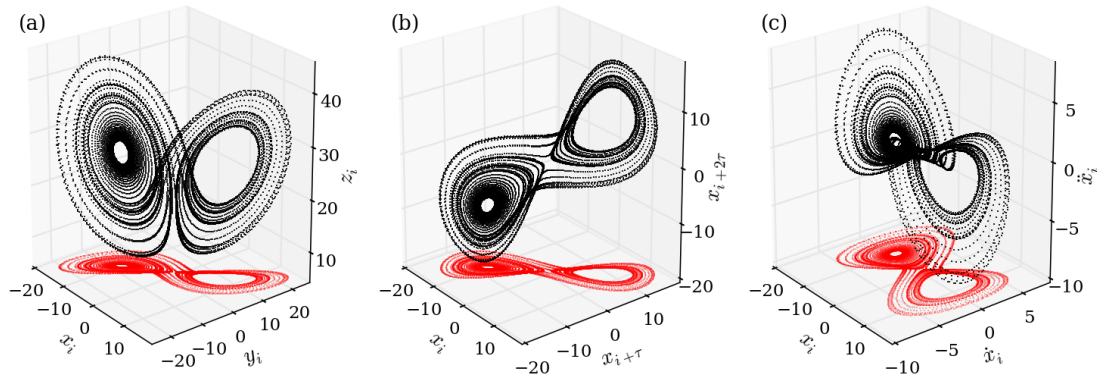[3]The SA algorithm can be used to minimize *any* cost function.

Figure 2.1: Embedding of a time series. (a) 10000 points of the Lorenz system at a step length $h = 0.005$. The data is plotted in 3D, with the xy projection shown in red. (b) Time delay embedding of the x coordinates from (a) in $d = 3$ dimensions with delay time $\tau = 18$ time steps. (c) Embedding of the x coordinates with the first two time derivatives $\dot{x}_i = (x_{i+1} - x_{i-1})/2h$ and $\ddot{x}_i = (x_{i+1} - 2x_i + x_{i-1})/h^2$.

In Fig. 2.1(a), the three-dimensional (3D) Lorenz system (see Sec. 2.3.1) is plotted. The attractor, which resembles a butterfly, is clearly visible. This attractor is reconstructed from only the x coordinates of the data using the method of delay coordinates in Fig. 2.1(b). For this simple example, the embedding dimension $d = 3$ was chosen in order to compare it with the original attractor. The delay time τ is chosen by visual inspection of the result. In Fig. 2.1(c), an embedding using time derivatives is shown. For the smooth Lorenz system, this method also works very well for reconstructing the topology of the attractor.

2.2.3 Measures for Nonlinearity derived from Networks

In order to test a time series for nonlinearity with the help of network analysis, it is first embedded in an artificial phase space. Then, a network is created from the embedded time series and different network measures are obtained. These network measures are compared to the measures from networks which correspond to surrogate data. In Donges et al. [2011], a similar approach was applied to the sliding-window method, where the "surrogates" were randomly drawn points from the embedded time series.

The adjacency matrix of the recurrence network [Donner et al., 2010] reads

$$A_{ij}(\kappa_{\mathrm{or}}) = \Theta\left(\epsilon(\kappa_{\mathrm{or}}) - \|\mathbf{y}_i - \mathbf{y}_j\|\right) - \delta_{ij}, \tag{2.3}$$

where the $\{\mathbf{y}_n\}$ are now the vectors in embedding space. The threshold $\epsilon(\kappa_{\mathrm{or}})$ of the recurrence network is chosen such that the global connectivity of the network from the embedded time series is $\kappa = \kappa_{\mathrm{or}}$. The same threshold is then used for the surrogate data. For the FT surrogates, the time series is rank-order remapped to a

Gaussian distribution before calculating the threshold of the recurrence network and the network measures.

The network measures average connectivity κ (Eq. 1.3), average clustering coefficient C (Eq. 1.5) and average path length L (Eq.1.7) are used for a statistical test to determine whether the null hypothesis of a linear time series has to be rejected. The idea is that if the original time series is nonlinear, its structure in embedding space will be different to that of its surrogates. This difference in structure is examined with the aid of network measures. If the measure of the original time series has the maximal or minimal value compared to 19 realizations of surrogate data, the null hypothesis is rejected. The size of a test is the probability that the time series has the maximal or minimal value by chance [Schreiber and Schmitz, 1997]. For the total of 20 data sets (the original data and 19 surrogates), the size of the test is 0.10. Repeating the test for a sufficiently large data set yields a rejection probability.

In the case of the AGN light curves (see Sec. 2.4), a statistical test is not appropriate, since the number of data sets is limited. Instead, the significance $S(X)$ of the deviation of the time series from the surrogates is calculated as

$$S(X) = \frac{|X_{\text{or}} - \langle X_{\text{surrogate}} \rangle|}{\sigma(\{X_{\text{surrogate}}\})}, \tag{2.4}$$

where X is a measure for nonlinearity, and X_{or} and $\{X_{\text{surrogate}}\}$ the values of the measure for the original time series and the surrogates. $\langle \cdot \rangle$ and $\sigma(\cdot)$ denote average value and standard deviation, respectively.

2.2.4   Nonlinear Prediction Error

The nonlinear prediction error (NLPE) is one of the most powerful measures for nonlinearity in time series [Sugihara and May, 1990]. In Schreiber and Schmitz [1997] it was found to be the one with the best overall performance for different nonlinear data sets. The underlying idea is that if a nonlinear time series is embedded with suitable parameters, points that are close to each other in phase space have a comparable time evolution. By averaging over the future positions of the neighbors of a point $\mathbf{y}_i$, an estimate $\mathbf{F}(\mathbf{y}_i, T)$ for the future position is obtained. The local nonlinear prediction error is then the distance from this estimate to the actual future position in phase space $\mathbf{y}_{i+T}$ (see Fig. 2.2). The NLPE is then defined as

$$\text{NLPE}(\{\mathbf{y}_n\}) = \frac{1}{\sqrt{n-T}} \sqrt{\sum_{i=0}^{n-T-1} [\mathbf{y}_{i+T} - \mathbf{F}(\mathbf{y}_i, T)]^2}, \tag{2.5}$$

where T is the lead time. T determines for how many time steps into the future of the points the prediction is made.

There are multiple ways of defining the predictor, here a *fixed mass* method is used:

$$\mathbf{F}(\mathbf{y}_i, T) = \frac{1}{g} \sum_{j \in U_g(i)} \mathbf{y}_{j+T}, \tag{2.6}$$
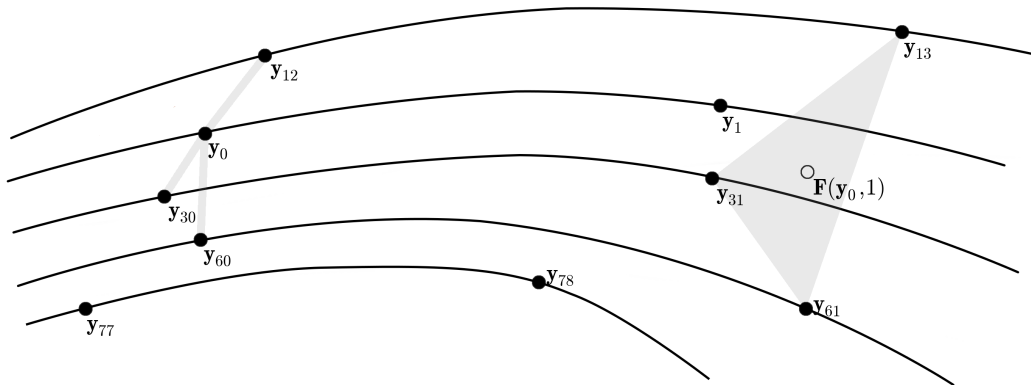
Figure 2.2: Sketch of the method of the nonlinear prediction error (NLPE). The attractor of a system is represented as lines, and the data points are numbered in an arbitrary way. The NLPE for $\mathbf{y}_0$ is obtained as follows. Find the 3 nearest neighbors of $\mathbf{y}_0$. Find the center $\mathbf{F}(\mathbf{y}_0, 1)$ of the future position of these neighbors a lead time $T = 1$ ahead. Calculate the Euclidean distance between $\mathbf{y}_1$ and $\mathbf{F}(\mathbf{y}_0, 1)$. The NLPE of the system is then obtained by averaging over the prediction errors of all points.

where $U_g(i)$ is the set of indices of the g nearest neighbors of point $\mathbf{y}_i$. Another possibility is the *fixed ball* method, where all particles nearer than a given radius are considered. Care has to be taken when choosing the radius, as it must not be too small for the points to have at least one neighbor in the sparse regions of the phase space, and be small enough to resolve the smaller scales in the dense regions.

Due to the smooth dynamics on the attractor, the error is expected to be smaller for a nonlinear time series than for a random time series with the same linear properties. Accordingly, the null hypothesis is only rejected in this case, reducing the size of the test to 0.05. For the AGN data, the significance of the test is calculated according to Eq. 2.4.

2.3 Analysis of the Lorenz System

Artificial, nonlinear data obtained from integrating the Lorenz system are examined in order to test the measure for nonlinearity based on network analysis. The null hypothesis that the time series is in agreement with a linear process is tested with different surrogate generating algorithms. The data is mixed with a linear process to test both the ability to reject the hypothesis when nonlinearities are present and the susceptibility to erroneously do so when the data is purely linear. In previous studies, noise of a given amplitude was simply added to the time series under study [Schreiber and Schmitz, 1997]. The embedding proposed here enables a throughout analysis in the range from a purely linear time series to a nonlinear signal.
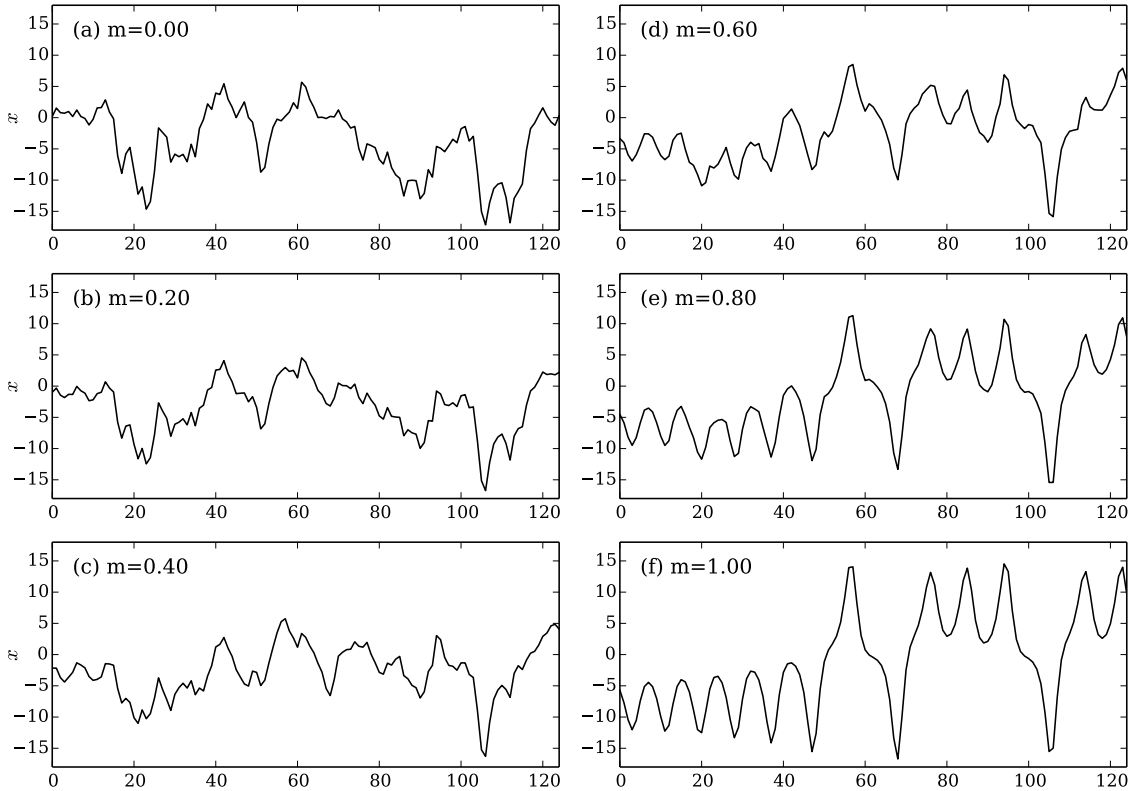
Figure 2.3: Lorenz system mixed with an AR process. The 125 first steps of a data set of 2000 steps are shown for different mixing parameters m. The step size for the Lorenz system is 0.08 time units.

2.3.1 Data Sets

The tests for nonlinearity are applied to the well known Lorenz system [Lorenz, 1963], which is given by the following equations:

$$\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= \sigma(y - x), \\
\frac{\mathrm{d}y}{\mathrm{d}t} &= x(\rho - z) - y, \\
\frac{\mathrm{d}z}{\mathrm{d}t} &= xy - \beta z.
\end{aligned} \tag{2.7}$$

The parameters are chosen to have the standard values $\sigma = 10$, $\beta = 8/3$ and $\rho = 28$. A 3D plot of the Lorenz system at a small step size can be seen in Fig. 2.1. In the following, data sets containing 2000 time steps with a larger step size of 0.08 time units are considered. The x coordinates $\{x_n\}$ of the data are subject of the analysis.

The univariate data $\{x_n\}$ are mixed with a linear autoregressive (AR) process $\{a_n\}$ in order to generate a new series $\{y_n\}$ which contains a fraction m of the nonlinear series,

$$y_n = m \cdot x_n + (m - 1) \cdot a_n, \tag{2.8}$$

where m is called the mixing parameter. m includes the whole range from a linear process to a nonlinear time series in the interval $m \in [0,1]$. Ideally, a measure for nonlinearity should not reject the null hypothesis for $m = 0$, and detect nonlinearities for $m > 0$. In this work, six values of m are used to evenly sample this interval. This sampling is already sufficient to depict differences in the measures for nonlinearity and the different surrogate generating algorithms.

The same distorted AR process as in Schreiber and Schmitz [1996] is used, it reads

$$
\begin{aligned}
a_n &= x_n \sqrt{|x_n|}, \\
x_n &= c x_{n-1} + \eta_n,
\end{aligned}
\tag{2.9}
$$

with $c = 0.9$ and noise η_n drawn from a Gaussian distribution. The first 125 steps of the AR process, corresponding to $m = 0$ of the mixed time series, are shown in Fig. 2.3(a). Other realizations for different m up to $m = 1.00$ (pure Lorenz) are shown in Fig. 2.3(b)-(f). Note that due to the different amplitude distributions of the Lorenz system and the AR process, $m = 0.5$ does not correspond to equal parts of linear and nonlinear contributions.

2.3.2   Finding Nonlinearities in the Lorenz System

The mixed Lorenz data is embedded in $d = 3$ dimensions. The delay time for the embedding is set to $\tau = 2$ time steps such that the attractor is clearly reconstructed for $m = 1$. As proposed in Sugihara and May [1990], the number of nearest neighbors for the predictor was set to $g = d + 1 = 4$. The best results of the NLPE were found for a lead time $T = 2$.

The discrimination power of the NLPE as a function of m can be seen in Fig. 2.4(a). The AAFT surrogates erroneously reject the null hypothesis at $m = 0$, where the series is a pure AR process, with a probability of more than $20\,\%$. This observation was used in Schreiber and Schmitz [1996] to motivate the advantages of the IAAFT algorithm which indeed shows a lower percentage of false rejections. In this paper, however, the performance of the algorithms was not compared for time series with nonlinearities. In Fig. 2.4(a) it can be seen that at $m = 0.2$, the rejection probability of the IAAFT algorithm is still around $10\,\%$ (which corresponds to the size of the test) while the FT and AAFT surrogates already detect nonlinearities. Stopping the IAAFT algorithm already after ten iterations has no effect, these surrogates [labeled IAAFT(10) in the Figure] yield results almost identical to the IAAFT surrogates. The FT algorithm shows a good performance, starting at about $10\,\%$ at $m = 0$ and showing a rejection probability of more than $20\,\%$ at $m = 0.2$. The SA algorithm shows a performance comparable to the IAAFT algorithm that is almost constant for $m \leq 0.2$.

The same embedding parameters are used for the network measures. The parameter controlling the threshold of the adjacency matrix (see Eq.2.3) is set to $\kappa_{\mathrm{or}} = 0.01$.
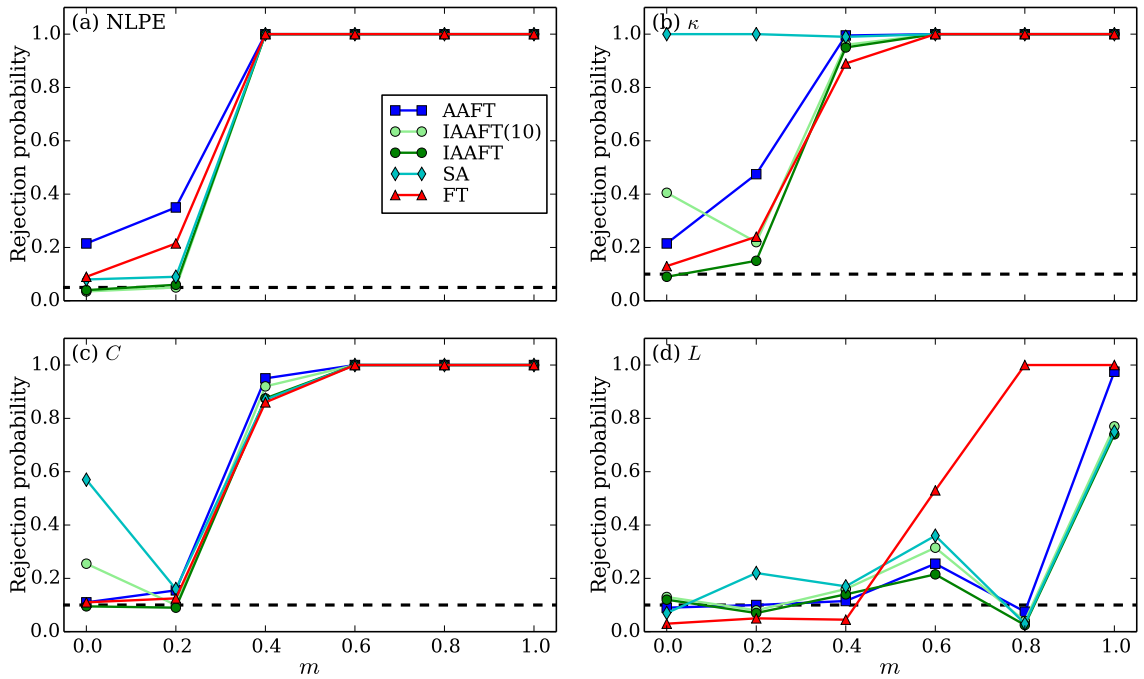
Figure 2.4: Discrimination power of the measures for nonlinearity for the mixed Lorenz data versus mixing parameter m. The rejection probability of 200 tests based on the NLPE (a) and network measures average connectivity κ (b), average clustering coefficient C (c) and average path length L (d). The size of the test is plotted as a dashed line. For the IAAFT algorithm, the surrogates were saved after 10 iterations [IAAFT(10)], before continuing until further iterations lead to no change in the surrogate (IAAFT). The SA algorithm stopped when the cost function was decreased to $E[C(\tau)] = 0.01$. Only 100 tests were made with the SA surrogates due to the numerical expense of the iteration scheme.

Figure 2.4(b) shows the discrimination power of the test based on the average connectivity κ of the network. The performance of most surrogate algorithms is similar to the NLPE. Stopping the IAAFT algorithm after ten iterations yields a relatively high false rejection probability of 40% at $m = 0$, which even decreases for increasing m. Iterating the IAAFT scheme until no change in the surrogates is obtained cancels this effect, the rejection probability for $m = 0.2$ is then quite small as in the case of the NLPE. This dependence of the rejection probability on the number of iterations shows that locally the surrogates change significantly even after ten iterations. The AAFT and FT algorithms show a comparable performance, with the former showing a higher rejection probability at $m = 0.2$. The SA surrogates fail to detect the linearity of the AR process.

In Fig. 2.4(c), the average clustering coefficient C yields a test with a slightly weaker performance compared to the NLPE. The surrogate generating algorithms AAFT, IAAFT and FT show similar performance. As in the case of the test based on κ, the IAAFT algorithm shows an increased false rejection probability at $m = 0$ if stopped after ten iterations. The SA algorithm also shows this increased rejection at $m = 0$. At $m = 0.2$, the rejection probability is still near 10% for all algorithms.

Finally, in Fig. 2.4(d), the average path length L is considered. The test shows a relatively weak performance with the first significant rejection of the FT algorithm at $m = 0.60$. The AAFT and IAAFT surrogates reject the hypothesis only when the data are almost pure Lorenz system. This global network measure fails to detect weak nonlinearities in the mixed Lorenz data.

2.3.3 Discussion

To explain the weak performance of the test based on the average path length L, it can be argued that the global properties of a (embedded) time series are not significantly different from its surrogate. Indeed, when κ_{or} is further decreased, the networks are very sparse and the average path length, taking into account only edges of the same component, becomes a more local measure. Accordingly, the performance of the test based on L improves. The network measures for a range of κ_{or} from 0.005 to 0.040 are shown in Fig. A.1 in the Appendix. Due to the weak performance of the test based on average path length L for detecting nonlinearities in the Lorenz system, and its strong dependence on κ_{or}, this network measure will not be considered for the application to real data in Sec. 2.4.

In Fig. A.1 it can also be seen that the network measures based on κ and C yield very similar results for the values of κ_{or} considered in the Appendix. The rejection probability for the test based on κ decreases slightly with increasing κ_{or}, while the performance of the test based on C increases. This can be understood as for small κ_{or}, the local structure is resolved very well, leading to the best performance of the local network measure κ. The clustering coefficient C, being a network measure on intermediate scales of the network, requires greater values of κ_{or}. In total, tests based on κ and C have a good performance comparable to the NLPE.

Calculating the rejection probability of the NLPE over the whole range of the mixing parameter shows that the main argument of Schreiber and Schmitz [1996] has to be reconsidered: While it is true that the IAAFT algorithm shows a smaller false rejection probability for linear time series, the rejection probability is smaller than that of other surrogate generating algorithms when weak nonlinearities are present. The mixing proposed here thus allows a mode detailed study of nonlinear time series, and enables a more profound comparison of different tests for nonlinearity.

The SA surrogates yield a relatively good performance for the NLPE, but they fail for the network test based on connectivity. The reason may be that the cost function has to be reduced further in order to mimic well the dynamics on the smallest scales. For a deeper insight, dedicated studies on the SA surrogates are necessary.

2.4   Analysis of Active Galactic Nuclei

Active galactic nuclei (AGN) light curves are analyzed as an example of real-world data. An AGN is a luminous region at the very center of a galaxy. It is powered by accretion onto a supermassive black hole. Ionization and recombination processes lead to X-ray emission from regions near to the event horizon of the black hole. The accretion disc contributes a continuum emission in the X-ray band that can be described by a power-law. Perpendicular to the accretion disc are radio jets. There are two classes of AGN, radio-loud AGN where the jet contributes to the emission, and radio-quiet AGN where this contribution can be neglected. Seyfert galaxies discussed here are a subclass of radio-quiet AGN which are again subdivided in different types ranging from 1 to 2, depending on the observed line widths. It was argued that the variety of AGN is partly the result of different aspect angles [Antonucci, 1993].

A spectral analysis is the typical approach to examine the observed X-ray emission. The observation of the gravitationally broadened iron Kα line was the first confirmation of a black hole in an AGN [Tanaka et al., 1995]. Spectral properties also permit an estimate of black hole spin and inner disc radii [Miller, 2007].

The contributions of a test for nonlinearity are twofold. First, linear models like global disk oscillation models [Titarchuk and Osherovich, 2000] can be rejected with the detection of nonlinearities. Second, the analysis of the (nonlinear) dynamics can put to test different nonlinear models. Fore example, it was argued that AGN are galactic black hole binary systems with their masses scaled up [McHardy et al., 2006]. Black hole binaries can be in the state of quasi periodic oscillation (QPO) [Remillard and McClintock, 2006]. One evidence of the relation between galactic black hole systems and AGN would be the detection and characterization of QPOs in AGN data [Antonucci, 1993]. A unified model for these objects of very different masses[4] and time scales would contribute greatly to the understanding of the physical processes

---

[4]The masses of the black holes of AGN are in the range of $10^6 - 10^9$ solar masses, as opposed to around 10 solar masses for galactic black holes.
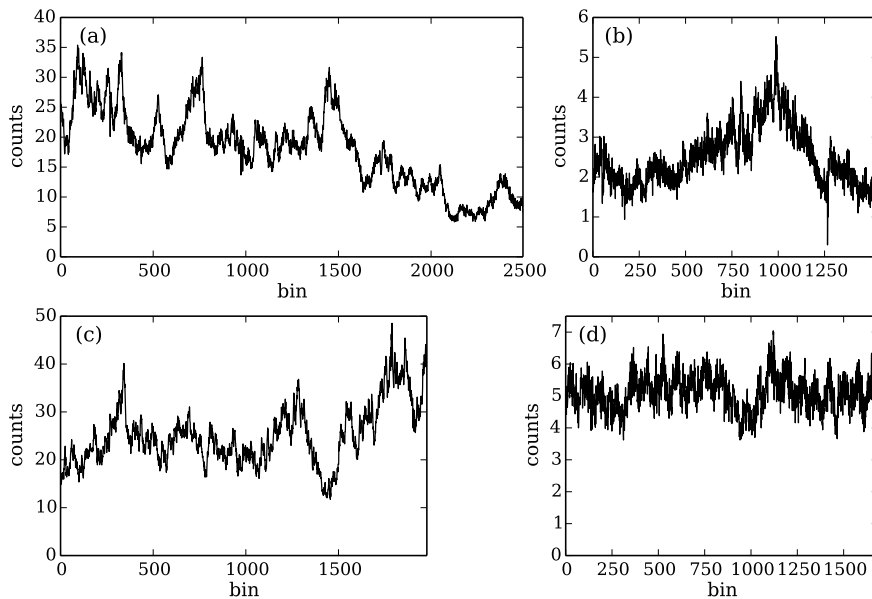
Figure 2.5: AGN light curves. (a) MCG. (b) Mrk. (c) Akn. (d) RE J. The data is binned with a bin size of 50 s. Since the characteristic times for the analysis for nonlinearities are given in time steps (see Sec. 2.4.2), the x axis labels of the plots show the bin number instead of time. The durations of the time series vary from about 80 ks (Mrk) to 120 ks (MCG).

close to a black hole.

2.4.1   Data Sets

The first time series considered here is a measurement of the Seyfert galaxy MCG-6-30-15 (MCG) taken in 2001 by the *X-ray Multi-Mirror Mission* (XMM-Newton) satellite [McHardy et al., 2005]. The time series is plotted in Fig. 2.5(a). With a duration of more than 120 ks it is the longest time series considered here.

The bright Narrow Line Seyfert 1 galaxy Mrk766 (Mrk) has been observed by all main X-ray observatories. Here, a measurement taken by the XMM-Newton satellite in 2005 is used. The light curve is shown in Fig. 2.5(b). Power density spectra have been analyzed in Markowitz et al. [2007], and a test for nonlinearity has previously shown a very significant outcome [Räth et al., 2012]. The significance of the detection of a QPO in another measurement on Mrk766 [Boller et al., 2001] was questioned by Benlloch et al. [2001].

The *Advanced Satellite for Cosmology and Astrophysics* (ASCA) observation of Ark 564 (Ark) in the X-ray band is also examined [Edelson et al., 2002]. As can be seen in 2.5(c), starting and ending point of this time series have very different values.

Finally, the time series of RE J1034+396 (RE J), taken by the XMM-Newton satellite in 2007, shows an oscillation with a periodicity of about one hour [Gierliński et al., 2008]. It is to date the only AGN light curve that shows periodic oscillations clearly visible by eye.
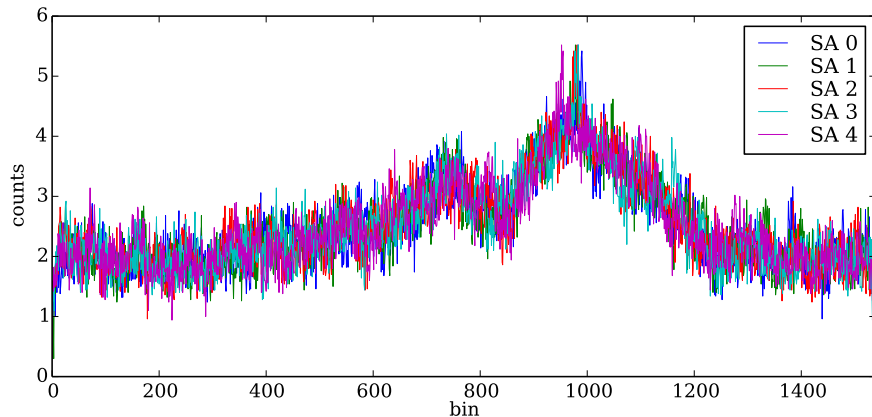
Figure 2.6: The first five realizations of SA surrogates for the Mrk data shown in one plot. The algorithm stopped when the cost function was decreased to $E[C(\tau)] = 0.02$.

2.4.2 Finding Nonlinearities in Active Galactic Nuclei Light Curves

As can be seen in Fig. 2.6, the SA algorithm tends to exactly reproduce the original time series of the Mrk, and cannot be used for a test for nonlinearity. Reproducing both the autocorrelation (with the Metropolis steps) and the amplitude spectrum (by starting the algorithm with a shuffle of the original time series) essentially leads to a copy of the original time series.

The embedding dimension of $d = 3$ is kept for the analysis of the AGN data. Since the optimal delay time cannot be detected by optical inspection of the embedded time series, ranges covering both first minimum and first zero crossing of the autocorrelation function were chosen. RE J with its relatively fast oscillations has the smallest range, starting at $\tau = 50$ time steps. The delay times of MCG, being long and with relatively slow dynamics, range from $\tau = 250$ to $\tau = 400$ time steps. The delay times can be read off the x axis of the plots of Fig. 2.7. The lead time of the NLPE is set to $T = 5$ time steps.

The same delay times as for the NLPE are used for the test based on the network measures. $\kappa_{\mathrm{or}} = 0.01$ as in Sec. 2.3, with further tests for different values in the Appendix. The IAAFT algorithms gave almost identical results if stopped after ten iterations or iterated until the end for all tests.

The significances of the tests of the MCG light curve are shown in the top row of Fig. 2.7. The NLPE shows significant nonlinearities of more than 5σ for the FT and AAFT surrogates[5]. The network test based on κ yields a higher significance for the AAFT surrogates, and a lower significance for the FT surrogates as compared to the NLPE. The test based on C shows results comparable to the NLPE. The

---

[5]Note that even though the significance as defined in 2.4 is a dimensionless measure, it is common practice to repeat the σ.
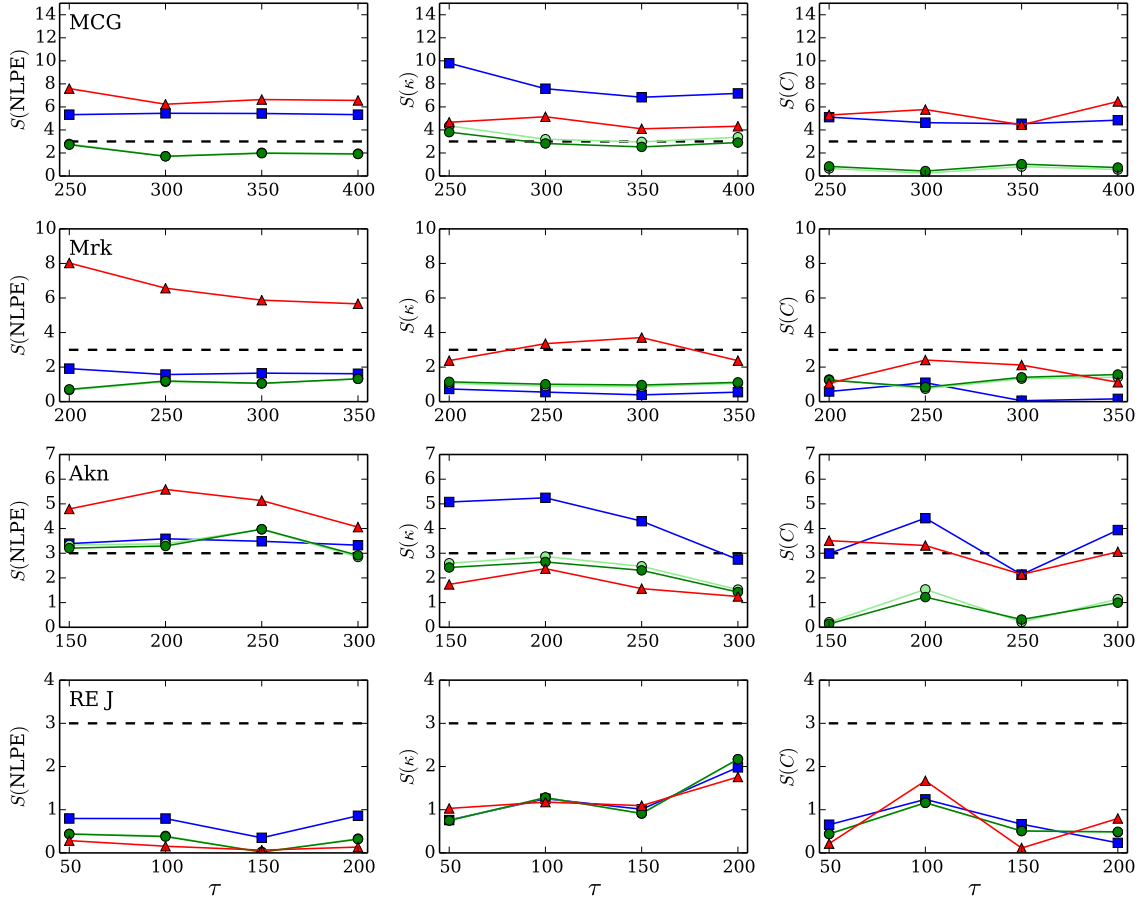
Figure 2.7: Significances of the NLPE (first column), connectivity (second column) and clustering coefficient (third column) versus delay time τ of the embedding. First row: MCG, second row: Mrk, third row: Akn, fourth row: RE J. The significances were calculated for 200 realizations of the surrogates and are colored as in Fig. 2.4. The 3σ detection limit is plotted as a dashed line.

significances of the IAAFT surrogates are below the detection limit of 3σ for all three measures for nonlinearity.

As for the Mrk, the significances of the FT surrogates vary from $\simeq 7\sigma$ for the NLPE to near and below 3σ for the network measures (see Fig. 2.7, second row). The significances of AAFT and IAAFT algorithms are below the detection limit for all tests considered here. Supposing that the significance of the NLPE for the FT surrogates is not an error, this suggests that both AAFT and IAAFT surrogates contain nonlinearities that prevent the detection of nonlinearities in the original data.

The analysis of the Akn data with the NLPE yields significant values of about 5σ for the FT surrogates, while the significances of the network measures are below (test based on κ) or near (based on C) the detection limit. The AAFT surrogates show an opposite behavior: Slightly above the detection limit for the NLPE, they reach significances of 5σ and 4σ for the network measures. Finally, the IAAFT algorithm
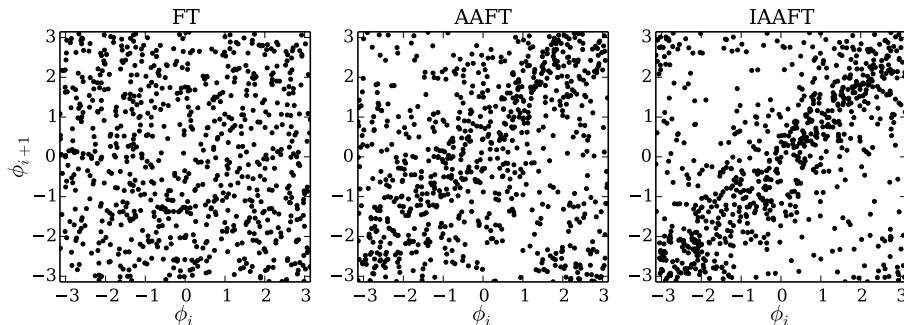
Figure 2.8: Phase maps with shift $\Delta = 1$ for surrogates of the Mrk data. For each surrogate generating algorithm, the surrogate with the highest cross correlation c_{PM} (see Eq. 2.10) out of 20 realizations is chosen. The values of c_{PM} are FT: 0.06, AAFT: 0.25, IAAFT: 0.43.

yields low significances.

The analysis of the RE J time series yields no significant values for any of the tests for nonlinearity (see Fig. 2.7, bottom row). This case may be compared to $m = 0$ for the Lorenz system, where the results of the NLPE were confirmed by the network measures.

The performance of the network tests for different κ_{or} for MCG, Mrk and Akn are shown in Appendix A. The significances for the RE J were below the detection limit for all values of κ_{or} considered here and are not shown. The significances do not change dramatically in the range of κ_{or} considered here, as did the rejection probability for the mixed Lorenz data in Sec. 2.3. The significances of the test based on κ decrease with increasing connectivity κ_{or} of the networks, which also agrees with the case of the Lorenz data. For the test based on C, the significance may decrease (Akn data) or increase (FT surrogates of Mrk data).

2.4.3 Phase Correlation in the Surrogates

To further investigate the differences of the surrogate generating algorithms, so-called phase maps are considered. The phases of the Fourier modes $\{\phi_i\}$ are plotted versus the phases $\{\phi_{i+\Delta}\}$ of the modes that where shifted by Δ. For a linear time series, the $\{\phi_n\}$ are independent, and thus scattered uniformly in the square bounded by $\pm\pi$. Structure in a phase map shows that the $\{\phi_n\}$ are not independent, which means that the phases contain information about the time series. Phase maps are a readily implemented test for phase information, however, they do not necessarily show a significant structure for nonlinear data. One way to quantify the correlation in the phase maps is to calculate the cross correlation between $\{\phi_i\}$ and $\{\phi_{i+\Delta}\}$:

$$c_{PM}(\Delta) = \frac{\langle \phi_i \phi_{i+\Delta} \rangle}{\sigma(\{\phi_i\})\sigma(\{\phi_{i+\Delta}\})}. \tag{2.10}$$

In Fig. 2.8, phase maps of the three surrogate generating algorithms FT, AAFT and IAAFT are shown. Since the SA surrogates have not been used in Sec. 2.4.2 due
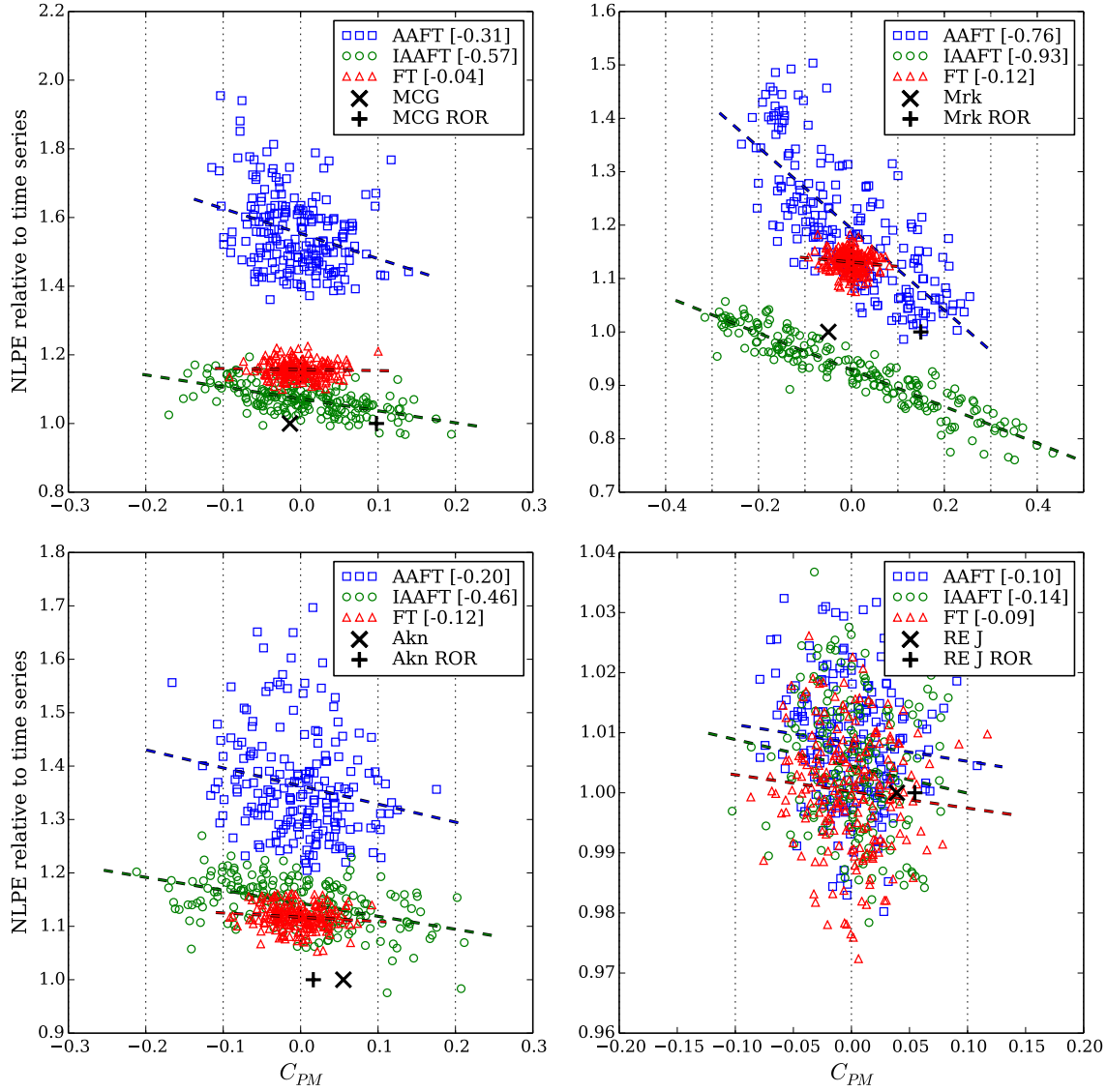
Figure 2.9: NLPE versus $c_{\mathrm{PM}}(\Delta = 1)$ for 200 surrogates of the AGN data. The NLPE is calculated for the following delay times: MCG: $\tau = 300$, Mrk: $\tau = 250$, Akn: $\tau = 200$, RE J: $\tau = 100$. The NLPE is divided by the value of the NLPE of the corresponding time series which are plotted as an x mark (original time series) and a plus sign (rank-order remapped time series). The values of the cross correlation between NLPE and c_{PM} are shown in square brackets in the legend. Root-mean-squared fits to the surrogates are shown as dashed lines to guide the eye.
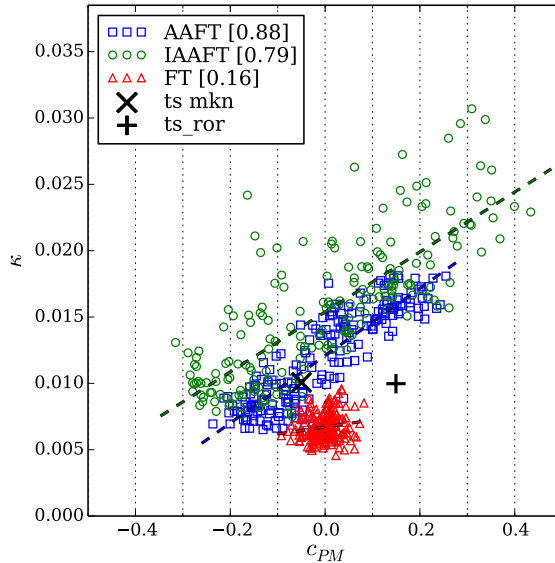
Figure 2.10: Average connectivity κ versus $c_{PM}(\Delta = 1)$ for 200 realizations of AAFT, IAAFT and FT surrogates for the Mrk. Delay time and legend are the same as in Fig. 2.9.

to their resemblance to the original time series, their phases will not be analyzed further. The phase maps of the SA surrogates showed no phase correlations in a first analysis. The FT scheme, which stops after randomizing the phases of the Fourier modes, is guaranteed to contain no phase information. Diagonal features are evident for the AAFT and IAAFT algorithms. The rank-ordering (AAFT) and the iteration scheme (IAAFT) thus reintroduce phase correlations to the surrogates which only at the beginning of the algorithms were truly linear. As was shown for AGN and financial market data in Räth et al. [2012], these nonlinearities lead to a non-detection of nonlinearities in the time series.

In order to examine the impact of the phase correlations on the measures for nonlineartiy, the NLPE of the surrogates of the AGN data is plotted versus the cross correlation c_{PM} of the phase maps in Fig. 2.9. There is a significant anticorrelation for AAFT and IAAFT surrogates for the three light curves where nonlinearities were detected in Sec. 2.4.2. In Fig. 2.10 the average connectivity of the corresponding networks is also plotted versus the cross correlation for surrogates of the Mrk data. The correlation in this case are as strong as the anticorrelation for the NLPE.

2.4.4   Discussion

Significant nonlinearities were found for the MCG, Mrk and Akn data. For the periodic light curve of RE J, however, no nonlinearities could be detected. It is thus still an open question if a QPO state of a galactic black hole corresponds to the variability observed for RE J [Middleton and Done, 2010].

The network measures proposed here perform as well as or slightly worse than the NLPE. This suggests once again that the NLPE is a very robust measure for

nonlinearity with good performance. Dynamic measures (see Sec. 2.1) seem to be more appropriate than the static measures, as could be expected for an analysis of the *dynamic* properties of a system. The network measures might thus rather be compared to other static measures like the correlation dimension. Still, the network analysis of nonlinear time series allows a deeper insight in the differences of the surrogate generating algorithms.

The FT surrogates yield the highest significance for the NLPE for all light curves considered here. They are, however, less significant than the AAFT surrogates for the network measures of MCG and Akn. One reason may be the strong impact on the structure of the time series of the rank-ordering to a Gaussian distribution.

The reproduction of surrogates almost identical to the original time series illustrates the weakness of the SA algorithm in a striking manner. It is impossible to make sure that no nonlinearities are induced during the iteration scheme. For unevenly sampled time series, a prime example for the application of SA surrogates, it seems more promising to find a basis that allows the generation of FT surrogates. Rossmanith et al. [2012] examined the impact of the Galactic plane on the detection of non-Gaussianities in the cosmic microwave background by using a set of orthogonal basis functions on the cut sky[6].

The analysis of the AGN data with the NLPE and network measures suggests that the IAAFT scheme is not useful when examining time series with rather long correlation times. The phase correlations introduced by the iteration scheme were detected in the phase maps and clearly show that there are induced nonlinearities in the surrogates. The fact that the structures found in the phase maps (anti)correlate with the measures for nonlinearity shows that these nonlinearities in the surrogates have an impact on the significance of the tests for nonlinearity. IAAFT surrogates also yielded a weak performance in Sec. 2.3, where, however, no obvious phase correlations were found in the phase maps.

Correlations in the phase maps have also been found for the AAFT surrogates. The AAFT surrogates show the highest significances exactly in the cases where the correlation between the NLPE and c_{PM} are not too strong (MCG: -0.31 and Akn: -0.20 as compared to Mrk: -0.76, see Fig. 2.9). In these two cases the significance is even higher for the test based on network measure κ. In the case of the Mrk data, however, the NLPE and the network measures strongly (anti)correlate with c_{PM} and yield significances below the detection limit.

The influence of the Fourier phases on properties of data in real-space should be examined further. Such studies are beyond the scope of this work which will continue with the application of network analysis to spatial structures, namely complex plasma crystals.

---

[6]The variations of the cosmic microwave background are much smaller than signals measured from our galaxy. Therefore, one can subtract what is thought of being part of this foreground, or cut the whole area influenced by Galactic signals

Chapter 3

Analysis of Complex Structures

3.1 Motivation

The results of this chapter have previously been published in Laut et al. [2014].
Complex (or dusty) plasmas are ionized gases containing micron-sized dust particles.
The particles are in general negatively charged since the electrons of the plasma have
a higher mobility than the ions, and can arrange in regular patterns.

Complex plasmas exist in various forms, from small clusters to large extended
systems [Thomas et al., 1994; Fortov et al., 2005; Ivlev et al., 2012]. Two- and
three-dimensional clusters are popular objects to study for two main reasons. First,
they are used as model systems to study generic phenomena such as self-organization
and transport, at the level of individual particles [Arp et al., 2004; Totsuji et al., 2005;
Melzer et al., 2010; Hyde et al., 2013]. Second, clusters can be used for diagnostic
purposes, e. g. probing the plasma parameters at the position of particles [Arp et al.,
2005; Carstensen et al., 2010; Nosenko et al., 2009].

The dust particles constituting a complex plasma cluster are highly charged
and therefore have to be externally confined. For instance, in Arp et al. [2005] the
particles were suspended in a short open glass tube placed on top of the rf electrode
in a gas discharge. The authors suggested that the nearly isotropic particle cluster
confinement was mainly determined by gravitational, electric and thermophoretic
forces. A contribution from the ion drag force due to streaming ions was not observed.
Dust particles suspended in a glass box can exhibit various structures, such as isolated
single linear chains [Kong et al., 2011], vertical strings [Wörner et al., 2012], zigzag
structures [Melzer, 2006], helical structures [Tsytovich et al., 2007], and Coulomb
clusters with onionlike shells [Arp et al., 2005; Totsuji et al., 2005].

Here, complex plasma clusters observed by Wörner et al. [2012] are analyzed using
complex networks. By considering the whole adjacency matrix with information
stored for each particle, this approach may remain applicable where many other
tools relying on some kind of averaging fail because of too weak statistics. The
experimental setup is explained in Sec. 3.2. In order to show the advantage of
network analysis over more straightforward methods, vertical strings of the clusters

are revisited in Sec. 3.3. They were identified in Wörner et al. [2012] by introducing a certain fixed threshold to their transverse extent. This simple approach has its evident advantages but also limitations such as erroneously including passing by particles and a somewhat arbitrary threshold. With the help of multislice networks [Mucha et al., 2010], strings are found and resolved throughout the whole time series in a natural way.

As a second application of networks to complex plasma clusters, the global structure of the clusters is analyzed in Sec. 3.4. In Wörner et al. [2012], the clusters were found to be of spheroidal shape, yet a competing cylindrical symmetry was also present due to streaming ions which create an anisotropy in the interparticle potential. Since there was no appropriate analysis method, the contributions of the competing symmetries in the clusters was not further analyzed in Wörner et al. [2012]. By creating networks from the structural properties of the clusters, the analysis now reveals an interplay between these two symmetries. Spherical and cylindrical ordering of the particles is examined by comparing network measures of the experimental data with null models. These findings are compared with an estimate of the radial confinement. To this end, a dynamical force balance is considered.

3.2   Driven Complex Plasma Clusters

The clusters of micron-size particles were suspended in a glass box mounted on top of the rf electrode in a capacitively coupled discharge in argon (see Fig. 3.1). Sinusoidal voltages were individually applied to the sides of the box, which results in an electric field at the position of the particles that has a constant magnitude and rotates in the horizontal plane. This "rotating wall" technique [Nosenko et al., 2009] can be used to manipulate complex plasma clusters. Depending on the frequency of the applied electric field, the clusters rotated in the direction of the electric field (albeit with much lower frequency) or remained stationary. Here, a rotating cluster driven at 5 kHz and a stationary cluster driven at 1 kHz will be analyzed.

The particle coordinates were measured using a 3D imaging method, namely the stereoscopic digital in-line holography [Kroll et al., 2008]. In two identical channels, expanded laser beams illuminated the particle cluster from two perpendicular directions. The diffracted light was registered directly by two CCD cameras operating at 50 frames per second over a time interval of 10 s. In the resulting images, each particle is represented by a system of concentric circles [see Figs. 3.1(b) and (c)]. The depth information is encoded in the intercircle spacing.

Whether rotating or not, the clusters are significantly compressed in the radial direction by the applied field. The specific mechanism of this compression is not clear. The ponderomotive force [Lamb and Morales, 1983], a naturally expected candidate, is of minor importance in this experiment. Indeed, the clusters driven at the frequencies of 1 kHz and 5 kHz are of the same size (see Fig. 3.8), whereas the ponderomotive force has an inverse quadratic dependence on frequency.
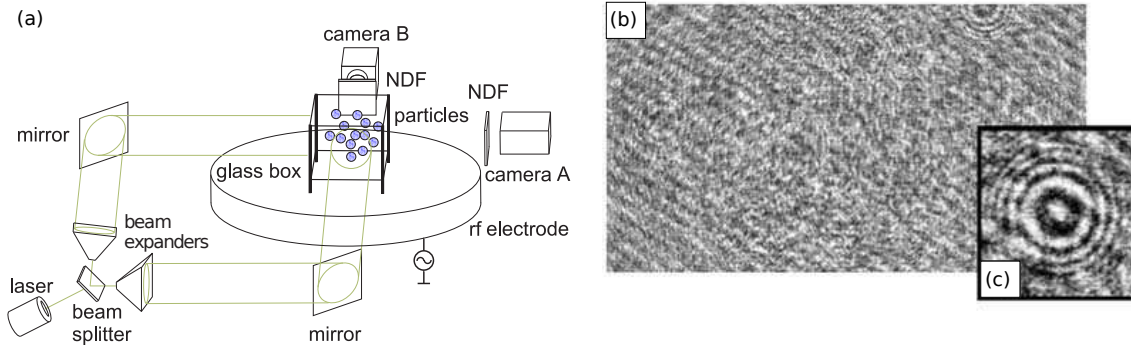
Figure 3.1: (a) Experimental setup. A 3D cluster of micron-size particles is suspended in the glass box mounted on top of the rf electrode in a capacitively coupled discharge. The particles are charged negatively and are manipulated by applying voltages on the conductive side plates of the box. By shifting sinusoidal signals on the adjacent plates by $\pi/2$, a rotating electric field is created at the position of particles. Expanded laser beams at $532\,\mathrm{nm}$ illuminate the particles from two perpendicular directions. The diffracted light is attenuated by neutral density filters (NDF) and registered by CCD cameras A and B. The interference patterns of the whole cluster (b) and of an individual dust particle (c) are also shown.

3.3 Particle Strings in the Clusters

3.3.1 Detecting Particle Strings with Community Finding Algorithms

As was motivated in Sec. 1.3, community detection in networks can be used to find functional units in complex structures. If the strength of a node is inversely proportional to the horizontal distance between two particles, the functional units at an appropriate scale are the vertical strings observed in the cluster. Weighting the nodes by this distance leads to a connected network where each node is connected to all other nodes as opposed to the networks in Chapter 2. Persistent strings can be found with the help of multislice networks. By considering the whole measurement over $10\,\mathrm{s}$, the dynamics of the strings of the cluster will be analyzed.

The network *slice* A_{ijt} is defined to be inversely proportional to the horizontal distance between particles i and j at time t,

$$A_{ijt} = \begin{cases} d_{\min}/d_{ij} & \text{for } i \neq j \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

Here, $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ is the horizontal distance between two particles at time t and $d_{\min}$ the minimal horizontal distance in the time series such that A_{ijt} is at most one. The connection between two particles is thus stronger, the nearer they are to each other. The resulting multislice network contains one *slice* A_{ijt} for each frame t of the experimental data.
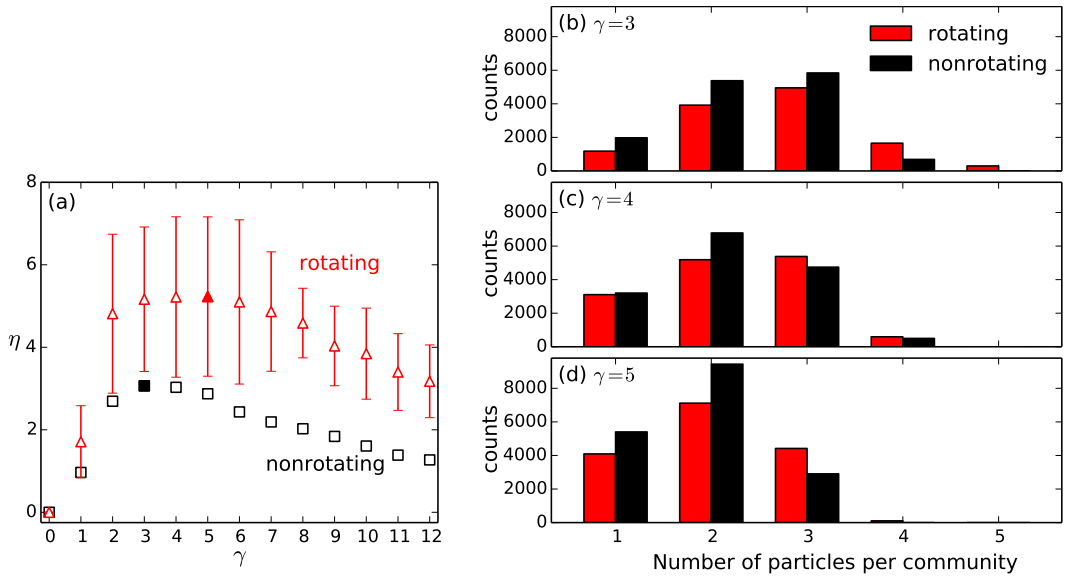
Figure 3.2: Dependency of detected communities on resolution parameter γ. (a) Ratio η of intercommunity distance to community size for different values of γ for the rotating (triangles) and nonrotating (squares) cluster. The maximum values are plotted as solid symbols. The standard deviations are plotted as error bars and are concealed by the symbols in the case of the nonrotating cluster. (b)–(d) Histograms of the number of particles per community for different values of γ. The absolute counts including all particles and time steps are shown.

In order to resolve the communities in time, each particle is connected with itself at the consecutive time step. Varying the strength of this connection changes the susceptibility of a particle to change the community. Here, the connection is set to $C_{jtr} = 0.1\delta_{t,r+1}$ for all particles. The cummunity finding algorithm then calculates the partition of the network that maximizes $Q_{\mathrm{multislice}}(\gamma)$ (see Eq. 1.14).

The resolution parameter γ is chosen as follows. It is considered optimal when the distances between the communities of a partition are largest compared to the extent of the communities, i. e., where

$$\eta = \langle d_{i,\mathrm{next}}/d_{i,\mathrm{same}} \rangle \tag{3.2}$$

is maximal. Here $d_{i,\mathrm{next}}$ is the horizontal distance of particle i to the next particle of another community and $d_{i,\mathrm{same}}$ is the average horizontal distance to the particles in the same community if the particle is in a community with more than one particle and infinity otherwise. The brackets denote the average over all particles and time steps.

As can be seen in Fig. 3.2(a), η varies weakly in the range of resolution parameter γ from 2 to 6 around the maxima for the clockwise rotating and nonrotating clusters. The value of η is significantly larger for the rotating cluster where the strings are more pronounced. The distribution of the number of particles per community is very sensitive to γ [see Figs. 3.2(b)–(d)]. The algorithm finds more small communities of
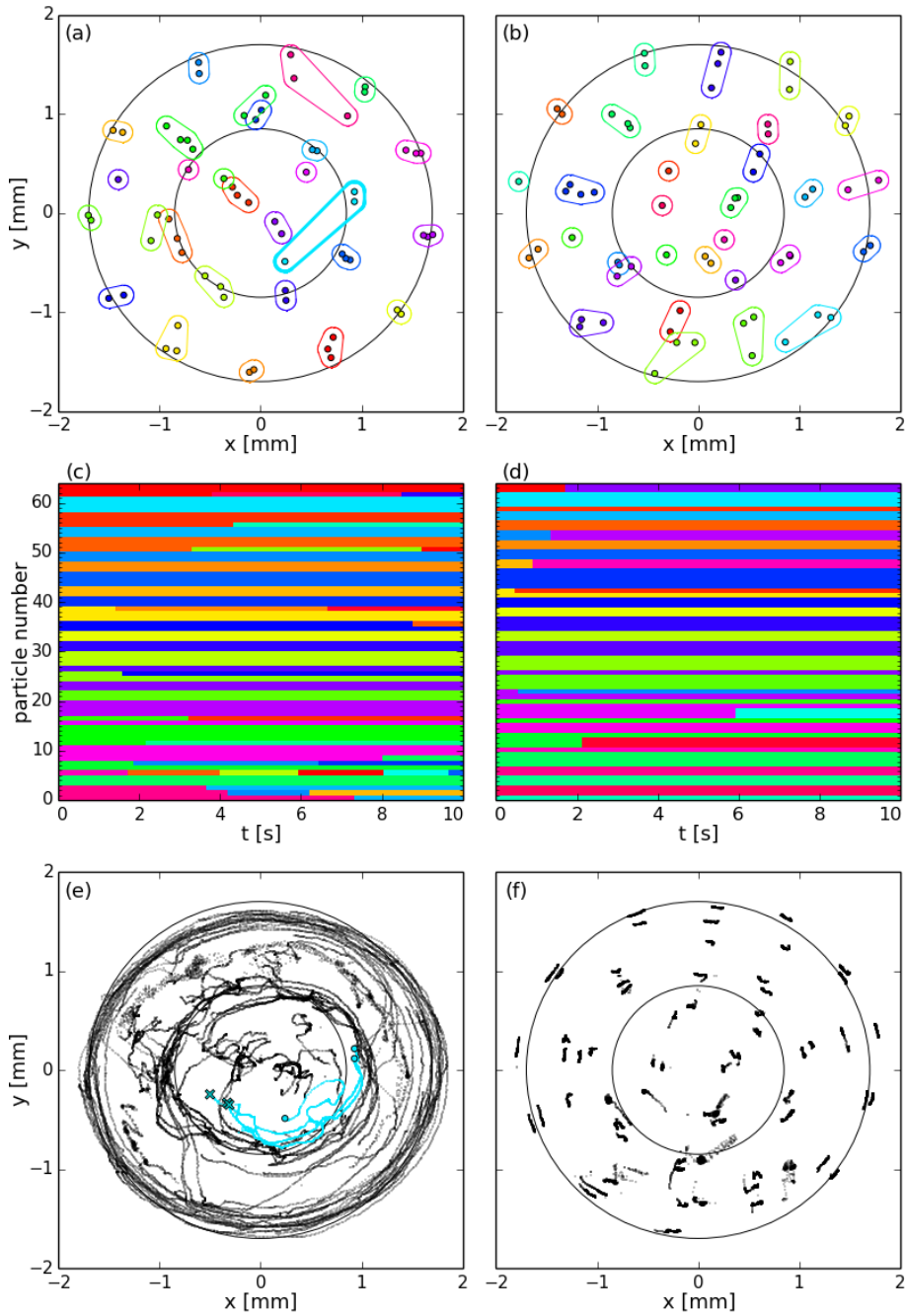
Figure 3.3: Time evolution of detected strings for resolution parameter $\gamma = 4$. Projection of the rotating (a) and nonrotating (b) cluster on the xy plane at $t = 0$. Particles which were found to be in the same string are grouped together and have the same color. An initially elongated community is highlighted for the rotating cluster. (c), (d): The string affiliation of particles over time for the rotating and nonrotating cluster. A transition from one string to another appears as a change in color of the line corresponding to the particle. (e), (f): Projection of the clusters on the xy plane with all frames superimposed. The highlighted community from (a) is plotted in blue with the positions of the particles at $t = 0$ and $t = 10\,\mathrm{s}$ represented by circles and crosses, respectively.

one or two particles in the case of the nonrotating cluster, while larger communities with at least three particles are more likely for the rotating cluster.

By considering the communities as the vertical particle strings in the cluster, one thus obtains persistent strings of different size depending on γ, in contrast to a more traditional approach of grouping particles whose distance is below a certain value as in Wörner et al. [2012]. The strings for $\gamma = 4$ and their evolution in time for the rotating and nonrotating clusters are shown in Fig. 3.3. It is evident from the plots of the string affiliations over time in Figs. 3.3(b) and (e) that in the case of the nonrotating cluster there are fewer transitions between the strings.

3.3.2 Discussion

The strings are quite robust against encounters with only occasionally passing by particles: During their passage, these roaming particles are not considered to be part of the string. The cases where the horizontal distance between the particles is small, while a large vertical distance prohibits physical correlation, do not persist in time. These events are thus not considered as strings by the community-finding algorithm. Indeed, the mean vertical distance between particles of the same string is $0.79\,\mathrm{mm}$ (rotating cluster) and $0.86\,\mathrm{mm}$ (nonrotating cluster), and the events where this distance is larger than $1.2\,\mathrm{mm}$ are not frequent (less than $3\,\%$) for both clusters.

By inspection of Fig. 3.3(a) alone, it is not obvious why, for example, the particle near the center belongs to the highlighted longish community. In order to understand this, the whole time series has to be considered, as this roaming particle quickly joins the two remaining particles and forms a persistent string. In Fig. 3.3(c), the position of the particles at all time steps are plotted. The particles of the highlighted community are plotted in blue again, with their positions at $t = 0$ represented by circles, and the positions at $t = 10\,\mathrm{s}$ represented by crosses.

This special feature of community assignment in networks may be a powerful tool within a wide range of possible applications.

3.4 Global Structure of the Clusters

3.4.1 Examining Competing Symmetries with Networks

The global structure of the clusters is analyzed by means of unweighted networks. In order to examine the cluster geometry, the adjacency matrix was chosen to connect two particles whose difference in cylindrical radii is small enough. The procedure is sketched in Fig. 3.4: In (a), the 60th particle of the cluster is plotted in blue and the particles with comparable cylindrical radii in red. The latter can be read off the 60th column of the corresponding adjacency matrix, (b). A representation of the resulting network can be seen in Fig. 3.4(c). At this time step the network consists of two main components, which correspond to the two ring regions of the cluster, and various smaller components.
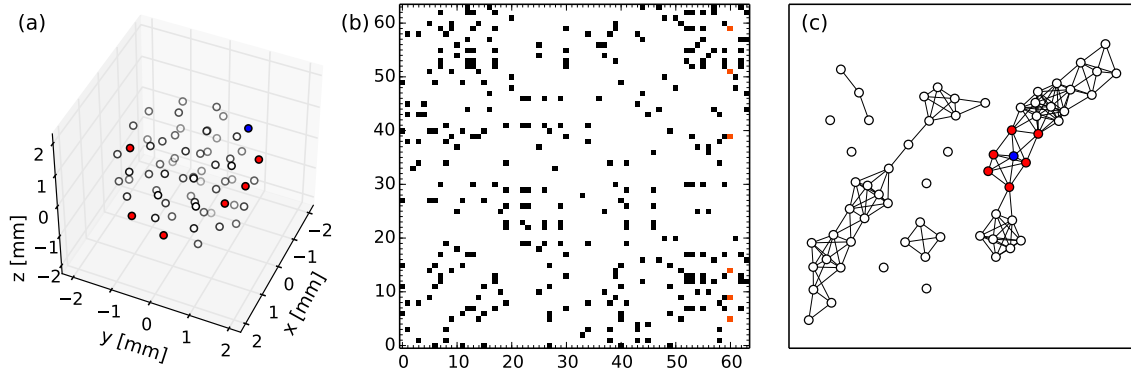
Figure 3.4: Method for generating an unweighted network from the data. (a) 3D plot of the clockwise rotating cluster. The 60th particle is plotted in blue and the particles j satisfying $|\rho_{60} - \rho_j| < \epsilon$ are plotted in red, they will be the neighbors in the corresponding network. As opposed to Fig. 1.1, neighbors are now not necessarily close to each other. (b) The corresponding adjacency matrix. The nodes connected to the 60th node can be read off the 60th column (or line) of the matrix. (c) Representation of the network. The nodes representing the particles have the same markers as in (a). The positions of the nodes is chosen such that the components of the graph are clearly visible

The network thus obtained is analyzed using the network measures defined in Sec. 1.2 and the results are compared to those of a network from a null model where a certain fraction of points is in perfectly spherical order, and the rest in cylindrical order. The ratio that shows the best agreement with the experimental data will be considered as the ratio of the competing spherical and cylindrical geometries of the cluster.

The adjacency matrix is defined similar to Eq. 1.1 as

$$A_{ij}^{\text{cyl}}(\epsilon) = \Theta\left(\epsilon - \|\rho_i - \rho_j\|\right) - \delta_{ij}, \tag{3.3}$$

where $\rho_i = \sqrt{x_i^2 + y_i^2}$ is now the cylindrical radius, and ϵ again an appropriate threshold. The threshold ϵ is chosen to be a fraction α of the mean difference in cylindrical radius: $\epsilon = \alpha\langle|\rho_i - \rho_j|\rangle$. The brackets denote the average over all particles and time steps. The fraction is set to $\alpha = 0.1$.

The null models are artificial structures with a predefined ratio $R = n^{\text{sph}}/n^{\text{cyl}}$ of the number of particles in a perfect spherical structure to the number of particles in a perfect cylindrical structure. The total number of particles $n = n^{\text{sph}} + n^{\text{cyl}}$ of the null models is equal to the number of particles in the experimental data. Two-shell null models are used with different ratios $R = 0, 1/3, 1/2, 1, 2, 3, \infty$. Each model is constructed as follows. The cylindrical structure consists of two concentric cylinders with the same (cylindrical) radii as the main components of the experimental data ($\rho_1 = 1.0\,\text{mm}$, $\rho_2 = 1.6\,\text{mm}$). For a given n^{cyl}, the ratio of the particle number in the inner cylinder to that in the outer cylinder is chosen to be equal to ρ_1/ρ_2. The two shells of the spherical structure have (spherical) radii $r_1 = 1.1\,\text{mm}$ and $r_2 = 1.7\,\text{mm}$. A uniformly distributed random noise of amplitude $0.15\,\text{mm}$ is added to the positions
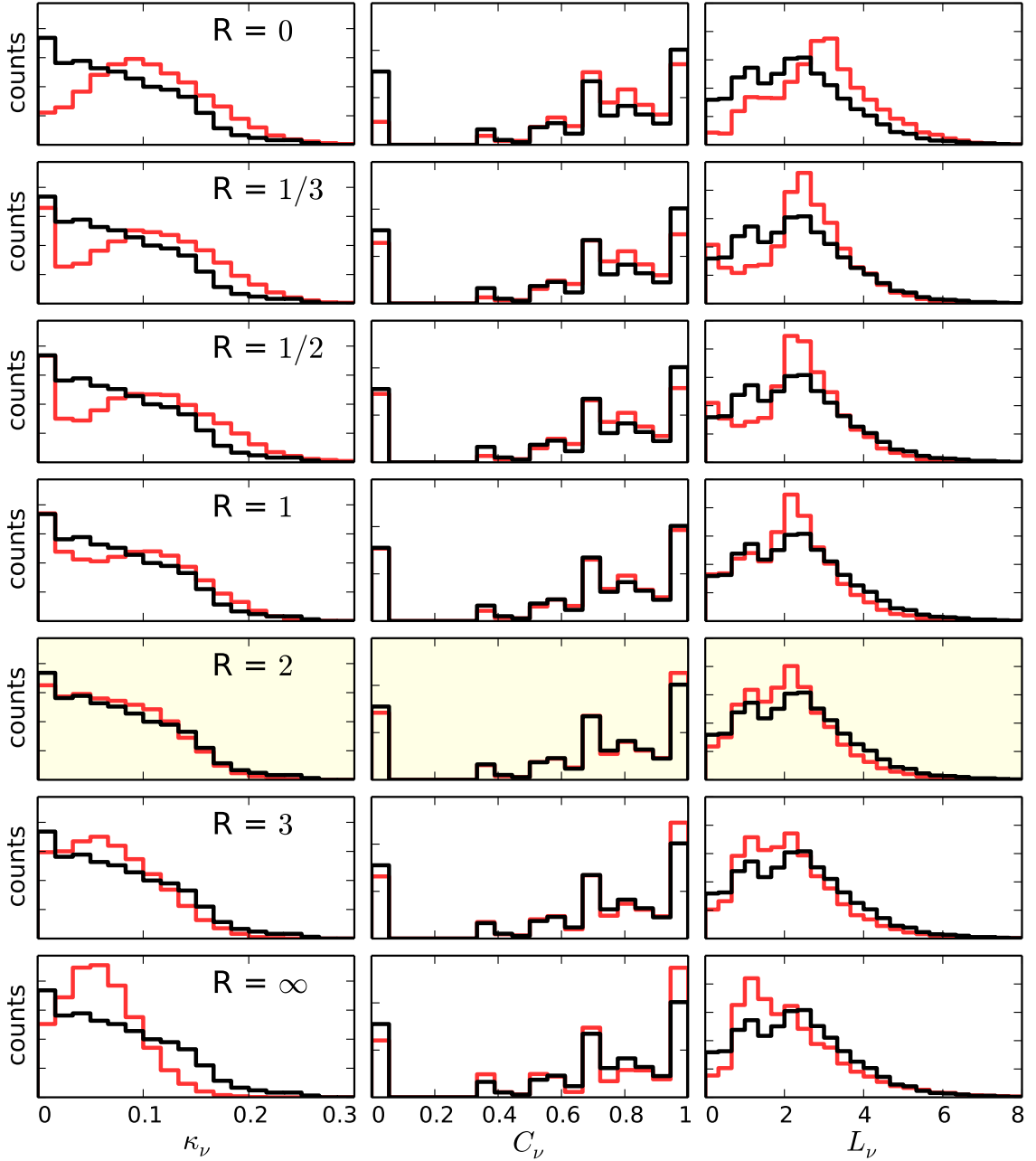
Figure 3.5: Comparison of network measures for the clockwise rotating cluster with the measures for null models of different ratios R. The histograms (for all particles and time steps) of measures on networks of the clockwise rotating cluster are plotted in black. They are identical in all rows. The null models are plotted in red with the ratio ranging from $R = 0$ (top row, purely cylindrical structure) to $R = \infty$ (bottom row, purely spherical structure), where R is the ratio of particles from a perfectly spherical geometry to particles from a cylindrical geometry. First column: connectivity κ_ν, second column: clustering coefficient C_ν, third column: path length L_ν. The row with the best agreement for all three network measures is highlighted with a pale yellow background. The good agreement of the measures in the fifth row is clearly visible. See Figs. B.1 and B.2 in the Appendix for comparisons for the counterclockwise rotating cluster and the nonrotating cluster.
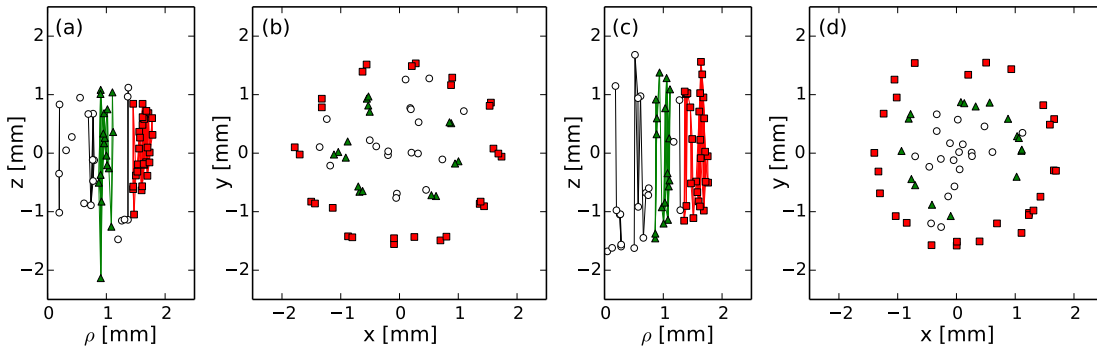
Figure 3.6: Comparison of experimental cluster to null model. Projections of the clockwise rotating cluster onto the ρz plane (a) and xy plane (b). The two main components of A^{cyl} are plotted as triangles (squares) and correspond to the inner (outer) ring regions of the cluster. For the sake of clarity the network edges are only plotted in panel (a). The network is identical to the one in Fig. 3.4(c), only the positions of the nodes are chosen differently: While in Fig. 3.4 a more abstract representation of the network was chosen to show its structure, here the positions correspond to the projections of the particles. (c), (d) The same for the null model whose network measures showed the best agreement with the cluster: The ratio of particles in a spherical arrangement to particles in a cylindrical arrangement is $R = 2$.

of all particles. See Figs. 3.6(c) and (d) for the projections of such a null model with $R = 2$ on the ρz and xy plane.

For each time step of the experimental data, the corresponding networks of the data and of a null model with a given ratio R are created and analyzed with the network measures degree centrality κ_ν, clustering coefficient C_ν and path length L_ν. As opposed to Sec. 2, the distributions and not averaged values of the measures will be considered. The histograms of these measures for all particles and all time steps of the data are compared, and the ratio R of the null model with the best agreement is considered to be the ratio of spherical to cylindrical order of the cluster.

In Fig. 3.5, the results for A^{cyl} of the clockwise rotating cluster are shown and compared to the null models. Given the good agreement for $R = 2$, one can argue that the spherical geometry of the cluster is two times more pronounced than the cylindrical geometry. The projections of the null model with $R = 2$ on the ρz and xy planes are plotted in Figs. 3.6(c) and (d) and compared to the clockwise rotating cluster [Figs. 3.6(a) and (b)].

In order to test the network approach, the analysis is repeated, this time comparing the spherical radii of the particles. The adjacency matrix now reads

$$A^{\mathrm{sph}}_{ij}(\epsilon) = \Theta\left(\epsilon - \|r_i - r_j\|\right) - \delta_{ij} \tag{3.4}$$

in the same notation as in Eq. 3.3 where r is the spherical radius and $\epsilon = \alpha\langle|r_i - r_j|\rangle$ with $\alpha = 0.1$. See Appendix B for comparisons of the network measures for A^{sph}.

Table 3.1: Ratios $R = n^{\mathrm{sph}}/n^{\mathrm{cyl}}$ of the number of particles in a perfect spherical structure to the number of particles in a perfect cylindrical structure, and anisotropic confinement parameters Ω_ρ and Ω_z of the clockwise and counterclockwise rotating, and nonrotating clusters. The values of R shown here provided the best agreement for network measures on the cylindrical (A^{cyl}) and spherical (A^{sph}) properties of the clusters. The results are averaged over 40 calculations, with the standard deviation indicated as error. The cylindrical (Ω_ρ) and vertical (Ω_z) confinement parameters [Eq. (3.8)] are calculated from the linear fitting parameters of the projections of the Yukawa forces (see Fig. 3.9).

	Cluster rotation		
Parameter	Clockwise	Counterclockwise	Nonrotating
R for A^{cyl}	2.0±0.2	2.0±0.2	3.0±0.2
R for A^{sph}	1.3±0.5	1.0±0.2	2.0±0.2
Ω_ρ [s^{-1}]	19±2	12±15	16.2±0.8
Ω_z [s^{-1}]	33±1	36±2	32.0±0.3

The results of the analysis of the rotating and nonrotating clusters for both adjacency matrices A_{ij}^{cyl} and A_{ij}^{sph} are summarized in Table 3.1. Even though the results for the cluster geometry are not the same for the different adjacency matrices, the general observation remains the same: One finds comparable results in the cases of clockwise and counterclockwise rotation, while the value of R is greater for the nonrotating cluster.

The dependence of the best ratio R on the noise amplitude of the null models is shown in Fig. 3.7. Noticeably, decreasing the noise amplitude generally shifts the optimal ratios R to higher values for A^{cyl} but towards lower values for A^{sph}. This can be understood as in the first case the difference in the cylindrical radii mostly determines whether nodes are connected or not. Hence, in order to find the best agreement, more particles of spherical structure have to be added when the noise level is reduced, increasing n^{sph} and thus increasing $R = n^{\mathrm{sph}}/n^{\mathrm{cyl}}$. In the second case of A^{sph}, decreasing the noise amplitude is to be compensated by adding more cylindrically structured particles, yielding a lower value for R.

Another possibility to examine the structure of a cluster is to merely plot the particle z positions vs. cylindrical radius ρ [Arp et al., 2004; Totsuji et al., 2005; Schella et al., 2013] (see Fig. 3.8). In this projection, spheres appear as semicircles and cylindrical structures as vertical lines. One can see that in the case of rotation [Figs. 3.8(a) and (b)] the particles at the bottom and the top of the cluster have moved outward, while the particles at $z \simeq 0$ appear to have moved inward compared to the nonrotating cluster, Fig. 3.8(c). The rotating clusters thus appear more cylindrical, while the spheres are more pronounced in the nonrotating case, but there is neither strong visual evidence nor a numerical result as in the case of network analysis.
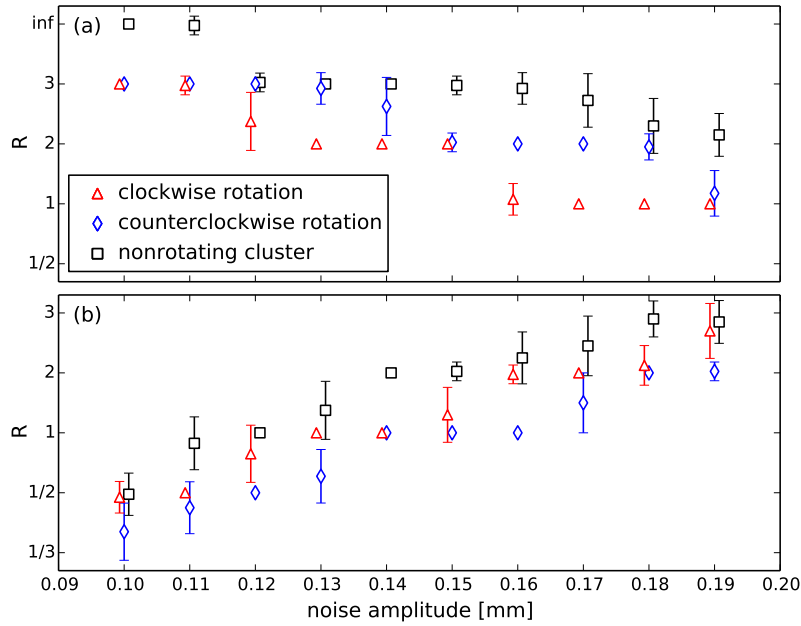
Figure 3.7: Ratio $R = n^{\mathrm{sph}}/n^{\mathrm{cyl}}$ vs. noise amplitude of null models for A^{cyl} (a) and A^{sph} (b). The results are averaged over 40 calculations, and the standard deviations are plotted as error bars. The results in Table 3.1 correspond to a noise amplitude of $0.15\,\mathrm{mm}$.
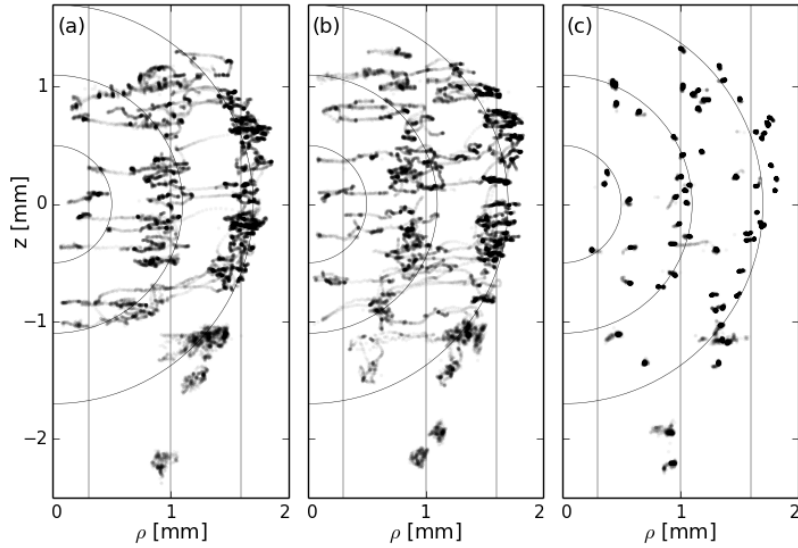

Figure 3.8: Projection of particle positions on the ρz plane for clockwise (a) and counterclockwise (b) rotating, and nonrotating (c) clusters. Vertical lines and semicircles are adjusted to the cylindrical structure of the rotating clusters and to the spherical structure of the nonrotating clusters, respectively, in order to guide the eye. Particle positions at each time step correspond to a transparent marker, leading to lines in the event of a string transition and solid structures if the particle stays in the same position throughout the time series.

3.4.2 Estimating Particle Confinement

The difference in structure between rotating and nonrotating clusters stems from the particulars of the mutual particle interactions and the forces confining the cluster. Even though it is a rather difficult task to explore the forces controlling the cluster dynamics in detail, it is instructive to assume that globally the system is in quasiequilibrium. The latter is determined by the balance of repulsion $\mathbf{F}^{\mathrm{rep}}$ via the Yukawa forces, neutral gas friction $\mathbf{F}^{\mathrm{fr}}$, inertial forces $\mathbf{F}^{\mathrm{in}}$, and confinement $\mathbf{F}^{\mathrm{conf}}$ provided by all other forces[1]:

$$\langle \mathbf{F}^{\mathrm{rep}} \rangle + \langle \mathbf{F}^{\mathrm{fr}} \rangle + \langle \mathbf{F}^{\mathrm{in}} \rangle + \langle \mathbf{F}^{\mathrm{conf}} \rangle = 0. \tag{3.5}$$

Here, the stochastic averaging is assumed to be performed.

The Yukawa forces, governing the mutual repulsion between the particles, can be directly calculated from the data:

$$\mathbf{F}_i^{\mathrm{rep}} = -\left(Ze\right)^2 \nabla_{\mathbf{r}_i} \sum_{j \neq i}^{N} \frac{\exp\left(-\left|\mathbf{r}_i - \mathbf{r}_j\right|/\lambda\right)}{\left|\mathbf{r}_i - \mathbf{r}_j\right|}, \tag{3.6}$$

where $\mathbf{r}_i$ is the position of particle i, $Ze = 50000e$ is the particle charge (e is the elementary charge) and $\lambda = 0.4\,\mathrm{mm}$ is the screening length as proposed in Wörner et al. [2012]. The radial and vertical projections $F_{\rho,z}^{\mathrm{rep}}$ of the Yukawa force computed from the experimental data are shown in Fig. 3.9. The figure also contains the best-rms fits introduced by the relations

$$F_\rho^{\mathrm{rep}} = A_\rho \rho + B_\rho \rho^2, \quad F_z^{\mathrm{rep}} = A_z z, \tag{3.7}$$

with A_ρ, A_z, B_ρ the parameters and ρ the cylindrical radius. In the case of counterclockwise rotation, the particles occasionally approach each other very closely, leading to a wide spread of the estimated Yukawa forces [see Fig. 3.9(b) and (e)]. This leads to a large uncertainty of the parameters in the insets of Fig. 3.9 (e); the approach of the dynamic force balance may therefore not be applicable there.

The friction force is $\mathbf{F}^{\mathrm{fr}} = -M\gamma_{\mathrm{Eps}}\mathbf{v}$, where M is the particle mass, γ_{Eps} is the Epstein drag coefficient [Epstein, 1924] and $\mathbf{v}$ is the particle velocity. The value of the friction force at the cluster periphery has been estimated in Wörner et al. [2012] to be smaller than $10\,\mathrm{fN}$. Compared to the Yukawa forces in the same region of the cluster, which are on the order of 0.5–$2\,\mathrm{pN}$, the friction force is of minor importance and can be neglected in the balance of Eq. (3.5).

The inertial forces in the rotating frame of the cluster consist of the centrifugal force $\mathbf{F}^{\mathrm{centri}}$ and the Coriolis force $\mathbf{F}^{\mathrm{Cor}}$ [Hartmann et al., 2013]. The former is readily estimated as $F^{\mathrm{centri}} = M\omega^2\rho < 1\,\mathrm{fN}$, where $M = 1.1 \times 10^{-12}\,\mathrm{kg}$ is the particle mass [Wörner et al., 2012] and $\omega \simeq 0.4\,\mathrm{s}^{-1}$ is the rotation speed of the cluster. To estimate

---

[1]The confinement is provided by gravity, electric field of the plasma sheath, rotating electric field, ion drag force due to streaming ions, and ion wake mediated interaction.
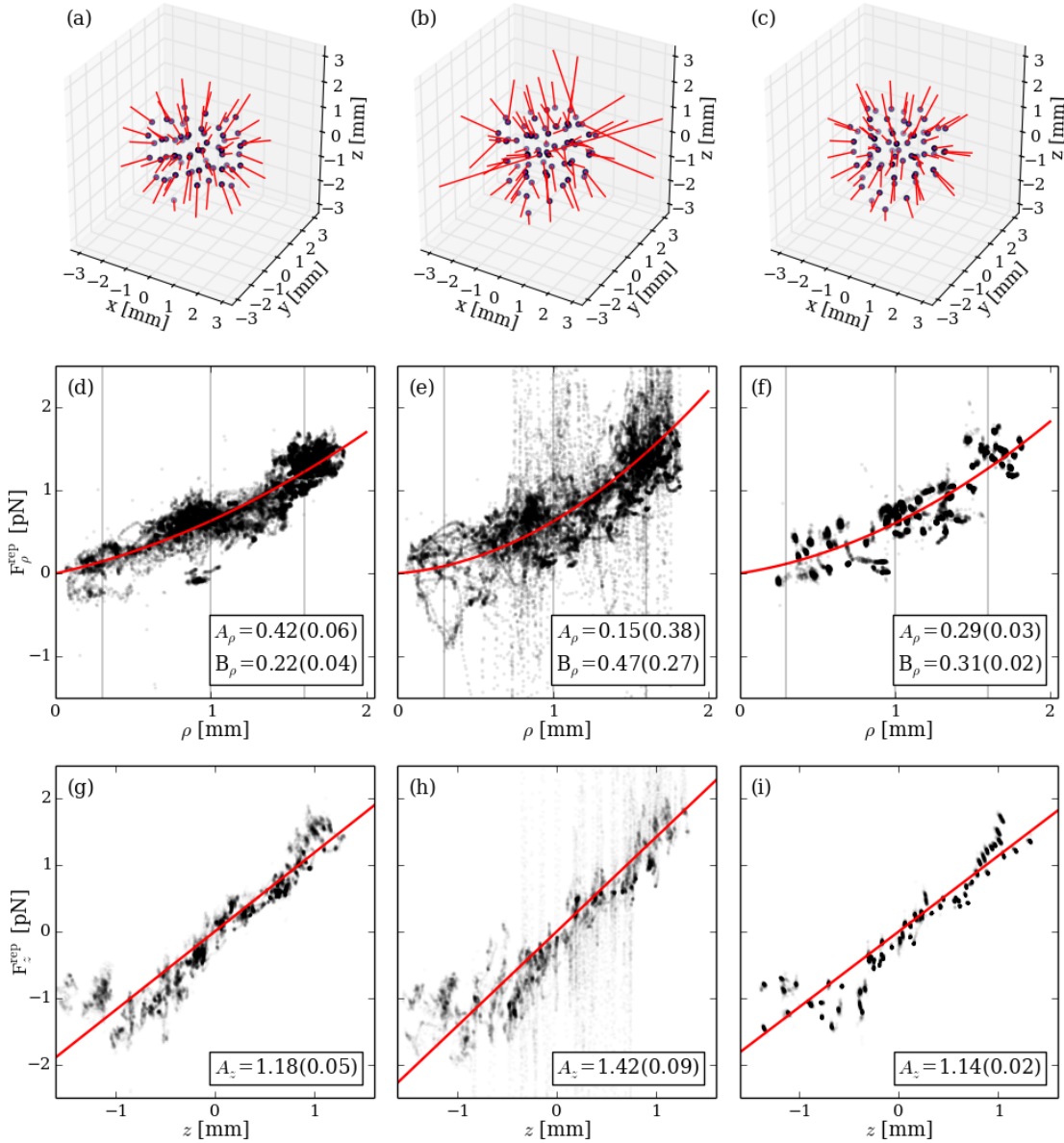
Figure 3.9: Repulsive Yukawa-type forces for the clockwise (a) and counterclockwise (b) rotating, and nonrotating (c) clusters are shown as red lines (in arbitrary units) pointing from the 3D positions of the particles. (d)–(f) The cylindrical projection of the Yukawa forces F_ρ^{rep} vs. the cylindrical radius ρ. The solid lines are the best-rms-fit parabolas [Eq. (3.7)] to the force profiles. The insets show the mean values of A_ρ in pN/mm and B_ρ in pN/mm^2 which are calculated for each time step, with the standard deviations indicated as errors. The vertical lines are the same as in Fig. 3.8. (d)–(f) The same for the vertical profiles of the Yukawa forces F_ρ^{rep} , here, the few particles suspended below the cluster at $z \simeq -2$ mm (see Fig. 3.8) are left out. The insets show the parameter A_z in pN/mm of the linear fit to the profiles.

the particle velocities, five consecutive frames are averaged, as the jitter on some particles due to imaging processes complicates the calculation of the instantaneous velocity. The Coriolis force, which is maximal during transitions of a particle from an inner to an outer string or *vice versa*, then yields $F^{\text{Cor}} = 2M|\boldsymbol{\omega} \times \mathbf{v}| < 3\,\text{fN}$. Hence, compared to the Yukawa forces, the contribution of inertial forces can be neglected. The confinement force profiles can thus be estimated as $F_{\rho,z}^{\text{conf}} \simeq -F_{\rho,z}^{\text{rep}}$.

Given the values of the fitting parameters A_ρ introduced above, the radial confinement near the center of the cluster can be estimated as

$$F_\rho^{\text{conf}} \simeq -M\Omega_\rho^2 \rho, \ \ M\Omega_\rho^2 \equiv A_\rho, \tag{3.8}$$

where Ω_ρ is the cylindrical confinement parameter. The values of Ω_ρ as well as the values of the vertical confinement parameter Ω_z (obtained similarly with the relation $M\Omega_z^2 \equiv A_z$) are shown in Table 3.1.

It is not surprising that the vertical confinement force is systematically stronger than the cylindrical one, as can be naturally expected due to the stronger forces of the sheath electric field and gravity compressing the cluster vertically [Wörner et al., 2011]. Note that a stronger vertical confinement was also found in Stokes et al. [2008] for "dust molecules" consisting of only two particles.

Neglecting the case of counterclockwise rotation due to the wide spread of data points, this approach yields a stronger cylindrical confinement for the clockwise rotating cluster than for the nonrotating one (see Table 3.1). This stronger confinement, increasing the *cylindricity* of the rotating cluster, is in good agreement with the findings from the network analysis of Sec. 3.4.1. To study individual contributions to $\mathbf{F}^{\text{conf}}$ from, e. g., the ion drag force, dedicated measurements are needed.

From the estimate of the centrifugal and the Coriolis force it follows, furthermore, that the structural changes in the rotating cluster are not due to the cluster rotation *per se*. It must rather be inferred that the electrostatic confinement of the whole cluster changes with the applied frequency, allowing for string formation within the rotating cluster.

3.4.3   Discussion

Both the network analysis of the global structure and the estimate of the particle confinement show that the cylindrical symmetry is more important in the case of cluster rotation than in the case of nonrotation (see Table 3.1). Dynamically driven clusters favor the formation of vertical strings which were observed in Sec. 3.3. As can be seen in Fig. 3.2(a), the detected strings are also more compact in the case of cluster rotation.

As to the network analysis, it can be seen in Fig. 3.7 that the results depend sensitively on the noise amplitude of the artificial null models. The exact values of the ratios R thus have to be treated with caution, but they are a convenient tool for comparing structures. Considering Fig. 3.7, it is not possible to name the exact value of the ratio of the competing spherical and cylindrical geometries of the clusters, but

it is evident that the nonrotating cluster systematically shows a higher value of R. The geometry is thus more spherical in the case of the nonrotating cluster, while the rotating clusters have a more pronounced cylindrical geometry.

For the Yukawa interactions used for the estimate of the particle confinement (see Eq. 3.5), constant particle charge Ze and screening length λ are assumed. It is straightforward to see that different values of λ and Z only change the scale of the profiles in Fig. 3.9, and not the general shape. While Yukawa interaction is valid and widely used in two-dimensional systems, an additional contribution from the wake-field interaction is not negligible in 3D [Lampe et al., 2000; Ivlev et al., 2012]. An adequate description of the wake field interaction is not straightforward and the subject of current research. Nonetheless, this simplified model of the Yukawa-type interaction gives a hint for the origin of the differences in symmetry revealed by network analysis. A more detailed analysis should take into account the mass dispersion of the particles as well as the anisotropy in the interparticle interaction.

Chapter 4

Conclusion

The test for nonlinearity based on network measures of recurrence networks proposed in Chapter 2 has proven a bit weaker than the nonlinear prediction error which is a dynamic measure for nonlinearity. The network-based test could possibly be improved by comparing the distributions of the network measures containing the whole information for each node instead of the average values. Another route would be to consider visibility graphs instead of recurrence networks [Lacasa et al., 2008; Luque et al., 2009]. There, the time series is considered as landscape, and a visibility criterion determines whether points are connected or not, concentrating more on the dynamic aspects of the time series.

Induced phase correlations were found in surrogate data produced by the established algorithms AAFT and IAAFT. The influence of these correlations on tests for nonlinearity were shown both for the nonlinear prediction error and for the network measure average connectivity. In order to gain a deeper insight into the relation of correlations in the Fourier phases on one hand and the properties of data in real space on the other hand, further nonlinear time series should be studied. A broader range of light curves of active galactic nuclei could already give additional insights. Other candidates are stock markets, showing a multitude of interesting nonlinear properties. Also, equations that reliably produce artificial time series with phase correlations would be an informative subject, but are not known.

In Chapter 3, two promising applications of networks to complex plasmas were described. To my knowledge, it is the first network analysis of spatial structures that are not obviously network-like. The detection of vertical strings with the aid of a community finding algorithm has proven an elegant way of examining stable units in complex structures. A further advantage of this method was shown to be the ability to examine the time evolution of a system as a whole with the aid of multislice networks, yielding information about the dynamics of the system. These advantages were shown for the detection and analysis of vertical strings in dynamically driven complex plasma clusters.

Multislice networks could possibly also be used to identify (quasi)particles in more complicated structures. Complex plasmas offer a multitude of interesting possible further applications, like the onset of synchronization in a melting two-dimensional

crystal [Couëdel et al., 2014] or the dynamics of defects that have discontinued displacements and can be created and annihilated.

Network analysis also enabled a throughout study of the global structure of the clusters. For the relatively small clusters of about 60 particles a significant difference between the well-studied case of nonrotating clusters and dynamically driven clusters was demonstrated. Driving the clusters with external fields leads to a more cylindrical structure.

The analysis could be repeated for any structure of competing or emerging symmetries if appropriate null models exist. One candidate is the forthcoming PK-4 setup onboard of the International Space Station to study complex plasmas under microgravity conditions [Thoma et al., 2007]. For example, the emergence of anisotropic structures in these strongly coupled systems can be examined [Ivlev et al., 2011].

Appendix

Appendix A
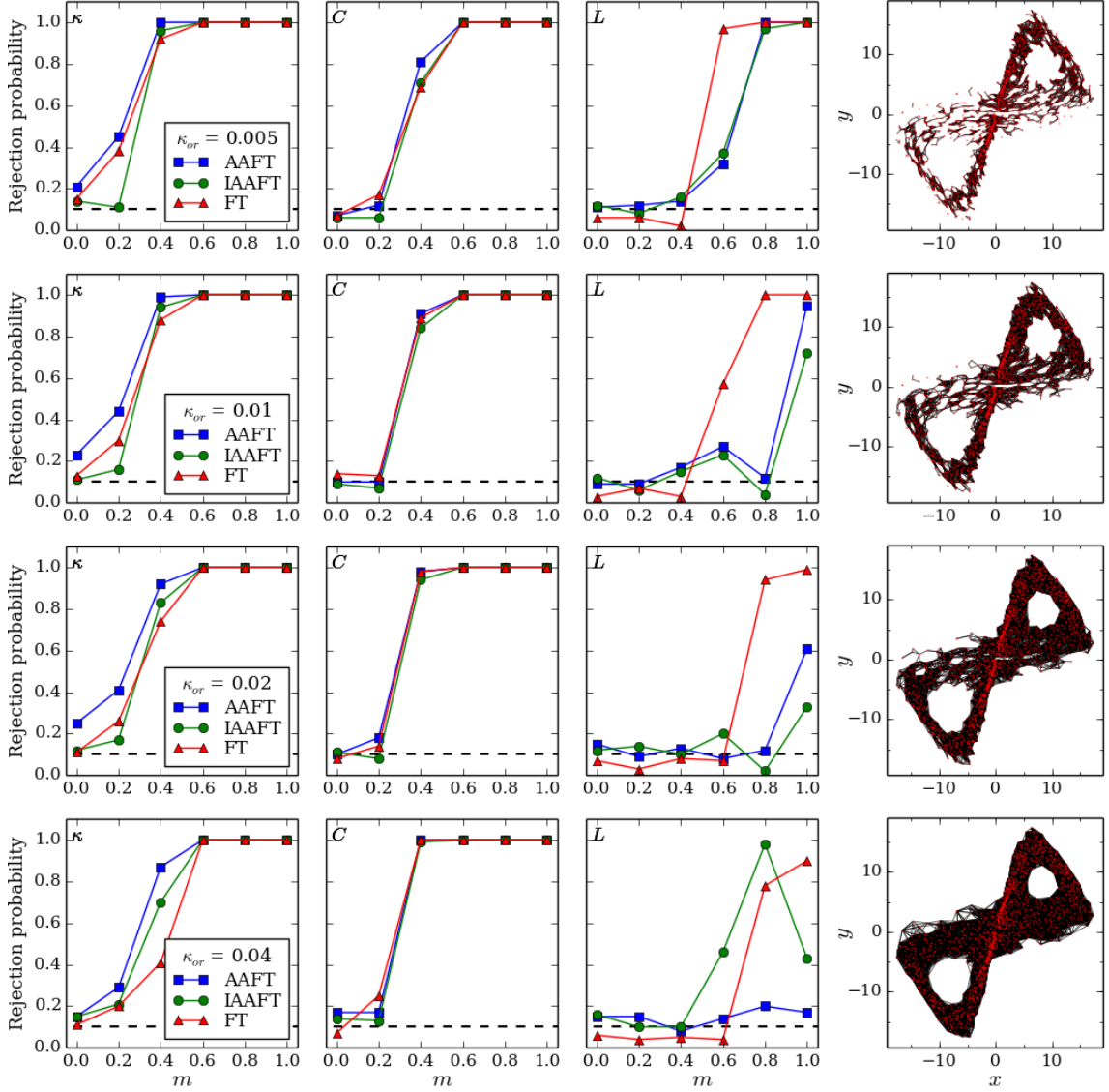
Further Plots of Network Measures on Time Series

Figure A.1: Discrimination power of 100 tests based on network measures with different κ_{or} applied to the mixed Lorenz data (see Sec. 2.3.2). The value of κ_{or} is given in the legend of the leftmost plot and is valid for the whole line. Higher values of κ_{or} lead to networks that are more connected as can be seen in a representation of a network of a time series with $m = 1$ (pure Lorenz) where the positions of the nodes are the projections of the corresponding point in phase space onto the xy plane. The network measures are indicated at the top left corner of each plot.
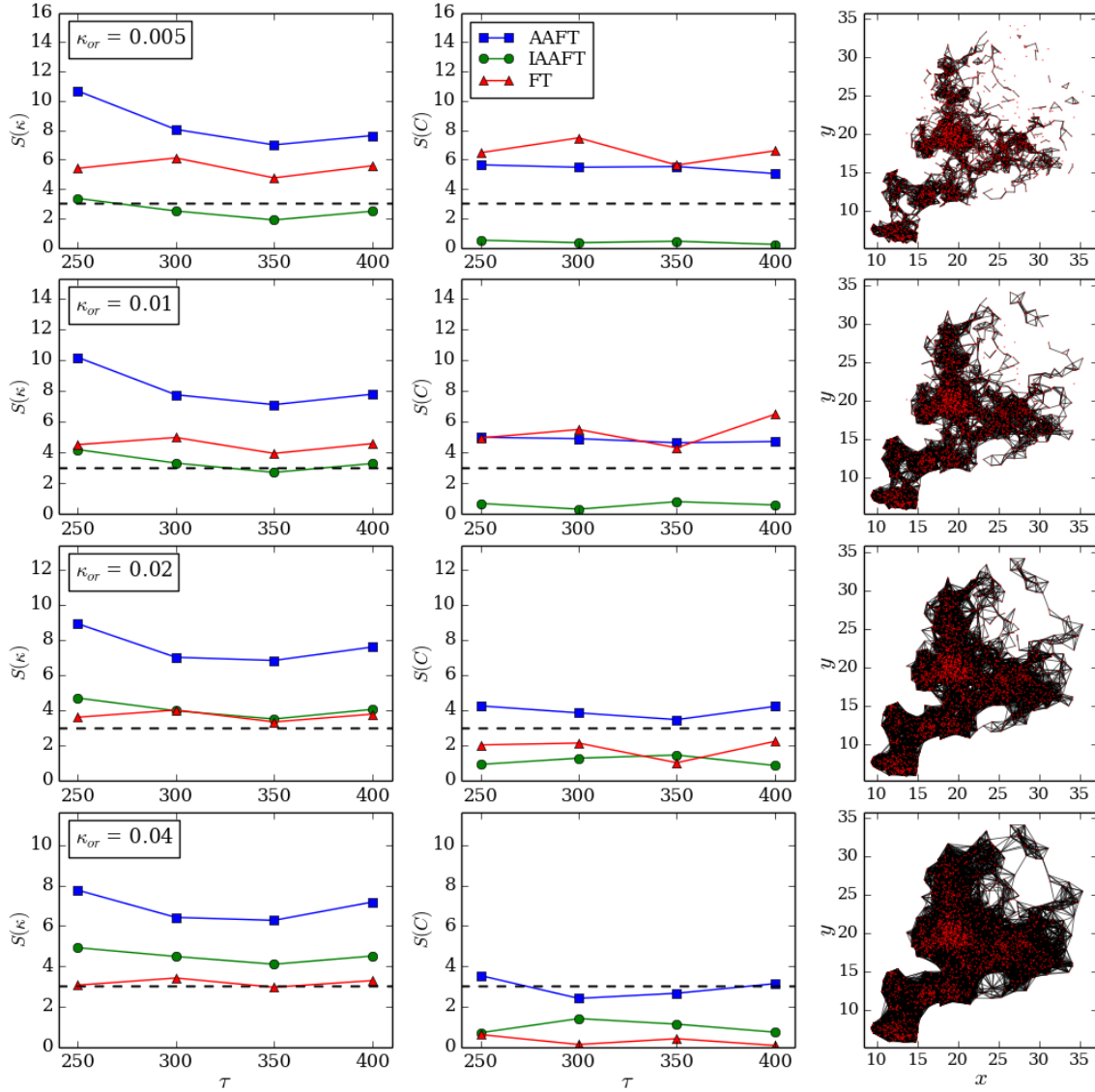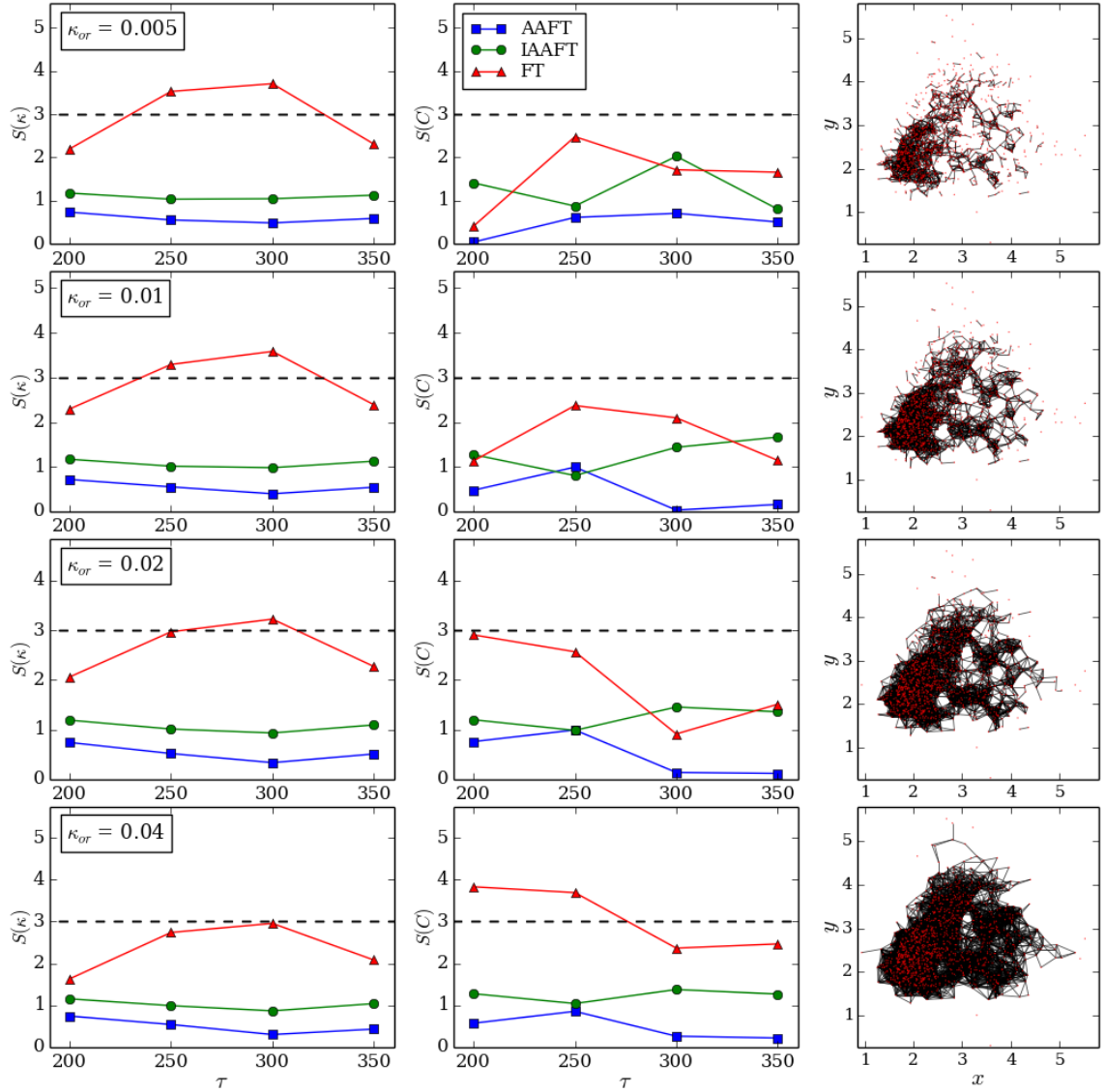
Figure A.2: Significances of tests based on network measures for the MCG data as a function of delay time τ (see Sec. 2.4.2). The value of κ_{or} is given in the leftmost plot and is valid for the whole line. First row: connectivity κ, second row: clustering coefficient C. 200 realizations of the surrogates were used. The network of the data is shown in the rightmost plot; the nodes' positions are the projections of the embedded data onto the xy plane.

Figure A.3: Significances of tests based on network measures for the Mrk data as a function of delay time τ (see Sec. 2.4.2). The value of κ_{or} is given in the leftmost plot and is valid for the whole line. First row: connectivity κ, second row: clustering coefficient C. 200 realizations of the surrogates were used. The network of the data is shown in the rightmost plot; the nodes' positions are the projections of the embedded data onto the xy plane.
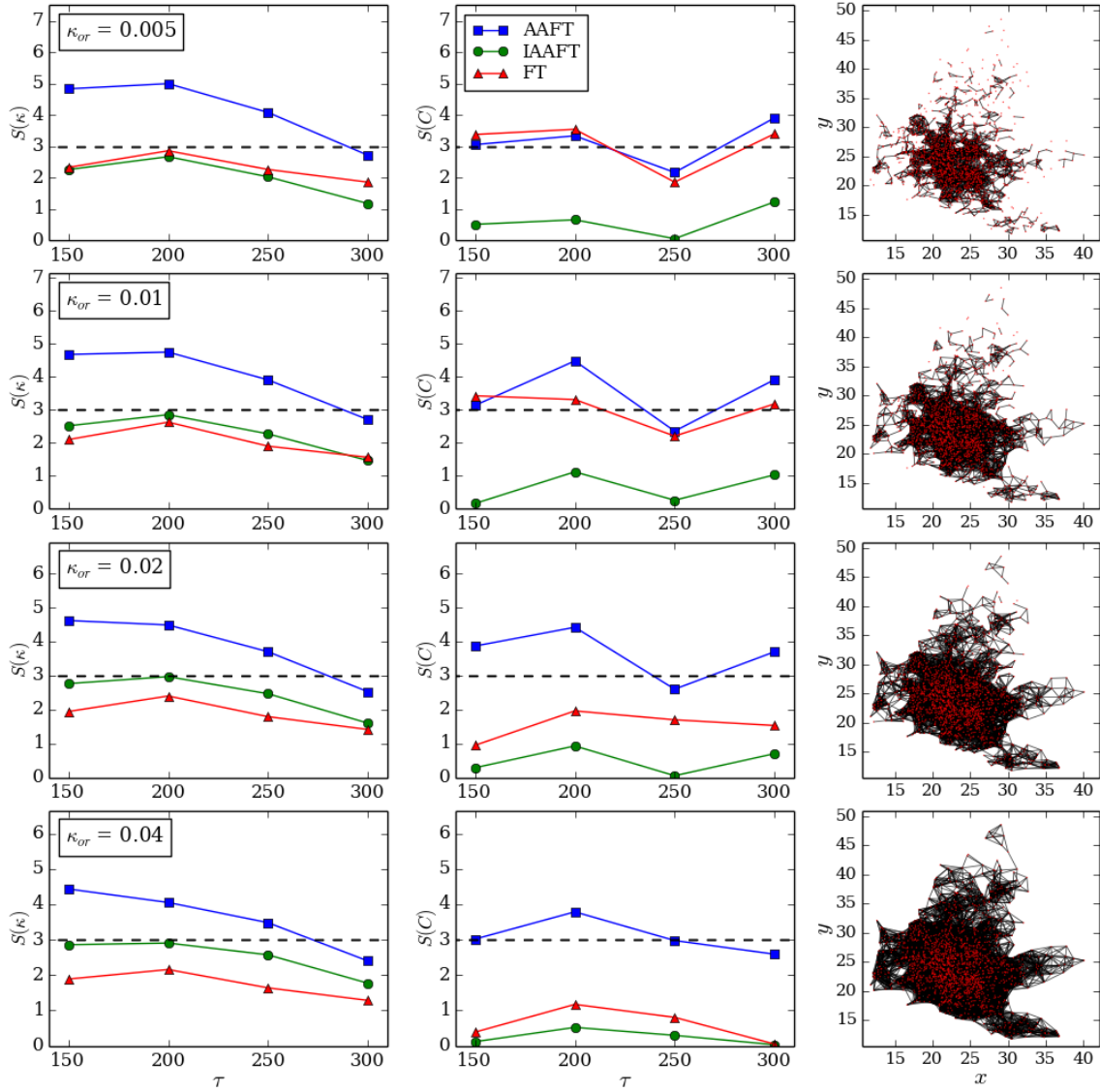
Figure A.4: As Fig. A.3 for the Akn data.

Appendix B

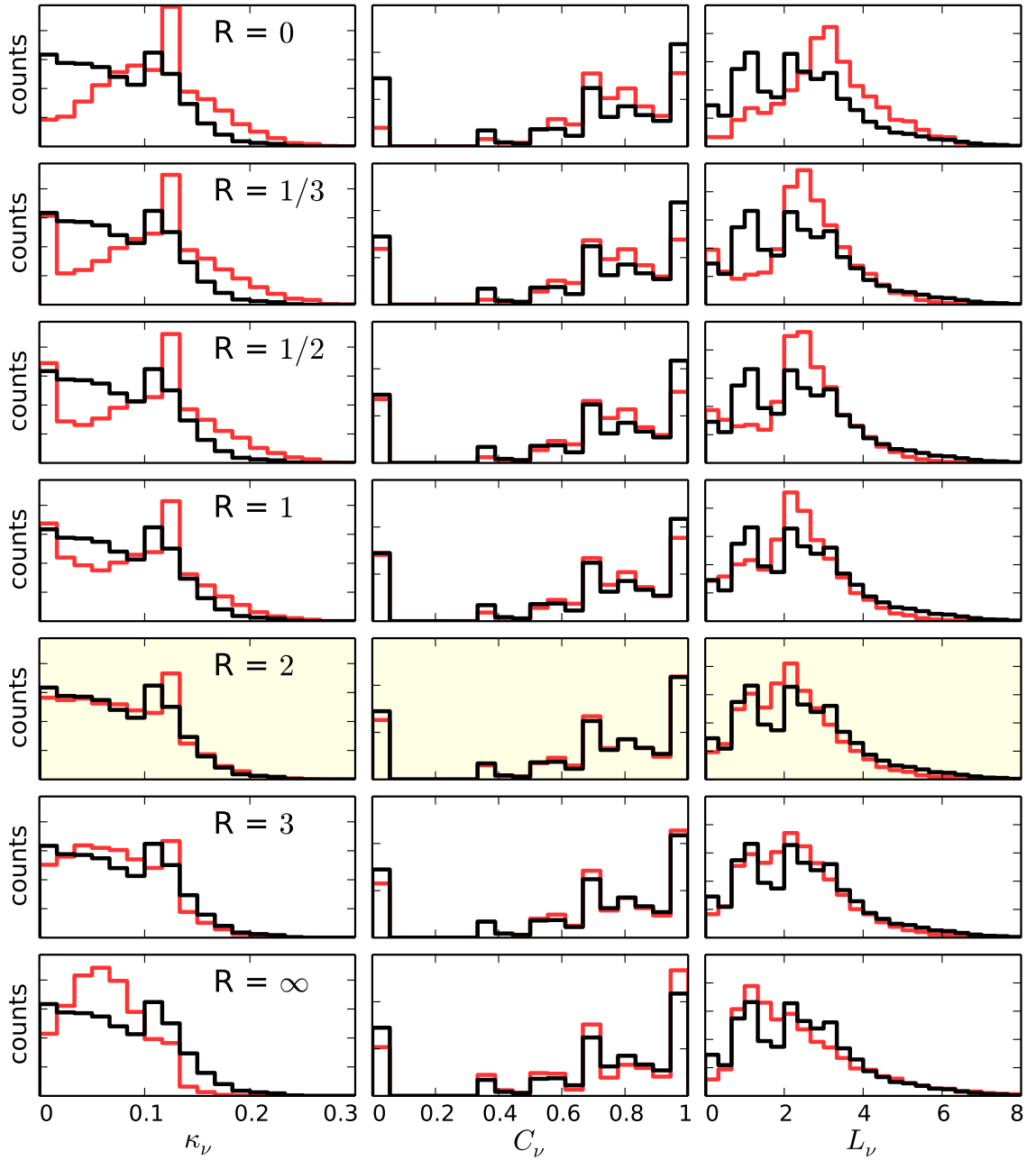# Further Plots of Network Measures on Complex Plasma Clusters

Figure B.1: Comparison of network measures on A^{cyl} for the counterclockwise rotating cluster (see Sec. 3.4.1). The histograms (for all particles and time steps) of measures on networks of the clockwise rotating cluster are plotted in black. They are identical in all rows. The null models are plotted in red with the ratio ranging from $R = 0$ (top row, purely cylindrical structure) to $R = \infty$ (bottom row, purely spherical structure). First column: connectivity κ_ν, second column: clustering coefficient C_ν, third column: path length L_ν. The row with the best agreement for all three network measures is highlighted with a pale yellow background.
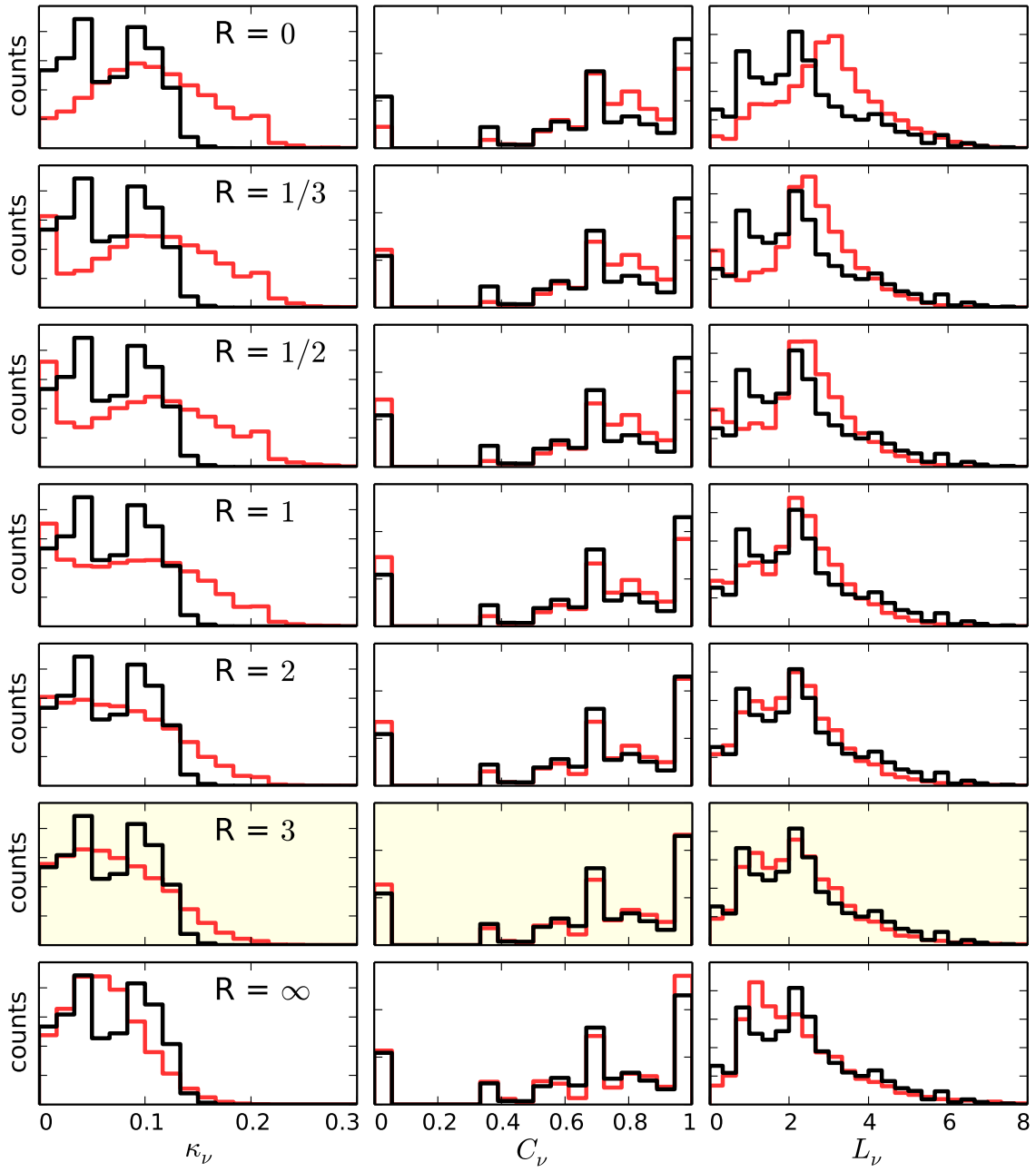
Figure B.2: As Fig. B.1 for measures on A^{cyl} for the nonrotating cluster.
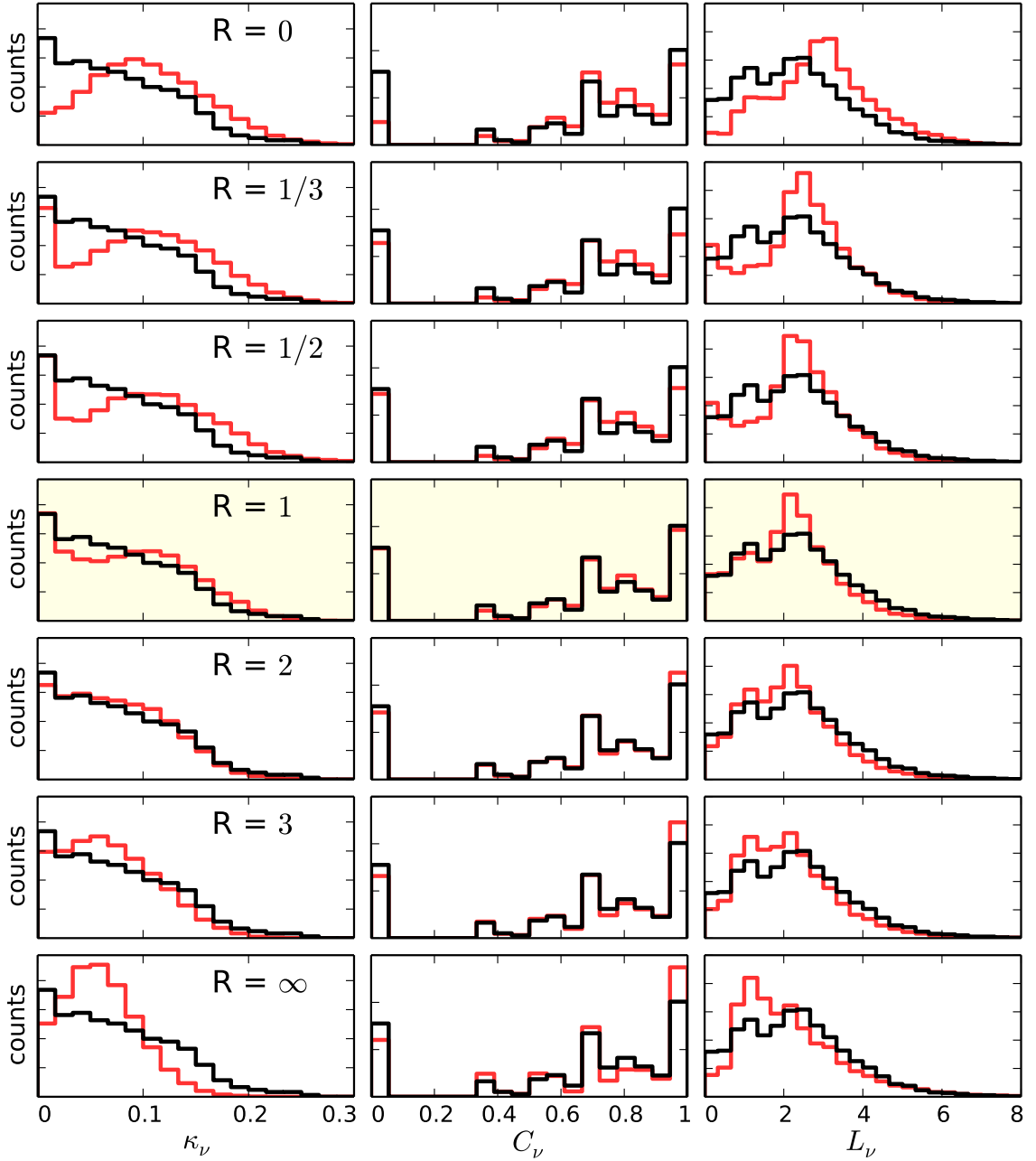
Figure B.3: Comparison of network measures on A^{sph} for the clockwise rotating cluster (see Sec. 3.4.1). The histograms (for all particles and time steps) of measures on networks of the clockwise rotating cluster are plotted in black. They are identical in all rows. The null models are plotted in red with the ratio ranging from $R = 0$ (top row, purely cylindrical structure) to $R = \infty$ (bottom row, purely spherical structure). First column: connectivity κ_ν, second column: clustering coefficient C_ν, third column: path length L_ν. The row with the best agreement for all three network measures is highlighted with a pale yellow background.
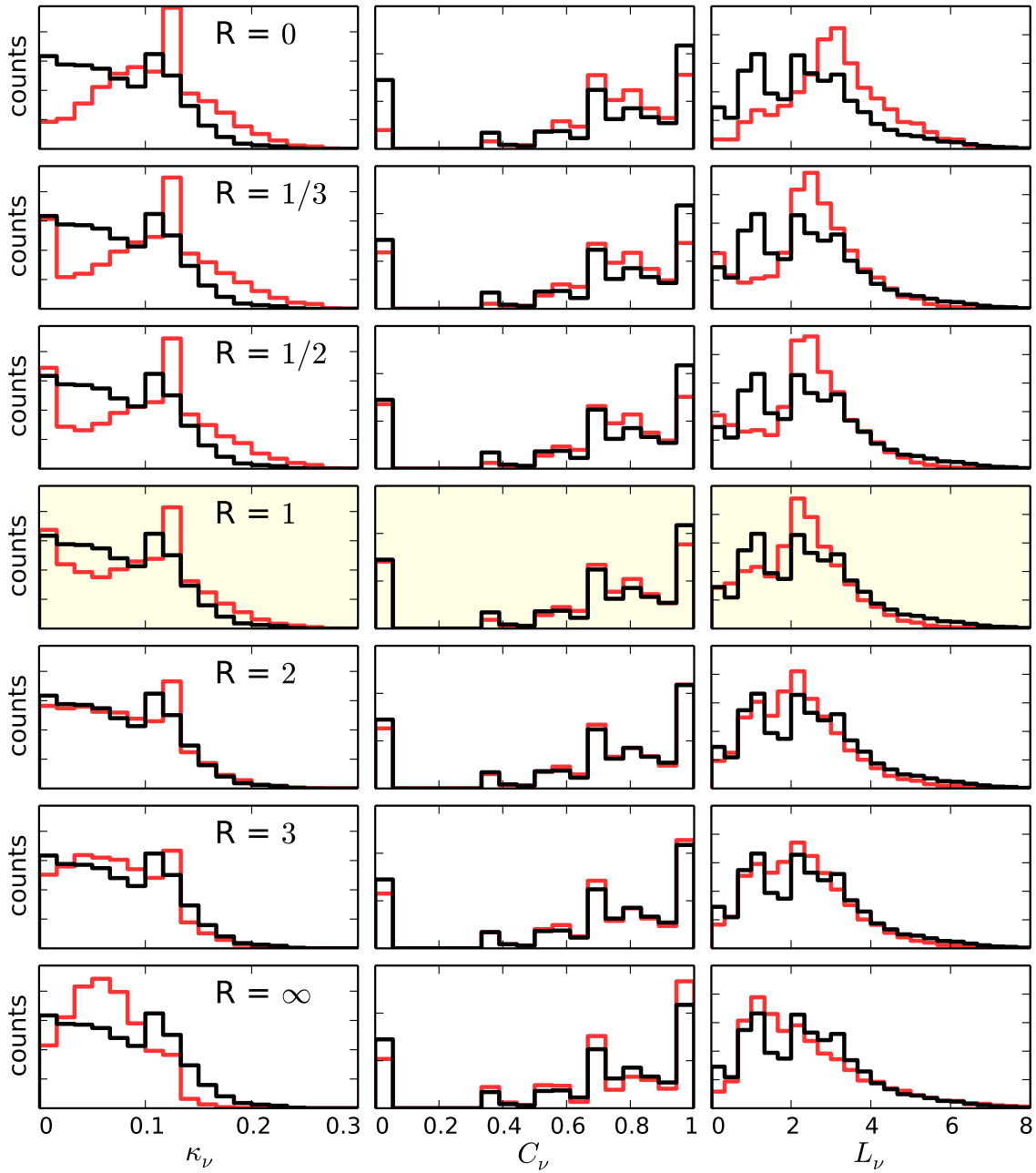
Figure B.4: As Fig. B.3 for measures on A^{sph} for the counterclockwise rotating cluster.
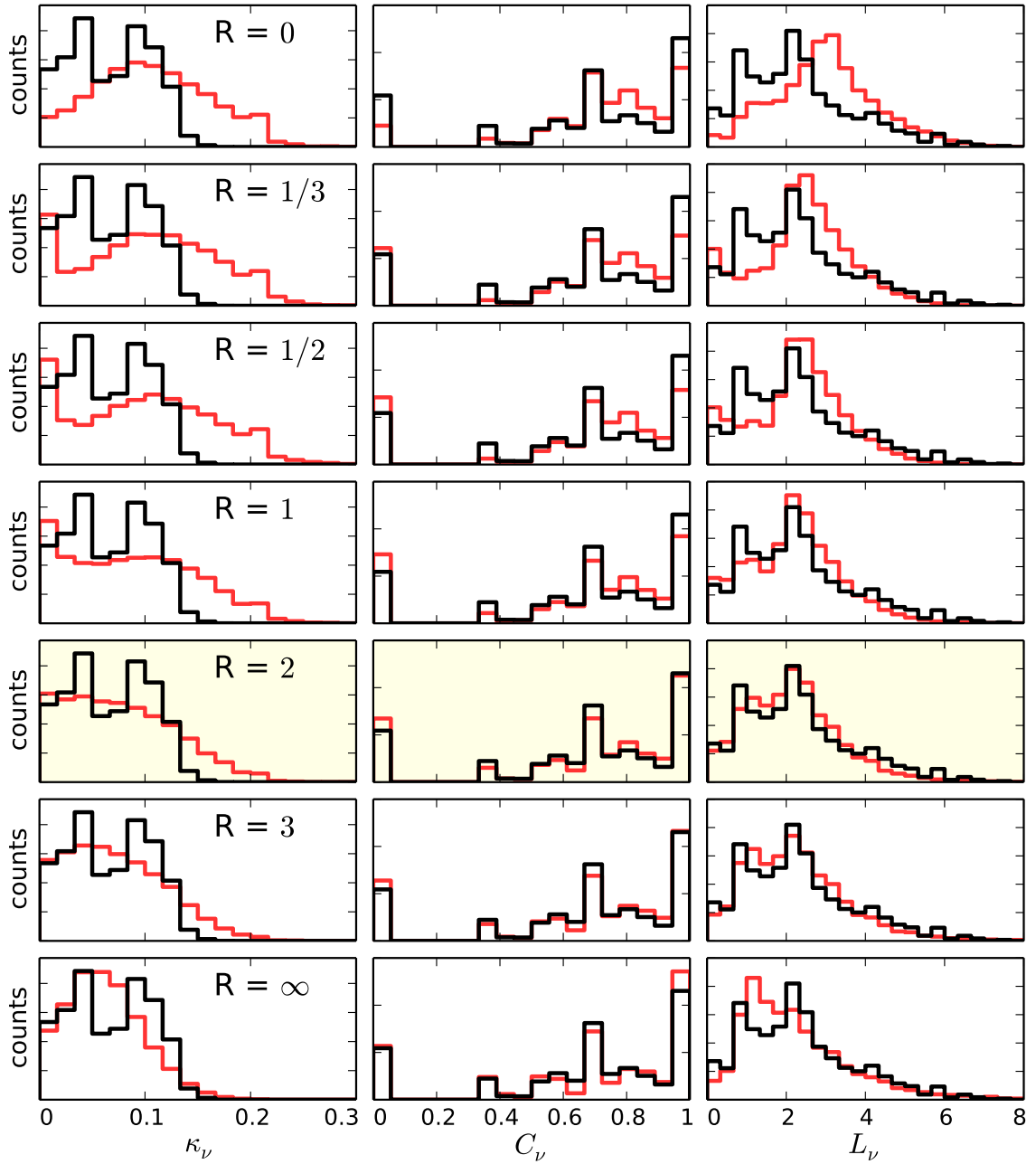
Figure B.5: Comparison of network measures on A^{sph} for the nonrotating cluster (see Sec. 3.4.1). The histograms (for all particles and time steps) of measures on networks of the clockwise rotating cluster are plotted in black. They are identical in all rows. The null models are plotted in red with the ratio ranging from $R = 0$ (top row, purely cylindrical structure) to $R = \infty$ (bottom row, purely spherical structure). First column: connectivity κ_ν, second column: clustering coefficient C_ν, third column: path length L_ν. The row with the best agreement for all three network measures is highlighted with a pale yellow background.

Bibliography

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):47–97.

Albert, R., Jeong, H., and Barabási, A.-L. (1999). Internet: Diameter of the World-Wide Web. *Nature*, 401(6749):130–131.

Antonucci, R. (1993). Unified models for active galactic nuclei and quasars. *Annu. Rev. Astron. Astrophys.*, 31:473–521.

Arp, O., Block, D., Klindworth, M., and Piel, A. (2005). Confinement of Coulomb balls. *Phys. Plasmas*, 12:122102.

Arp, O., Block, D., Piel, A., and Melzer, A. (2004). Dust Coulomb balls: three-dimensional plasma crystals. *Phys. Rev. Lett.*, 93:165004.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Bassett, D. S., Porter, M. A., Wymbs, N. F., Grafton, S. T., Carlson, J. M., and Mucha, P. J. (2013). Robust detection of dynamic community structure in networks. *Chaos*, 23(1):013142.

Benlloch, S., Wilms, J., Edelson, R., Yaqoob, T., and Staubert, R. (2001). Quasi-periodic oscillation in Seyfert galaxies: Significance levels. The case of Markarian 766. *Astrophys. J.*, 562(2):L121–L124.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theor. Exp.*, 2008(10):P10008.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Phys. Rep.*, 424(4):175–308.

Boller, T., Keil, R., Trümper, J., OBrien, P., Reeves, J., and Page, M. (2001). Detection of an X-ray periodicity in the narrow-line Seyfert 1 galaxy Mrk 766 with XMM-Newton. *Astron. Astrophys.*, 365:L146–L151.

Carstensen, J., Greiner, F., and Piel, A. (2010). Determination of dust grain charge and screening lengths in the plasma sheath by means of a controlled cluster rotation. *Phys. Plasmas*, 17:083703.

Couëdel, L., Zhdanov, S., Nosenko, V., Ivlev, A. V., Thomas, M. H., and Morfill, G. E. (2014). Synchronization of particle motion induced by mode coupling in a two-dimensional plasma crystal. To be published in Phys. Rev. E.

Donges, J. F., Donner, R. V., and Kurths, J. (2013). Testing time series irreversibility using complex network methods. *Europhys. Lett.*, 102(1):10004.

Donges, J. F., Donner, R. V., Trauth, M. H., Marwan, N., Schellnhuber, H.-J., and Kurths, J. (2011). Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. *Proc. Natl. Acad. Sci. U.S.A.*, 108(51):20422–20427.

Donner, R. V., Zou, Y., Donges, J. F., Marwan, N., and Kurths, J. (2010). Recurrence networks - a novel paradigm for nonlinear time series analysis. *New J. Phys.*, 12(3):033025.

Eckmann, J.-P., Kamphorst, S. O., and Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhys. Lett.*, 4(9):973–977.

Edelson, R., Turner, T., Pounds, K., Vaughan, S., Markowitz, A., Marshall, H., Dobbie, P., and Warwick, R. (2002). X-ray spectral variability and rapid variability of the soft X-ray spectrum Seyfert 1 galaxies Arakelian 564 and Ton S180. *Astrophys. J.*, 568(2):610–626.

Epstein, P. S. (1924). On the resistance experienced by spheres in their motion through gases. *Phys. Rev.*, 23(6):710–733.

Erdös, P. and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5:17–61.

Fortov, V. E., Ivlev, A. V., Khrapak, S. A., Khrapak, A. G., and Morfill, G. E. (2005). Complex (dusty) plasmas: Current status, open issues, perspectives. *Phys. Rep.*, 421(1):1–103.

Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215–239.

Gierliński, M., Middleton, M., Ward, M., and Done, C. (2008). A periodicity of 1 hour in X-ray emission from the active galaxy RE J1034+ 396. *Nature*, 455(7211):369–371.

Grassberger, P. and Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica 9 D*, 9:189–208.

Hartmann, P., Donkó, Z., Ott, T., Kählert, H., and Bonitz, M. (2013). Magnetoplasmons in rotating dusty plasmas. *Phys. Rev. Lett.*, 111:155002.

Hyde, T. W., Kong, J., and Matthews, L. S. (2013). Helical structures in vertically aligned dust particle chains in a complex plasma. *Phys. Rev. E*, 87(5):053106.

Ivlev, A., Löwen, H., Morfill, G., and Royall, C. P. (2012). *Complex Plasmas and Colloidal Dispersions: Particle-resolved Studies of Classical Liquids and Solids.* World Scientific, Singapore.

Ivlev, A. V., Thoma, M. H., Räth, C., Joyce, G., and Morfill, G. E. (2011). Complex plasmas in external fields: The role of non-Hamiltonian interactions. *Phys. Rev. Lett.*, 106(15):155001.

Kantz, H. and Schreiber, T. (2004). *Nonlinear time series analysis*, volume 7. Cambridge university press.

Kennel, M. B., Brown, R., and Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A*, 45(6):3403–3411.

Kong, J., Hyde, T. W., Matthews, L., Qiao, K., Zhang, Z., and Douglass, A. (2011). One-dimensional vertical dust strings in a glass box. *Phys. Rev. E*, 84(1):016411.

Kroll, M., Harms, S., Block, D., and Piel, A. (2008). Digital in-line holography of dusty plasmas. *Phys. Plasmas*, 15:063703.

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. U.S.A.*, 105(13):4972–4975.

Lamb, B. M. and Morales, G. J. (1983). Ponderomotive effects in nonneutral plasmas. *Phys. Fluids*, 26:3488–3496.

Lambiotte, R., Delvenne, J.-C., and Barahona, M. (2008). Laplacian dynamics and multiscale modular structure in networks. *arXiv preprint arXiv:0812.1770*.

Lampe, M., Joyce, G., Ganguli, G., and Gavrishchaka, V. (2000). Interactions between dust grains in a dusty plasma. *Phys. Plasmas*, 7:3851–3861.

Laut, I., Räth, C., Wörner, L., Nosenko, V., Zhdanov, S. K., Schablinski, J., Block, D., Thoma, H. M., and Morfill, G. E. (2014). Network analysis of three-dimensional complex plasma clusters in a rotating electric field. *Phys. Rev. E*, 89:023104.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20(2):130–141.

Luque, B., Lacasa, L., Ballesteros, F., and Luque, J. (2009). Horizontal visibility graphs: Exact results for random time series. *Phys. Rev. E*, 80(4):046103.

Markowitz, A., Papadakis, I., Arévalo, P., Turner, T., Miller, L., and Reeves, J. (2007). The energy-dependent X-ray timing characteristics of the narrow-line Seyfert 1 Mrk 766. *Astrophys. J.*, 656(1):116.

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(46):4246–4254.

McHardy, I., Gunn, K., Uttley, P., and Goad, M. (2005). Mcg–6-30-15: Long time-scale X-ray variability, black hole mass and active galactic nuclei high states. *Mon. Not. R. Astron. Soc.*, 359(4):1469–1480.

McHardy, I. M., Koerding, E., Knigge, C., Uttley, P., and Fender, R. P. (2006). Active galactic nuclei as scaled-up galactic black holes. *Nature*, 444(7120):730–732.

Melzer, A. (2006). Zigzag transition of finite dust clusters. *Phys. Rev. E*, 73(5):056404.

Melzer, A., Buttenschön, B., Miksch, T., Passvogel, M., Block, D., Arp, O., and Piel, A. (2010). Finite dust clusters in dusty plasmas. *Plasma Phys. Controlled Fusion*, 52(12):124028.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys*, 21(6):1087–1092.

Middleton, M. and Done, C. (2010). The X-ray binary analogy to the first AGN quasi-periodic oscillation. *Mon. Not. R. Astron. Soc.*, 403(1):9–16.

Milgram, S. (1967). The small world problem. *Psychol. Today*, 2(1):60–67.

Miller, J. (2007). Relativistic X-ray lines from the inner accretion disks around black holes. *Annu. Rev. Astron. Astrophys.*, 45:441–479.

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., and Onnela, J.-P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science*, 328(5980):876–878.

Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E*, 74(3):036104.

Newman, M. E. J. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(2):026113.

Nosenko, V., Ivlev, A. V., Zhdanov, S. K., Fink, M., and Morfill, G. E. (2009). Rotating electric fields in complex (dusty) plasmas. *Phys. Plasmas*, 16:083708.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). Geometry from a time series. *Phys. Rev. Lett.*, 45:712–716.

Porter, M. A., Onnela, J.-P., and Mucha, P. J. (2009). Communities in networks. *Not. Am. Math. Soc.*, 56(9):1082–1097.

Räth, C., Gliozzi, M., Papadakis, I. E., and Brinkmann, W. (2012). Revisiting algorithms for generating surrogate time series. *Phys. Rev. Lett.*, 109:144101.

Remillard, R. A. and McClintock, J. E. (2006). X-ray properties of black-hole binaries. *Annu. Rev. Astron. Astrophys*, 44:49–92.

Rossmanith, G., Modest, H., Räth, C., Banday, A. J., Górski, K. M., and Morfill, G. E. (2012). Probing non-Gaussianities in the cosmic microwave background on an incomplete sky using surrogates. *Phys. Rev. D*, 86(8):083005.

Ruelle, D. and Takens, F. (1971). On the nature of turbulence. *Comm. Math. Phys.*, 20(3):167–192.

Schella, A., Mulsow, M., Melzer, A., Schablinski, J., and Block, D. (2013). From transport to disorder: Thermodynamic properties of finite dust clouds. *Phys. Rev. E*, 87(6):063102.

Schreiber, T. (1998). Constrained randomization of time series data. *Phys. Rev. Lett.*, 80(10):2105.

Schreiber, T. and Schmitz, A. (1996). Improved surrogate data for nonlinearity tests. *Phys. Rev. Lett.*, 77(4):635.

Schreiber, T. and Schmitz, A. (1997). Discrimination power of measures for nonlinearity in a time series. *Phys. Rev. E*, 55(5):5443.

Schreiber, T. and Schmitz, A. (2000). Surrogate time series. *Physica D*, 142(3):346–382.

Stokes, J. D. E., Samarian, A. A., and Vladimirov, S. V. (2008). Dynamics of two particles in a plasma sheath. *Phys. Rev. E*, 78(3):036402.

Strogatz, S. H. (2006). *Nonlinear dynamics and chaos*. Perseus Publishing.

Sugihara, G. and May, R. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344(6268):734–734.

Takens, F. (1981). Detecting strange attractors in turbulence. In *Lecture Notes in Mathematics*, pages 366–381. Springer.

Tanaka, Y., Nandra, K., Fabian, A., Inoue, H., Otani, C., Dotani, T., Hayashida, K., Iwasawa, K., Kii, T., Kunieda, H., et al. (1995). Gravitationally redshifted emission implying an accretion disk and massive black-hole in the active galaxy MCG-6-30-15. *Nature*, 375(6533):659–661.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Doyne Farmer, J. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 58(1):77–94.

Thoma, M. H., Fink, M. A., Hofner, H., Kretschmer, M., Khrapak, S. A., Ratynskaia, S. V., Yaroshenko, V. V., Morfill, G. E., Petrov, O. F., Usachev, A. D., et al. (2007). PK-4: Complex plasmas in Spacethe next generation. *IEEE Trans. Plasma Sci.*, 35(2):255–259.

Thomas, H., Morfill, G. E., Demmel, V., Goree, J., Feuerbacher, B., and Möhlmann, D. (1994). Plasma crystal: Coulomb crystallization in a dusty plasma. *Phys. Rev. Lett.*, 73(5):652.

Titarchuk, L. and Osherovich, V. (2000). The global normal disk oscillations and the persistent low-frequency quasi-periodic oscillations in X-ray binaries. *Astrophys. J. Lett.*, 542(2):L111–L114.

Totsuji, H., Ogawa, T., Totsuji, C., and Tsuruta, K. (2005). Structure of spherical Yukawa clusters: A model for dust particles in dusty plasmas in an isotropic environment. *Phys. Rev. E*, 72(3):036406.

Tsytovich, V. N., Morfill, G. E., Fortov, V. E., Gusein-Zade, N. G., Klumov, B. A., and Vladimirov, S. V. (2007). From plasma crystals and helical structures towards inorganic living matter. *New J. Phys.*, 9(8):263.

Wackerbauer, R., Witt, A., Atmanspacher, H., Kurths, J., and Scheingraber, H. (1994). A comparative classification of complexity measures. *Chaos, Solitons Fractals*, 4(1):133–173.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393(6684):440–442.

Wörner, L., Nosenko, V., Ivlev, A. V., Zhdanov, S. K., Thomas, H. M., Morfill, G. E., Kroll, M., Schablinski, J., and Block, D. (2011). Effect of rotating electric field on 3D complex (dusty) plasma. *Phys. Plasmas*, 18:063706.

Wörner, L., Räth, C., Nosenko, V., Zhdanov, S. K., Thomas, H. M., Morfill, G. E., Schablinski, J., and Block, D. (2012). String structures in driven 3D complex-plasma clusters. *Europhys. Lett.*, 100(3):35001.

Xu, X., Zhang, J., and Small, M. (2008). Superfamily phenomena and motifs of networks induced from time series. *Proc. Natl. Acad. Sci. U.S.A.*, 105(50):19601–19605.

Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

_____
Ort, Datum

_____
Unterschrift