Information-theoretic approach for the characterization of interactions in nonlinear dynamical systems

Dissertation zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
mathematisch-naturwissenschaftlichen Fakultät
der
Universität Bonn

vorgelegt von
Anton Chernihovskyi
aus Aschgabat

Bonn, 2010

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. K. Lehnertz
2. Gutachter: Prof. Dr. K. Maier

Tag der mündlichen Prüfung: 15. April 2011

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
`http://hss.ulb.uni-bonn.de/diss_online` elektronisch publiziert.

Erschienen 2011

In dieser Dissertation eingebunden:

- Zusammenfassung

Zusammenfassung

Mit Hilfe der Zeitreihenanalyse können Interaktionen zwischen natürlichen dynamischen Systemen anhand experimenteller Daten charakterisiert werden. In den letzten Jahren wurde eine Reihe von Maßen vorgestellt, die darauf abzielen, neben der Interaktionsrichtung auch die Interaktionsstärke zu bestimmen. Die zur Charakterisierung von Interaktionsrichtungen konzipierte Transferentropie zeichnet sich gerade durch eine besonders hohe Rauschtoleranz gegenüber anderen Maßen aus.
Ziel der vorliegenden Arbeit ist es, zwei Limitationen, die die Interpretierbarkeit der Charakterisierungen mit der bisher vorgeschlagenen Transferentropie einschränken, zu untersuchen und auszuräumen. Zum einen wird ein Verfahren entwickelt und implementiert, mit dem langreichweitige Korrelationen besser beobachtet werden können, zum anderen werden Korrekturen vorgeschlagen, die den Einfluss so genannter statischer Korrelationen berücksichtigen.
Bei Charakterisierungen von Interaktionsrichtungen mit Hilfe der Transferentropie konnten langreichweitige Korrelationen nur durch die Abschätzung von hochdimensionalen Wahrscheinlichkeitsräumen berücksichtigt werden. Für diese Abschätzung sind sehr viele Datenpunkte innerhalb des Beobachtungsintervalls notwendig, was bei Felddaten, gemessen an unbekannten Systemen, mit der Annahme der Stationarität in einem Beobachtungsintervall konkurriert. Um diese Beschränkung zu umgehen, wird in dieser Dissertation eine Verallgemeinerung des Konzepts der Entropie im Sinne von Lempel-Ziv auf das Maß der Transferentropie übertragen. Hierdurch können langreichweitige Korrelationen ohne die Abschätzung eines hochdimensionalen Wahrscheinlichkeitsraums bestimmt werden.
Zeitgleiche Korrelationen der zugrunde liegenden Signale - so genannte statische Korrelationen - können die Interpretierbarkeit der Charakterisierung einschränken. Zur Berücksichtigung statistischer Korrelationen mit den bisher vorgestellten Maßen war ebenfalls eine mit einem großen Rechenaufwand verbundene Abschätzung hochdimensionaler Wahrscheinlichkeiten notwendig. In der vorliegenden Dissertation wird eine Korrektur der Transferentropie zur Abschätzung der statischen Korrelationen vorgeschlagen, ohne höherdimensionale Terme berechnen zu müssen.
Durch die in dieser Arbeit vorgestellten Maße und Korrekturen kann die Charakterisierung der Interaktionsrichtung verbessert werden. Dabei wird anhand prototypischer Modellsysteme mit chaotischen Dynamiken demonstriert, dass die Charakterisierungen mit Hilfe der vorgeschlagenen Maße und Korrekturen gerade bei Systemen, die ohne Zeitversatz interagieren, besser interpretierbar sind. Weiterhin wurden Interaktionsstärke und Interaktionsrichtung an Zeitreihen hirnelektrischer Aktivität von Epilepsiepatienten bestimmt und mit Charakterisierungen der Transferentropie verglichen. Hierbei lässt sich zusammenfassen, dass sich mit den in dieser Arbeit vorgestellten Maßen Kontraste unterschiedlicher Interaktionsrichtungen besser auflösen lassen.

Contents

1. Introduction **1**

2. Theoretical foundations **7**
2.1. Deterministic approach to dynamical systems 7
2.1.1. Continuous and discrete dynamical systems 7
2.1.2. Stability of dynamical systems 8
2.1.3. State space reconstruction and nonlinear time series analysis 9
2.1.4. Characterizing chaotic behavior in nonlinear dynamical systems . . . 10
2.2. Stochastic approach to dynamical systems 11
2.2.1. Random variables and stochastic processes 12
2.2.2. Random variables in the state space of dynamical systems 17
2.2.3. Characterization of dynamical systems with Kolmogorov-Sinai entropy 18
2.3. Symbolic representation of dynamical systems 19
2.3.1. Application of symbolic dynamics to time series analysis 20
2.4. Kolmogorov complexity and data compression 22
2.4.1. Lempel-Ziv complexity 23
2.4.2. Lempel-Ziv complexity for multivariate data analysis 24

3. Estimating Kolmogorov-Sinai entropy of chaotic dynamical system **27**
3.1. Symbolic representation of tent map 28
3.1.1. Symbolic representation with permutation partition 29
3.1.2. Symbolic representation with threshold-crossing partition 33

4. Characterization of interactions in dynamical systems **39**
4.1. Characterizing strength of interactions with symbolic mutual information . 40
4.2. Characterizing directionality of interactions with symbolic transfer entropy . 45
4.2.1. Corrected symbolic transfer entropy 46
4.2.2. Entropy transfer between time series of dynamical model systems . . 51
4.2.3. Entropy transfer between noise-contaminated time series 58
4.3. Directional interactions in multivariate time series 63

5. Characterizing interactions in electroencephalograms of epilepsy patients **73**
5.1. Epilepsy and electrical activity of the epileptic brain 73
5.2. Characterizing strength of interactions in electroencephalographic recordings 76
5.3. Characterizing directions of interactions in electroencephalographic recordings 80

6. Estimating entropy transfer between dynamical systems exhibiting long-term memories **93**

6.1. Directional interactions between Rössler oscillators 93
6.2. Conditional LZ-complexity and algorithmic transfer entropy 98
6.3. Estimating algorithmic transfer entropy between model dynamical systems . 101

7. Summary and outlook **113**

A. Appendix **117**
A.1. Deterministic dynamical model systems 117
A.2. Mean phase coherence 121
A.3. Entropy of random variables and stochastic processes 122
A.4. Measuring symbolic transfer entropy 125

1. Introduction

In experiments, one is often interested in testing some hypothesis or making inferences on the basis of temporal and/or spatial patterns observed in experimental data. Linear methods of time series analysis provide us a solid toolkit for the characterization of various important properties of experimental data such as dominant frequencies, linear correlations, etc [Ste75, OS09]. All these methods are based on the assumption that the intrinsic dynamics of the investigated system is governed by a linear equation. Since periodic oscillations and an exponential growth/decay are the only possible solutions of linear equations, all irregular behavior in such time series is usually associated with random external input. However, developments in nonlinear dynamical systems theory provided evidence that the irregular behavior may also arise as a result of a chaotic evolution of nonlinear dynamical systems with purely deterministic equations of motion (see e.g. [ER85, EFS98]). This theoretical finding stimulated the development of the quickly growing field of nonlinear time series analysis [KS03]. A group of nonlinear time series analysis methods associated with the symbolization of experimental time series is often referred to the symbolic time series analysis [Bl89, EFS98]. One of the main steps of the symbolic analysis includes the construction of coarse-grained or symbolic representation of raw data. In this representation a real-valued time series is transformed into a sequence of positive integers which are usually called symbols. The resulting symbol sequence is then treated as a representation of the original time series which retains much of the important temporal information. In general, there are several practical advantages of using such a symbolic representation of data. From an experimental point of view such a representation can provide us a computationally efficient and robust against noise way to deal with experimental data. From a theoretical point of view a symbolic representation of data allows us to directly apply a plethora of information-theoretic methods to characterize interactions between underlying dynamical systems.

It is known that the existence of dynamical instabilities and irregular time evolution is a main signature of chaotic dynamical systems [ER85, Ott94, EFS98, Str01, KS03]. In contrast to linear dynamical systems, where small causes always lead to small effects, a tiny uncertainty about the initial state of a chaotic system can lead to the unpredictability of its future states despite a deterministic time evolution. This phenomenon cannot be reliably captured by methods of linear time series analysis. A variety of approaches to characterize such dynamical instabilities of chaotic systems have been developed in the field of nonlinear time series analysis [HKS99, KS03]. For instance, symbolic time series analysis provides a quantitative approach to this problem. It allows us to address the question "How much information do we gain, on average, about the future state of the system by observing its present and entire past?". The theoretical investigations of this question led to the development of the notion of *Kolmogorov-Sinai* (KS) entropy which pro-

vides a measure for the amount of uncertainty generated by a dynamical system per time unit [Kol59, Sin59, CGG89]. The development of nonlinear time series analysis provided several robust and reliable methods to estimate this important characteristic of nonlinear dynamical systems from experimental time series [GP83a, KS03]. One of these methods was developed within the framework of symbolic time series analysis. A symbolic representation of experimental time series allows us to estimate KS-entropy of an underlying dynamical system by applying an information-theoretic toolkit developed by Claude Shannon in his seminal paper on the mathematical theory of communication [Sha48]. Shannon introduced two measures of uncertainty associated with either a random variable or a stochastic process. The first information-theoretic measure is nowadays referred as *Shannon entropy* and characterizes the average amount of information that is gained during the measurement of a single realization of this variable. The second measure is referred as *Shannon entropy rate* and characterizes the average amount of information that is produced by a stochastic process per time unit [Hon02]. If we assume that a symbolic representation of some real-valued time series exhibits a series of realizations of some random variable then the Shannon entropy provides an estimate for KS-entropy of the underlying dynamical system. To numerically compute Shannon entropy of some random variable one has to estimate an empirical probability distribution which is usually defined as the relative frequency of occurrence of different symbols. In real-world applications, an experimental time series may, in general, exhibit long-term memories (i.e., long-term temporal correlations) such that its symbolic representation cannot be represented as a series of realizations of some random variable but has to be represented as a single realization of some high-order Markov process [EFS98]. In this case, KS-entropy of an underlying dynamical system has to be approximated with the Shannon entropy rate of a corresponding order. A numerical analysis of Shannon entropy rates of high orders requires an estimation of high-dimensional empirical probability distributions and therefore demands a large amount of data that is not always available in real-world applications. An insufficient amount of data may lead to an undersampling of empirical probability distributions and, as a result, to significant statistical and systematic errors of obtained estimates of the Shannon entropy rate [Gra88, HSE94, SG96, Rou99]. In many applications one can, however, neglect the influence of long-term temporal correlations in data and approximate KS-entropy of an underlying dynamical system with a low-order estimator of the Shannon entropy rate. A complementary approach to estimate the entropy rate of a stochastic process has been developed within the framework of algorithmic information theory [CGG89, CT91, EFS98, LV08]. This approach is based on the notions of algorithmic and Lempel-Ziv complexities of a symbol series [LZ76, ZL77]. In contrast to the Shannon entropy rate the algorithmic approach does not require the estimation of empirical probability distributions and thus may provide an advantage for the estimation of the entropy rate in experimental data exhibiting long-term temporal correlations.

In general, the problem of deriving a symbolic representation of experimental data is usually application specific and yet lacks a generally acceptable solution [BSLZ01, DFT03]. The most explicit way for the symbolization of experimental data involves an equidistant partitioning of the dynamical range of observables into a finite number of intervals. By labeling each interval with a specific symbol allows us to transform a real-valued time series into a sequence of symbols and thus to obtain a symbolic representation of data. In general,

equidistant partitioning is not always optimal and has to be modified for each application. An alternative way for the symbolization of real-valued time series has been proposed in [BP02] where the authors introduced the concept of permutation symbols representing high-order differences between sequential measurements. Further theoretical investigations [AKK05, AK07] of this symbolic representation showed that the Shannon entropy of permutation symbols (which, according to [BP02], is referred to as permutation entropy rate) obtained from a real-valued time series is indeed related to KS-entropy of an underlying dynamical system. However, these investigations also demonstrated that the permutation entropy rate is only asymptotically (as the number of permutation symbols goes to infinity) related to KS-entropy of a dynamical system.

As it was mentioned above, a nonlinear dynamical system can generate entropy at a nonzero rate that is quantified by KS-entropy. For a dynamical system consisting of several components, an important information on its internal structure can be obtained by measuring to which extent the individual components generate and exchange entropy among each other. In the context of time series analysis the task of inferring causal or directional interactions between several components of a dynamical system (or between several dynamical systems) from experimental time series is a very challenging and important scientific problem. The existence of directional interactions between two dynamical systems can usually be identified by the presence of correlations between a past (present) state of the first system and a future state of the second system, correspondingly. Such correlations are usually referred to as a class of *dynamic correlations* because they reflect the dynamical structure (or evolution) of both systems. However, in many real-world applications, experimental time series can also be correlated in such a way that the present states of two dynamical systems appear to be functionally related to each other. In contrast to dynamic correlations, these correlations do not reflect the dynamical structure of the systems and only characterize the similarity between time series. Following [Sch00] such correlations can be called *static correlations*. As was originally pointed out in [Gra01] and then quantitatively addressed in [Sch00] the existence of such correlations in experimental data can lead to an incorrect inference of the directionality of interactions between two dynamical systems.

A more strict definition of causality of interactions has been introduced by Granger [Gra01]. In his work Granger proposed a list of restrictions which the notion of causality should fulfill to be logically consistent. According to his definition two events are considered as causally interconnected if the forecast error of the first (second) event can be reduced when the knowledge about the outcome of the second (first) event is taken into account. To provide a mathematical definition of causality Granger exploited the framework of autoregressive processes. This approach is nowadays widely used to infer the directionality of interactions in experimental data [BKK04, DCB06]. Recent findings demonstrated that the notion of causality can also be formulated under the information-theoretic framework. The notion of transfer entropy was formulated in [Sch00] (see also [PV07]) as a measure of entropy transfer between two joint stochastic processes. Recent theoretical analysis indicated that the notions of transfer entropy and Granger causality are closely related and in some cases are equivalent with each other [BBS09]. Thus, the estimation of entropy transfer between several dynamical systems allows one to characterize the directionality of interactions between them. Transfer entropy between two time series is usually estimated by using a

so called kernel estimator [Sch00, HSPVB07] or, as an alternative, by using the recently proposed symbolic transfer entropy [SL08]. As it was already mentioned above, permutation entropy rate introduced in [BP02] allows one to estimate KS-entropy of a dynamical system and therefore to characterize the amount of entropy produced by this system per time unit. The symbolic transfer entropy [SL08] extents the notion of permutation entropy rate and provides an approach to estimate an amount of entropy transfer and thus to infer the directionality of interactions between dynamical systems. Similarly as in the case of a high-order Shannon entropy rate considered above, a numerical analysis of high-order transfer entropies between two dynamical systems requires the estimation of high-dimensional empirical probability distributions. In many real-world applications where the amount of data is limited, this can result in an undersampling of obtained empirical probability distributions and therefore in significant statistical and systematic errors of estimated values of transfer entropy. As was originally pointed out in [Sch00], for most practical applications the entropy transfer between two dynamical systems can only be estimated by using a first-order estimator of transfer entropy. In this case, the influence of static correlations and long-term dynamic correlations in experimental data cannot be completely taken into account. Nevertheless, as it was shown in [SL08, SL09], the application of the first-order estimator of entropy transfer allows one to correctly characterize the directionality of interactions between different dynamical model systems as well as in experimental data.

The human brain is a complex network of a vast number of neurons [KSJ00]. The neurons are intrinsically nonlinear dynamical systems which are capable to generate a variety of patterns of electrical activity. Electroencephalography is an important tool in neuroscientific research and especially in clinical practice to measure the patterns of electrical activities of large populations of neurons at a high temporal resolution. Electroencephalography is used for diagnostic purposes and in the presurgical evaluation of epilepsy patients [EP97]. Epilepsy represents one of the most common neurological disorders and is associated with its cardinal symptom, the epileptic seizure. From neurophysiology it is known that interactions between different brain regions reflect a variety of physiologic and pathophysiologic states of the human brain [KSJ00]. Thus, the analysis of interactions in electroencephalographic recordings of epilepsy patients represents an important and widely growing field in neuroscientific research as well as in clinical practice [Kre99, MLDE00, Buz06, LMO$^+$07, OMWL08, Leh08]. The hippocampus is a neuroanatomical structure which plays an important role in long-term memory and spatial navigation [Eic00]. It is known that damage of the hippocampus can result in anterograde amnesia, i.e., in a loss of the ability to create new memories. In humans, this neuroanatomical structure supports declarative memory formation [FEG$^+$99, KSJ00, Eic00, FKL$^+$01, MFA$^+$05, JW07, WAL$^+$10]. Thus, the analysis of interactions in electroencephalographic recordings of epilepsy patients can be very important for the understanding of mechanisms of long-term memory formation in humans.

The main aim of this thesis is – by estimating the amount of Shannon entropy transfer – to characterize the direction of interactions between dynamical model systems as well as in experimental data. We start in chapter 2 with a detailed mathematical description of main concepts and notions of information and dynamical systems theories. We discuss several important techniques of symbolic time series analysis and present a brief introduction into algorithmic information theory. In chapter 3 we address the question as to how and

to what extent different symbolic representations of nonlinear dynamical systems (in our case chaotic maps) can be used to estimate analytically given KS-entropies of these systems. In chapter 4 we apply an information-theoretic approach to characterize the strength and direction of interactions between dynamical model systems. We introduce and then investigate the symbolic mutual information as a measure to characterize the strength of interactions between dynamical systems. Next, we investigate the problem of inference of the directionality of interactions between dynamical model system by using the symbolic transfer entropy introduced in [SL08]. Additionally, by following [MK02] we investigate the influence of finite sample effect on numerical estimates of the symbolic transfer entropy and introduce the notion of the corrected symbolic transfer entropy. This estimate of entropy transfer extends the notion of the effective transfer entropy proposed in [MK02] and allows one to additionally correct the influence of static correlations between time series. In chapter 6 we investigate the problem of entropy transfer between time series exhibiting long-term memories (i.e., long-term temporal correlations). We introduce a complementary approach to estimate high-order entropy transfer by extending the notion of mutual LZ-complexity proposed in [ZRB05]. We then demonstrate that – based on the concept of Lempel-Ziv complexity – this approach does not require the reconstruction of high-dimensional empirical probability functions and can directly be applied to estimate high-order transfer entropies in time series of, in general, arbitrary lengths. In chapter 5, we first characterize the strength of interactions between multi-channel electroencephalographic recordings of epilepsy patients with focal epilepsies undergoing the presurgical diagnostics. For this purpose we apply the symbolic mutual information introduced in chapter 4. Next, by estimating the symbolic transfer entropy as well as the corrected symbolic transfer entropy between different electroencephalographic recordings we investigate the problem of inferring the directionality of interactions in the hippocampus. Finally, we compare the values of both measures of transfer entropy with each other and briefly discuss obtained results.

2. Theoretical foundations

2.1. Deterministic approach to dynamical systems

2.1.1. Continuous and discrete dynamical systems

It is well known that the notion of dynamical systems first appeared when Newton introduced the concept of ordinary differential equations to describe the time evolution of natural phenomena. Nowadays, it is accepted that many physical systems can be described in terms of state variables $\vec{x} = (x_1, ..., x_n) \in \mathbf{R}^n$ whose time evolution is determined by first-order ordinary differential equations [Lan88, Ott94, Str01, KS03]. Thus, a deterministic dynamical system is defined as a set of state variables or a state vector $\vec{x}(t)$ that describes the state of a system at some time instant t and a dynamical law that governs the temporal evolution of corresponding state variables.

$$d\vec{x}(t)/dt = \vec{f}(\vec{x}, \vec{\mu}, t) \tag{2.1}$$

Dynamical systems are called deterministic if there is a unique successive state to every initial state $\vec{x}_0 = \vec{x}(t_0)$. The function $\vec{f}$ can be either linear or nonlinear. In the former case a corresponding dynamical system is called linear, and nonlinear otherwise. The vector $\vec{\mu}$ comprises a set of control or bifurcation parameters. Dynamical systems are called stationary if $\vec{f}(\vec{x}, \vec{\mu}, t) \equiv \vec{f}(\vec{x}, \vec{\mu})$ and non-stationary otherwise.
In order to achieve some insight into qualitative properties of a system - without solving the differential equations - it is sometimes convenient to represent its dynamics in some abstract n-dimensional state space, where each dimension corresponds to a particular state variable. Thus, each point in state space specifies the state of a system and vice versa. As the system evolves in time - guided by the velocity field $\vec{v} = \vec{f}(\vec{x}, \vec{\mu}, t)$ - it produces a trace or a trajectory $\vec{x}(t)$ in state space. Then we can study the dynamics of the system by analyzing the dynamics of the system's trajectory in the state space. A dynamical system is called dissipative when a state space volume that contains all possible initial conditions $\vec{x}_0$ will be contracted in the course of time, otherwise it is called conservative. In dissipative systems all trajectories will eventually (after some transient time) approach and then remain within some subset of points in state space. If this subset is invariant under the dynamical evolution law then it is called attractor of the system. A subset of all initial conditions that asymptotically leads to the same attractor is called its basin of attraction. Suppose a trajectory is approaching a fixed-point $\vec{x}_*$ such that $\vec{v}(\vec{x}_*) = 0$. This point corresponds to an equilibrium state of the system. If small disturbances in the vicinity of $\vec{x}_*$ are damping out in the course of time, then this point is called a stable fixed-point attractor. Self-sustained oscillations of a dynamical system can be observed as a closed loop or limit cycle

attractor. As in the case of a stable fixed-point, small disturbances in the vicinity of stable limit cycle will be suppressed. Limit cycles are inherently nonlinear phenomena and can only be observed in systems with at least two state variables. Attractors with a more complicated geometry and topology are also possible, ranging from an n-torus to fractal structures. The latter is called a strange attractor and corresponds to a deterministic chaotic dynamics whereas the former one represents quasi-periodic oscillations. Attractors of dissipative chaotic systems generally have a very complicated structure with a non-integer dimension that is a paramount signature of the fractality of the set. Thus, the knowledge of the geometry and topology of the attractor (e.g. its dimension) provides a deeper understanding of the nature of the underlying dynamical system.
Deterministic dynamical systems with discrete time are defined by a recursive relation or by a map

$$\vec{x}_{n+1} = \vec{f}(\vec{x}_n, \vec{\mu}) \tag{2.2}$$

where the state vector at time step $\vec{x}_n$ is mapped onto the state vector at time $\vec{x}_{n+1}$ ($n \in \mathbf{N}$). Thus, by setting some initial state vector $\vec{x}_0$ it is possible - by iteration - to compute the state vector $\vec{x}_n$ at an arbitrary time step in the future. In analogy with continuous time systems the equilibrium state $\vec{x}_{n_*}$ is defined as $\vec{x}_{n_*} = \vec{f}(\vec{x}_{n_*}, \vec{\mu})$ as well as the periodic dynamics (e.g. with period-T) is defined as $\vec{x}_{n+T} = \vec{f}(\vec{x}_n, \vec{\mu})$ for any n. Note, that the mapping function $\vec{f}$ might not necessarily be bijective and thus the reverse mapping of $\vec{x}_{n+1}$ into $\vec{x}_n$ is not always uniquely defined. This results in partial information loss about the state of the system during one forward iteration and thus in a rather complex or chaotic dynamical behavior, i.e., unpredictability of the future despite a deterministic time evolution.

2.1.2. Stability of dynamical systems

Stability is a central issue in the analysis of the temporal evolution of dynamical systems [KS03, Ott94, MP00]. It is natural to ask whether a small change in the initial conditions of a dynamical system will asymptotically lead to similar behavior or whether two nearby segments of the trajectory in the state space will remain in the neighborhood of each other in the course of time. Formally, an equilibrium point $\vec{x}_e$ of a dynamical system in the state space is called Lyapunov stable if for any $\epsilon > 0$ there exist $\delta \equiv \delta(e) > 0$ such that, if $\|\vec{x}_e - \vec{x}(t_0)\| < \delta$, then $\|\vec{x}_e - \vec{x}(t)\| < \epsilon$ for any $t > t_0$. The Lyapunov characteristic exponent λ of a dynamical system is a quantity that characterizes the rate of separation of infinitely close segments $\vec{x}_1$ and $\vec{x}_2$ of the trajectory in the course of time, i.e., $\delta x(t) = e^{\lambda t}\delta x(t_0)$, where $\delta x = \|\vec{x}_1 - \vec{x}_2\|$. Thus, a positive value of the Lyapunov exponent characterizes the exponential divergence or instability in the state space of dynamical systems. In general, there exists a whole spectrum of characteristic exponents that are as many as the dimensionality of the state space. Thus, the dynamical system can exhibit instability only along some directions of the state space but can remain stable along another directions. The existence of at least a single positive Lyapunov exponent indicates a chaotic behavior in the temporal evolution of dynamical systems. In this situation even a tiny uncertainty about the initial state of the system will grow exponentially fast in time. The

resulting long-term unpredictability is a consequence of the inherent instability in chaotic dynamical systems and this property is called sensitive dependence on initial conditions.
Another important property of dynamical systems is their structural stability, i.e., their robustness against slight changes of control parameters $\vec{\mu}$. In fact, many natural systems are not structurally stable, and even tiny perturbations of control parameters will lead to a topological and geometrical change of the attractor. This results in drastic changes in the asymptotic behavior of a dynamical system. Critical values of control parameters $\vec{a}_c$, at which qualitative changes of the system's attractors occur, are called *bifurcation points* [MP00, Izh07]. Thus, structurally unstable dynamical systems should reside in the vicinity of a bifurcation point. As an example, the transition from an initial equilibrium state to self-sustained oscillations can be regarded as a process in which, due to changes in the control parameters, a stable fixed-point attractor loses its stability while some limit cycle suddenly becomes stable.

2.1.3. State space reconstruction and nonlinear time series analysis

The spectrum of characteristic exponents and the dimension of the attractor of the dynamical system are dynamic invariants, i.e., their values are invariant under smooth transformations of the state space. This important property guarantees that estimates of these quantities should remain constant irrespective of the details of the measurement process and of the reconstruction of the state space making these invariants very attractive for practical applications. Based on these considerations various algorithms for the estimation of these quantities from experimental data have been proposed[KS03, MP00]. However, in order to proceed with the analysis of nonlinear dynamical systems, one has to first deal with a problem of a reconstruction of the state space from experimental time series $x(t) = M(\vec{X}(t))$. M is called a measurement function and is a projection of the original state space X onto some manifold (e.g. Cartesian axes). But, what is a reconstruction? Formally, an attractor Γ is reconstructed if we can build a vector space U with a topologically equivalent attractor Γ_e. Indeed, it is well known that a reconstruction of the original state space is not necessary and it is often sufficient to construct a new space with an attractor that is topologically equivalent to the original one. To do this one has to define a smooth map $\phi : \phi(\Gamma) \subset U$ such that ϕ is a diffeomorphism[1] between Γ and $\phi(\Gamma)$. Such a map ϕ is called an embedding of a manifold Γ into the space U. It is still not clear how the space U can be constructed and how to find an appropriate embedding. The solution for this problem was found by Takens. According to the *Takens theorem* the reconstructed space U can be formed with the help of so called delay vectors $\vec{X}_d(t) = (x(t), x(t-\tau), \ldots, x(t-(m-1)\tau))$ where m is called embedding dimension and τ is a time delay [Tak81]. This theorem is based on results of Whitney which state that the embedding dimension $m = 2d + 1$, where d is the dimension of the attractor (see [KS03] and references therein). In general, the value of the time delay τ is not the subject of the Takens theorem but a proper choice of this parameter is quite important. It is important to point out that the nonlinearity of the measurement function

[1] A diffeomorphism is an invertible function that maps a differentiable manifold onto another, such that both the function and its inverse are smooth.

M (i.e., if the function M is not homeomorphic[2]) can lead to partial loss of information about the topology of the attractor.

2.1.4. Characterizing chaotic behavior in nonlinear dynamical systems

The invention of the delay embedding technique has drastically enlarged the field of nonlinear time series analysis and especially numerical analysis of nonlinear dynamical systems. It was demonstrated that various natural processes which were before considered as stochastic ones can indeed satisfactorily be described as deterministic chaotic dynamical systems [KS03, HKS99, MP00]. The dynamical side of deterministic chaos manifests itself in the sensitive dependence of time evolution on initial conditions. Numerical estimates of the spectrum of Lyapunov exponents provide nowadays a standard toolkit to characterize and quantify the degree of deterministic chaos in dynamical systems and time series. The existence of at least a single positive Lyapunov exponent indicates chaotic behavior in temporal evolution of dynamical systems. The presence of dissipation in many macroscopic systems can lead to the situation where the state vector of the system will eventually settle within a low dimensional attractor of an, in generally, high dimensional state space. Thus, an estimate for the attractor dimension reconstructed from a time series can provide a good measure for its complexity. The generalized correlation integral provides nowadays a standard way to estimate the dimension of a set in high dimensional spaces [KS03]. The estimation of the correlation integral for various nonlinear dynamical systems demonstrated the existence of attractors of non-integer (or fractal) dimensions, i.e., a geometric signature of chaotic dynamics. In the following we provide several important examples of nonlinear dynamical systems exhibiting chaotic behavior.

The Hénon map is a prototypical two-dimensional invertible iterated map proposed by Hénon [H76]. For the canonical set of parameters (see section A.1) this map is chaotic. With this choice of parameters, a set of initial conditions will be asymptotically mapped into a subset of points of a two dimensional plane known as the Hénon attractor. This attractor is a fractal set and has a non-integer dimension. The Lorenz oscillator is a three-dimensional dynamical system that exhibits a chaotic flow. This system of differential equations was introduced by Lorenz to describe convection rolls arising in the atmosphere [Lor63]. The geometrical representation of the Lorenz system in state space is the well known Lorenz attractor. Similarly to the Hénon attractor, the Lorenz attractor is also a fractal structure and has a non-integer dimension. Table 2.1 contains numerical values of the correlation dimensions and Lyapunov spectra for the Hénon and Lorenz systems [GP83b, Gra83a, Gra83b, WSSV85, SS85].

The computation of the Lyapunov spectrum allows to perform a cross-check for the value of the fractal dimension of the attractor. Kaplan and Yorke conjectured that values of the Lyapunov spectrum and the fractal dimension are closely related (see [KS03] and references therein). Consequently, the Kaplan-Yorke (KY) dimension is given by

[2]A function is called homeomorphic if it is bijective (one-to-one), continuous and has a continuous inverse. It is known that a homeomorphic function preserves the topology of a set [MP00].

System	Parameter values	Corr./KY dimension	Lyapunov spectrum
Hénon map	$a = 1.4$ $b = 0.3$	$D_{corr} = 1.21$ $D_{KY} = 1.26$	$\lambda_1 = 0.61$ $\lambda_2 = -2.30$ $[bit/iter.]$
Lorenz oscillator	$\sigma = 16.0$ $R = 45.92$ $b = 4.0$	$D_{corr} = 2.06$ $D_{KY} = 2.07$	$\lambda_1 = 2.11$ $\lambda_2 = 0.00$ $\lambda_3 = -27.41$ $[bit/s.]$

Table 2.1.: Lyapunov spectra as well as correlation D_{corr} and Kaplan-Yorke D_{KY} dimensions of the Hénon map and the Lorenz oscillator.

$$D_{KY} = k + \frac{\sum_{i=1}^{k} \lambda_i}{|\lambda_{k+1}|} \tag{2.3}$$

where k is the maximum integer such that the sum of the k largest Lyapunov exponents is still non-negative. The KY-dimension, in many cases, is used to test the consistency of estimated values of the correlation dimension (for further details see [KS03, HKS99]).

2.2. Stochastic approach to dynamical systems

In many cases dynamical systems are too complex to be described in terms of a system of differential equations or maps. Various physical systems in nature possess a very large number of degrees of freedom and thus cannot be adequately analyzed by directly solving differential equations that describe the time evolution of their constituents. For example, a gas at normal conditions contains about $N_L = 2.43 \times 10^{16}$ particles per cubic millimeter. This obvious limitation of the mechanical view on macroscopic systems has led to the development of statistical physics that is a branch of physics which describes emergent properties[3] of such systems upon microscopic (atomic) dynamics [Lan64].
As was observed by Brown, the microscopic motion of particles suspended in a fluid is erratic. This observation led to the concept of Brownian motion that is one of the important concepts in Statistical Physics. Succeeding theoretical investigations (Einstein, Smoluchowski, Langevin) have led to the development of a stochastic approach for the description of this macroscopic phenomenon [Ein05, Lan08]. With this approach one tries to dispense from a pure deterministic description of the system's evolution in state space and, additionally, to rely on probabilistic concepts. These studies, among others, have shown that in many cases a dynamical system cannot be considered as completely isolated and the influence of external perturbations has to be taken into account. Such perturbations are then introduced by adding random terms into the equations of motion. Generally speak-

[3]Emergent properties of macroscopic physical systems (temperature, pressure, etc.) are typical for many-body systems. They result from the interaction of their constituents and do not exist on the microscopic level.

ing, a dynamical system is called stochastic if for any initial state there is a probability distribution of its possible successive states.

2.2.1. Random variables and stochastic processes

Discrete and continuous random variables

A fundamental concept for any statistical treatment is that of a random variable. We speak of the probability of an event, i.e., a number is assigned to each event such that this number represents the probability of this event to take place. In 1933 A.N. Kolmogorov formulated a complete system of axioms for a mathematical definition of probability [Hon02, vK07]. Let Ω be a set of all outcomes of a (possibly imaginary) experiment and $\omega \in \Omega$ is a single realization of this experiment. To each event A in the space of events $\mathbf{B}$ (the so called Borel space [4]) we assign some real number $P(A)$ that defines the probability of this event. This assignment possesses the following properties:

- $0 \leq P(A) \leq 1$ for all $A \in \mathbf{B}$,
- $P(\Omega) = 1$,
- Let $A_i \in \mathbf{B}$ be a partitioning of the space $\mathbf{B}$ by countably many disjoint sets $A_i \cap A_j = \emptyset$ for $i \neq j$, then $P(\cup A_i) = \sum P(A_i)$.

A discrete random variable is thus defined as a collection of possible elementary events together with their probabilities $\xi \equiv (a_1, a_2, \ldots; P(a_1), P(a_2), \ldots)$. A realization of a random variable is given by one of the elementary outcomes that is produced with the probability which has been assigned to this event.
The concept of a random variable can be extended to the case where the Borel space $\mathbf{B}$ contains intervals and points on the real axis $x \in \mathbf{R}$. The probability for a real number x to be enclosed in the interval $[x_1, x_2]$ is defined as

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} g(x)\, dx \tag{2.4}$$

where $g(x)$ is called the probability density function, i.e., for small enough dx, $g(x)dx$ is the probability of the event $x \in (x, x + dx)$. In a similar way one defines a d-dimensional random vector $\vec{x} \equiv (x_1, \ldots, x_d) \in \mathbf{R}^d$. In this case the Borel space $\mathbf{B}$ contains points and, in general, arbitrary domains in $\mathbf{R}^d$. Thus, the probability for each component of a vector $\vec{x} \equiv (x, y, \ldots)$ to be enclosed in the intervals $x \in (x_1, x_2), y \in (y_1, y_2), \ldots$ is defined as

$$P(x_1 \leq x \leq x_2; y_1 \leq y \leq y_2; \ldots) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} \ldots g(x, y, \ldots)\, dxdy \ldots \tag{2.5}$$

The probability density $g(x, y, \ldots)$ is now a function on $\mathbf{R}^d$ and is called the joined probability density function for random variables $X, Y, \ldots$. By integrating over some random variables of a given joint probability density function one obtains the so called marginal

[4] The Borel space is a set $\mathbf{B}$ such that $\Omega \in \mathbf{B}$, if $A \in \mathbf{B}$ then $\bar{A} \in \mathbf{B}$, and if $A_1, A_2 \in \mathbf{B}$ then $A_1 \cup A_2 \in \mathbf{B}$

probability density function, i.e., the probability density function of the remaining random variables. For instance $g(x) = \int\int \ldots dydz \ldots g(x, y, z, \ldots)$.
In probability theory one defines $A|B$ for an event A which is taking place under the condition that B is given. By definition, the conditional density function for the random variable X under the condition that the outcome of Y has been given is defined as $g(x|y) = g(x, y)/g(y)$. We say that a random variable X is statistically independent of Y if and only if $g(x, y) = g(x)g(y)$ and thus $g(x|y) = g(x)$.

Stochastic processes

Stochastic processes provide an important mathematical framework to describe various dynamical processes that cannot be adequately represented by a sequence of independent and identically distributed (i.i.d.) random variables [Hon02, vK07]. If one considers a random variable which depends on time, one is led to the concept of a stochastic process. For simplicity, let us suppose that the time flow is discrete. A stochastic process is then a sequence of discrete random variables, i.e., $\{Z(t_i)\}_{i=1}^{N}$. Thus, a stochastic process $Z(t)$ within the time interval $[t_1, \ldots, t_N]$ is uniquely defined with the density function $g(z(t_1), \ldots, z(t_N))$. In general, the state of a stochastic process at time step t_n, depends on its complete history $g(z(t_n)|z(t_{n-1}), \ldots, z(t_0))$. In many practical situations it is however quite reasonable to assume that a stochastic process has a limited memory. In this case its state at time step t_n is independent on its complete history and is uniquely defined when only the state of the process at time step t_{n-1} is given, i.e., $g(z(t_n)|z(t_{n-1}), \ldots, z(t_0)) = g(z(t_n)|z(t_{n-1}))$. A process with this property is called Markov process (chain) of first order [5]. The Markov property is defined as an idealization that allows one to completely specify a stochastic process by the density function for the state of the process at the initial time $g(z(t_0))$ along with the conditional density $g(z(t_n)|z(t_{n-1}))$ which is also called transitional density function. The stochastic process is called stationary if its transitional density function does not depend on time. Thus, for a stationary Markov process of first order we have

$$g(z(t_n), z(t_{n-1}), \ldots, z(t_0)) = g(z(t_n)|z(t_{n-1}))g(z(t_{n-1})|z(t_{n-2})) \ldots g(z(t_1)|z(t_0))g(z(t_0)).$$

In the case of a sequence of i.i.d. random variables, it is obvious that the transitional density function is reduced to $g(z(t_n)|z(t_{n-1})) = g(z(t_n))$. If all properties of a stationary stochastic process (e.g. its statistical moments of arbitrary orders) are identical when computed by either time or ensemble averages then this process is called *ergodic*.

Entropy of discrete and continuous random variables

Entropy is a key concept associated with a random variable. Let ξ be a discrete random variable with a set of N possible realizations $\{a_i\}_{i=1}^{N}$ occurring with probabilities $p(a_i)$. Entropy of a random variable is defined as

[5] It is straightforward to generalize this definition to the Markov process of order k as $g(z(t_n)|z(t_{n-1}), \ldots, z(t_0)) = g(z(t_n)|z(t_{n-1}), \ldots, z(t_{n-k}))$.

$$H(\xi) = -\sum_{i=1}^{N} p(a_i)\log(p(a_i)). \tag{2.6}$$

Entropy provides a measure of average information that is gained during the measurement of a single realization of the random variable ξ [Sha48, CT91]. If the base of the logarithm is 2, entropy is measured in bits. Moreover, the proper definition of entropy requires the condition $0\log(0) = 0$ to be fulfilled.
The notion of entropy can also be defined for continuous random variables. Let us consider some continuous random variable ξ that is defined by the probability density function $g(a)$. The differential entropy $h(\xi)$ of a continuous random variable ξ is defined as

$$h(\xi) = -\int_S g(a)\log(g(a))da, \tag{2.7}$$

where S is a support set of ξ (i.e., a set $\{a\}$ with $g(a) > 0$). It is easy to see that both notions of entropy for discrete and continuous random variables are related. The differential entropy can be approximated with the discrete entropy. To do this, let us divide the support set S in N bins of length Δ. In this way we introduce a new discrete random variable ξ_d that has N realizations $\{a_i\}_{i=1}^{N}$ each occurring with probabilities $p(a_i) = \int_{i\Delta}^{(i+1)\Delta} g(a)da$. It can be shown [CT91] that the differential entropy $h(\xi)$ and discrete entropy $H(\xi_d)$ are asymptotically ($\Delta \to 0$) related :

$$H(\xi_d) = h(\xi) + log(1/\Delta). \tag{2.8}$$

This relation indicates the difference between discrete and differential entropies. The discrete entropy of a continuous random variable is not uniquely defined and moreover it diverges with vanishing bin size ($\Delta \to 0$). Thus, in order to specify, with an infinite precision, a realization of a continuous random variable, on average, an infinite amount of information has to be provided [6].

Entropy of multivariate random variables

The definition of entropy can be extended to the case of two or more random variables (see [CT91, Mat00] and section A.3). The joined entropy of two random variables ξ and η with a joined probability distribution $p(a, b)$ is given by

$$H(\xi, \eta) = -\sum_{(i,j)=1}^{N} p(a_i, b_i)\log(p(a_i, b_i)) \tag{2.9}$$

or alternatively

$$H(\xi, \eta) = H(\xi) + H(\eta) - I(\xi, \eta). \tag{2.10}$$

Here, $I(\xi, \eta)$ denotes the so called *mutual information* between ξ and η. If the random variables ξ and η are statistically independent ($p(a, b) = p(a)p(b)$) then from the definition

[6] This property has important consequences in dynamical systems theory (see section 2.2.2).

of the entropy (see Eq. 2.9) it follows that $H(\xi,\eta) = H(\xi)+H(\eta)$ and thus $I(\xi,\eta) = 0$. The conditional entropy of the random variable ξ with a given realization of the random variable η is defined as $H(\xi|\eta) = H(\xi,\eta) - H(\eta)$ and analogically for $H(\eta|\xi)$. It is easy to see that if the random variables ξ and η are statistically independent then $H(\xi|\eta) = H(\xi)$ as well as $H(\eta|\xi) = H(\eta)$. If, in contrary, the random variables are statistically interdependent then we can define (see Eq. 2.10) the mutual information between ξ and η as

$$I(\xi,\eta) = H(\xi) + H(\eta) - H(\xi,\eta) = H(\xi) - H(\xi|\eta) = H(\eta) - H(\eta|\xi). \tag{2.11}$$

The mutual information characterizes the degree of interdependence between two random variables. It can be shown (see section A.3) that the mutual information is a symmetric measure of interdependence between two (or in general arbitrary many) random variables, i.e., $I(\xi,\eta) = I(\eta,\xi)$.
The notion of mutual information can be extended to the case of continuous random variables. Let us consider two interdependent continuous random variables ξ and η which are characterized by the joined probability function $g(a,b)$. The mutual information between ξ and η is then given as

$$I(\xi,\eta) = h(\xi) - h(\xi|\eta) = h(\eta) - h(\eta|\xi). \tag{2.12}$$

where $h(\xi)$ and $h(\eta)$ are differential entropies for the variables ξ and η. The conditional differential entropy $h(\xi|\eta)$ is a differential entropy of the variable ξ when the outcome of the variable η is given. The conditional differential entropy $h(\eta|\xi)$ is defined analogous. In contrast to the differential entropy, a discrete approximation of the mutual information between continuous random variables converges to its actual (continuous) value under the refinement of partitions, i.e., with $\Delta \to 0$ [CT91, KS02]. Indeed, according to equation 2.8 the discrete approximation of the mutual information is given as $I(\xi_d,\eta_d) = h(\xi) + log(1/\Delta) - h(\xi|\eta) - log(1/\Delta) = I(\xi,\eta)$ for any partition size Δ.

Entropy rate of stochastic processes

The entropy rate is a key concept associated with a stochastic process [CT91, vK07]. If we have a sequence of N random variables it is natural to ask: how does the joined entropy of the sequence grows with N? From the definition given above, a stochastic process is defined as a time sequence of random variables $\{Z(t_i)\}_{i=1}^{N}$ that is uniquely characterized by a joined probability density function $g(z(t_N),\ldots,z(t_1))$. Formally, the entropy rate of a stochastic process Z is defined as

$$dH(Z) = \lim_{N\to\infty} \frac{1}{N} H(z(t_N),\ldots,z(t_1)), \tag{2.13}$$

or alternatively

$$dH(Z) = \lim_{N\to\infty} H(z(t_N)|z(t_{N-1}),\ldots,z(t_1)). \tag{2.14}$$

Both definitions of the entropy rate provide a measure for an average amount of entropy that is generated by a stochastic process Z per time step. From the definition of the

joined entropy it follows that for a stationary stochastic process $dH(Z) \leq H(z(t_i))$ for any time step t_i. The equality is only reached when a stochastic process is a sequence of i.i.d. random variables. In general, the entropy rate has to be estimated in the limit of infinitely many time steps that reduces the practical applicability of this approach only to the case of i.i.d. variables. However, in many practical situation the problem can adequately be described by an $k-$ order Markov process. In this case the entropy rate can be defined as $dH(Z) = H(z(t_i)|z(t_{i-1}), \ldots, z(t_{i-k}))$.
The notion of the entropy rate $dH(Z)$ was originally introduced by Shannon in his famous work on foundations of information theory[Sha48]. He demonstrated that if we consider a message as a realization of some stochastic process then the entropy rate of this process will quantify to what extent a given message can be "compressed" without loosing any information. Indeed, let us consider some message of length N that is comprised of symbols taking from an alphabet of length A_{symb}. Next, if we define $\hat{N}$ as length of the maximally compressed version of our message then the entropy rate can be estimated as $dH \approx \frac{\hat{N}}{N} \log(A_{symb})$.

Entropy rate of bivariate stochastic processes

For the sake of simplicity, we can start by considering first order stationary Markov processes[7]. In this situation the joint entropy rate of two stochastic processes Z_1 and Z_2 is defined as

$$dH(Z_1, Z_2) = H(z_1(t_i), z_2(t_i)|z_1(t_{i-1}), z_2(t_{i-1})) \tag{2.15}$$

or alternatively (see section A.3)

$$dH(Z_1, Z_2) = dH(Z_1) + dH(Z_2) - TE(Z_1, Z_2) - TE(Z_2, Z_1) - dI(Z_1, Z_2) \tag{2.16}$$

where two-point interactions are taking into account by transfer entropies $TE(Z_1, Z_2)$ and $TE(Z_2, Z_1)$ (defined below) as well as by a mutual information rate $dI(Z_1, Z_2)$. The joint entropy rate $dH(Z_1, Z_2)$ of the stochastic processes provides a measure for the amount of entropy that is generated per time step by a bivariate process (Z_1, Z_2). Similar to the case of the joint entropy for two independent random variables the joint entropy rate of two independent stochastic process is $dH(Z_1, Z_2) = dH(Z_1) + dH(Z_2)$. If the stochastic processes are not independent it is rather useful to define the conditional entropy rates as $dH(Z_1|Z_2) = H(z_1(t_i)|z_1(t_{i-1}), z_2(t_{i-1}))$ and $dH(Z_2|Z_1) = H(z_2(t_i)|z_1(t_{i-1}), z_2(t_{i-1}))$. The conditional entropy rate thus characterizes the amount of net entropy that is generated per time step, for instance in process Z_1 only, i.e., when the state of process Z_2 at time step t_{i-1} is known. The difference between the entropy rate of the process $dH(Z_1)$ and its conditional entropy rate $dH(Z_1|Z_2)$ provides a measure for the amount of entropy per time step that has been transfered from process Z_2 to process Z_1, i.e.,

$$TE(Z_2, Z_1) = dH(Z_1) - dH(Z_1|Z_2) \tag{2.17}$$

[7]The generalization to the case of Markov processes of higher orders is given in the appendix (see section A.3).

and *vice versa* from process Z_1 to process Z_2

$$TE(Z_1, Z_2) = dH(Z_2) - dH(Z_2|Z_1). \tag{2.18}$$

As shown in the appendix (see section A.3) and in [Sch00, KS02, HSPVB07, SL08] the transfer entropy (or conditional mutual information used in [PV07]) is an asymmetric measure of interdependence, i.e., $TE(Z_2, Z_1) \neq TE(Z_1, Z_2)$ and can be used to reveal the asymmetry or direction of interaction between two interacting stochastic processes. In contrast to this, the mutual information rate is – per definition (see section A.3) – symmetric $dI(Z_2, Z_1) = dI(Z_1, Z_2)$ and characterizes the amount of common information (entropy) that is generated per time step between processes Z_1 and Z_2.
The given definition of the transfer entropy can also be applied to continuous stochastic processes. By converting the continuous stochastic variables into discrete representation (using some coarse graining procedure) it is possible to define a discrete approximation of transfer entropy. It is known that a discrete approximation of the mutual information is related and asymptotically converges to its actual (continuous) value [CT91]. However, a similar statement for the transfer entropy has not yet been proven [KS02].

2.2.2. Random variables in the state space of dynamical systems

The sensitive dependence on initial conditions is a main signature of deterministic chaos in dynamical systems. Due to the existence of instability of motion, the system's trajectory tries to explore its state space and sensitively reacts on external disturbances or fluctuations. This entails that probabilistic elements should enter in the description of chaotic dynamical systems in an essential manner. The knowledge of the geometrical and topological properties of the attractor provides a global picture of the long-term behavior of a dynamical system. A more detailed picture is, however, given by the probability density function $\rho(\vec{x})$ that is called probability measure [Lan64, ER85, KS03, Ott94, MP00]. This measure describes how frequently various parts of the state space are visited during the time evolution. Such a definition of a probability measure $\rho(\vec{x})$ allows one to estimate it from experimental time series. To do so, we can cover the state space with a grid of hyper-cubes of size ϵ and then calculate the amount of time $\tau(\epsilon)$ the orbit - starting from some initial condition - spends in every cube and divide it by the total observation time. Formally, for an infinitely small cube ϵ the probability measure is defined as

$$\rho(\epsilon) = \lim_{\epsilon \to 0} \lim_{t \to \infty} \frac{\tau(\epsilon)}{t}. \tag{2.19}$$

Both limits exist only if the dynamical system fulfills the so called ergodic hypothesis [ER85]. According to this hypothesis the system's trajectory eventually visits all points of the attractor in the course of time and the above defined probability measure (Eq. 2.19) does not depend on the choice of initial conditions. Another important property of the probability measure is its time invariance. In the study of complex dynamical systems the invariant ergodic probability measure is playing an important role: it allows one to calculate the statistical average of different observables of the system. Formally, the probability measure $\rho(\vec{x})$ is called the invariant probability measure of a dynamical system if it remains

constant under the time evolution law $\vec{f}(\vec{x})$ (see Eq. 2.1) [8]. Let us select an arbitrary part of the attractor $\Gamma_1 \subseteq \Gamma$ along with its history $\vec{f}^{-t}(\Gamma_1) \subseteq \Gamma$ that is defined as a time-backward transformation of Γ_1 for any $t > 0$. The invariant probability measure has then to satisfy the condition $\rho(\Gamma_1) = \rho(\vec{f}^{-t}\Gamma_1)$ [ER85]. The invariant probability measure that correspond to the stable invariant manifolds (or in other words to attractors) of dynamical systems can be observed in experiments and are called *natural measures*[MP00].

2.2.3. Characterization of dynamical systems with Kolmogorov-Sinai entropy

As was mentioned before, the knowledge of the dimension of the attractor along with the spectrum of Lyapunov exponents provides a set of dynamic invariants to describe and to characterize the global behavior of dynamical systems. Another important dynamic invariant of a deterministic dynamical system is its entropy. One of the first attempts to define an entropy for deterministic dynamical systems was done by Kolmogorov and Sinai [Kol59, Sin59, ER85, KS03]. By using the definition of the probability measure $\rho(\vec{x})$ (see Eq. 2.19) we can again cover the state space with a grid of hypercubes ξ_i of the size ϵ. Now, let us consider the system's trajectory as a stochastic process $\{\xi_i\}_{i=1}^{m}$ and define a probability $p(\xi_1, \xi_2, \ldots, \xi_m)$ that during the time evolution the trajectory will in succession visit these hypercubes. The entropy rate of this process or the Kolmogorov-Sinai entropy (KS-entropy[9]) is then defined as

$$h_{KS} = \lim_{\epsilon \to 0} \lim_{m \to \infty} \frac{1}{m} \sum_{\xi_1, \xi_2, \ldots, \xi_m} p(\xi_1, \xi_2, \ldots, \xi_m) log(p(\xi_1, \xi_2, \ldots, \xi_m)). \tag{2.20}$$

The KS-entropy is an important characteristic invariant of deterministic dynamical systems. It provides a measure for the amount of entropy generated by a system per time unit. The positivity of KS-entropy indicates the existence of dynamical instability, and the inverse absolute value provides an estimate for the time scale that is relevant for the predictability of the system. According to *Pesin's identity* [KS03] the sum of all positive Lyapunov exponents is an upper bound for KS-entropy

$$h_{KS} \leq \sum_{i:\lambda_i > 0} \lambda_i. \tag{2.21}$$

Equation 2.20 provides only an analytical approach to compute the KS-entropy for various dynamical systems. The numerical estimates of this measure were not invented for a long time. The development of nonlinear time series analysis provided several approaches to estimate KS-entropy, e.g. explicitly via the generalized correlation integral [GP83a] or

[8] In general, the time evolution of the probability measure under the evolution law (see Eq. 2.1) is given by the Frobenius-Perron equation $\rho_t(\vec{x}) = \int d\vec{y} \delta(\vec{x} - \vec{f}(\vec{y}, t))\rho_0(\vec{y})$ [ER85].

[9] The KS-entropy provides information about metric aspects of the state space of dynamical systems. For this reason it is also called a metric entropy. The topological aspects of the state space is covered by the so called topological entropy. Both entropies belong to a whole family of dynamic invariants that are called *Renyi entropies* (for more details see [KS03]).

implicitly by estimating values of all positive Lyapunov exponents (see [KS03] for a broad overview of the literature).

2.3. Symbolic representation of dynamical systems

As it is now widely accepted, in many cases there is a possibility to construct the so called symbolic representation of dynamical systems. In this representation one can, in general, provide a more practical definition of KS-entropy. Since Poincaré it is well known that differential equations can be viewed as discrete-time systems. In his analysis of the three-body problem, he proposed that the complex time evolution of this continuous dynamical system could be described by using a stroboscopic sampling of the system's trajectory in a multidimensional state space [Hol90]. To apply this method, which is nowadays known as Poincaré section, one needs first to form a suitably oriented surface in the state space of the system. Next, an invertible map (a stroboscopic map) on this surface is constructed by following the system's trajectory, i.e., the iterates of the map are given by the points where the trajectory intersects the surface in a specified direction. By applying this method Poincaré was able to convert the continuous flow in the state space to a smooth discrete-time mapping.
In addition to the discretization of the time flow the discretization of the state space itself is a natural extension of the concept of a Poincaré section. Indeed, formally, the state of the dynamical system is defined by continuous state variables which are real numbers. This implies that, in general, an infinite amount of information is required to completely (i.e., with an infinite precision) specify the state of a system. In practice, however, due to a finite precision of the measurement process the state of a system can only be specified by a set of discrete variables, i.e., by a finite set of natural numbers. Such a discretization of the state variables[10] results in the segmentation of the state space into a finite number of partitions. Within each partition infinitely many states of a system cannot be distinguished from each other. This observation led to the development of the mathematical discipline of symbolic dynamics. In the symbolic representation of a state space of a system each partition is labeled with some symbol and the time evolution of the dynamical system is represented by an, in general, infinitely long sequence of symbols. Nowadays, symbolics dynamics is an important branch of dynamical system theory and studies the relationships between the continuous and discrete representations of dynamical systems [DFT03, EFS98, Bl89]. One of the main results of symbolic dynamics comprises that the partitioning of the state space can be optimized in such a way to ensure the equivalence between symbolic and continuous representations of the system. In this case one can omit the limit of an infinitely small partition $\epsilon \to 0$ in the definition of KS-entropy (Eq. 2.20). This allows a rather practical definition of KS-entropy and for some partitioning scheme Π we have

$$h_{KS}^{symb} = \sup_{\Pi} \left\{ \lim_{m\to\infty} \frac{1}{m} \sum_{\xi_1,\xi_2,\ldots,\xi_m} p(\xi_1,\xi_2,\ldots,\xi_m) log(p(\xi_1,\xi_2,\ldots,\xi_m)) \right\} \tag{2.22}$$

[10] According to the Takens theorem the state variables can be reconstructed from experimental time series (see section 2.1.3).

where the supremum is taken over all possible partition schemes. However, as was already shown for many dynamical systems it is sometimes possible to define a so called *generating partition* Π_g such that $h_{KS}^{symb}(\Pi_g) = \sup_{\Pi}\left\{h_{KS}^{symb}(\Pi)\right\} = h_{KS}$. A next important class of the partitioning of the state space is defined via the so called *Markov partition* Π_m. In this case the system's trajectory can be represented as a Markov stochastic process that additionally eliminates the second limit in the definition of KS-entropy (Eq. 2.20). Unfortunately, there is up to now no widely accepted universal method for the definition of the generating partition and/or Markov partitions for an arbitrary dynamical system albeit several attempts have been taken to achieve this (see e.g. [BSLZ01, GK85, TK10]).
The symbolic representation of dynamical systems can often be stochastic even when its continuous counterpart is deterministic. Indeed, in case of deterministic dynamical systems with chaotic dynamics the initial uncertainty (due to partitioning of the sate space) about the state of the system will grow in time. In this situation even a deterministically generated trajectory in state space of a system will be represented by a symbolic sequence of stochastic nature (i.e., absence of any deterministic rules to generate this sequence). This results in a positive value of KS-entropy and, due to Pesin's identity (Eq. 2.21), in the existence of positive Lyapunov exponents, i.e., the main signature of deterministic chaos.

2.3.1. Application of symbolic dynamics to time series analysis

A main step of symbolic analysis includes a discretization of the raw data into a sequence of symbols. In the following, we briefly discuss several common methods for the construction of symbolic representation of raw data as well as an alternative approach which is based on the concept of permutation symbols proposed in [BP02].

Explicit partitioning of state space

The most explicit approach for a symbolization of data involves equidistant partitioning of the dynamical range of observables into a finite number of regions[11]. Next, each region is uniquely labeled with a specific symbol. The number of possible symbols defines the *alphabet length* of our symbolic sequence. Varying the size of the alphabet provides a means to select how much of the original information is retained in the resulting symbolic sequence. One of the simplest partitioning schemes is called threshold-crossing or binary partitioning. In this case the dynamical range of the observable is divided into only two partitions. To each of the partition one assigns a symbol and the time evolution of a real-valued time series is thus encoded as a sequence of binary symbols. Being rather useful in many practical applications (see [DFT03]) this partitioning scheme can, in some cases, lead to the incorrect characterization of the underlying dynamical behavior [BSLZ01]. In general, optimal choice of partitions is usually application-specific and ranges from equidistant partitions to equiprobable partitions (partitioning of the dynamical range of observables into the regions of equal probability). In this case the entropy of the generated symbolic

[11] In case of digital recordings the raw data is, from the very beginning, already discretized by an A/D-converter. But this original discretization is usually much more refined than that used in symbolic analysis.

sequence will be maximized which is one of the main requirements for the construction of a generating partition.

Let us now consider a more complicated case where our scalar time series $x(t)$ is generated by a system with n degrees of freedom and represents only a projection of the high dimensional state space ($X \in \mathbf{R}^n$) onto a one-dimensional manifold, i.e., $x(t) = M(\vec{X}(t))$. Here, M is again a measurement function. In this case one has to first apply a time delay embedding (see section 2.1.3) to reconstruct the state space and then perform a partitioning of it. Despite a seemingly simplicity, a practical realization of this approach is rather limited due to an exponential growth of possible partitions with embedding dimension m. Moreover, even in the simplest possible situation, where we divide our state space only in two partitions (i.e., a binary alphabet), the problem of finding an, in general, high-dimensional manifold that separates two generating partitions from each other can be extremely difficult. A detailed review of methods of symbolic time series analysis can be found in [DFT03].

Partitioning of state space by permutations

An alternative approach for the partitioning of a high-dimensional state space has been proposed in [BP02]. To proceed, let us again start with a scalar time series $x(t)$ which is generated by some system with n degrees of freedom. A delay reconstruction in m dimensions is then formed by the delay (embedding) vectors $\vec{X}(t) = (x(t), x(t-\tau), \ldots, x(t-(m-1)\tau))$ where m is the embedding dimension and τ is a time delay. Now, we are seeking for a symbolic representation of the reconstructed state space such that the time evolution of our system (i.e., the system's trajectory) is uniquely mapped onto a sequence of symbols $\pi_t = T(\vec{X}(t))$ with $\vec{X}(t) \in \mathbf{R}^n$ and $\pi_t \in \mathbf{N}^+$. To define the transformation T we first compare values of all m components of the state vector $\vec{X}$ with each other. It is known that there are, in general, $m!$ ways to order m different numbers. Each of $m!$ possible ordering patterns is then labeled with an integer number which, in the following, will be called the permutation symbol $\pi_t \in [1, 2, \ldots, m!]$. For instance, the order pattern $\{x(t) > x(t-\tau) > \ldots > x(t-(m-1)\tau)\}$ is assigned to the symbol $\pi_t = 1$ whereas $\{x(t) < x(t-\tau) > \ldots > x(t-(m-1)\tau)\}$ is assigned to the symbol $\pi_t = 2$, etc.. The resulting transformation T allows one to assign a sequence of permutation symbols to the time evolution of the state vector $\vec{X}(t)$. The number of ordering patterns defines the length of the alphabet A_{perm} of our symbols. The partitioning of the state space by this transformation can be easily demonstrated in the case of a two dimensional embedding $m = 2$. Indeed, in this case the two-dimensional embedding state space is divided in two partitions which are separated by the diagonal line $x(t) - x(t-\tau) = 0$. The definition of partitions for higher embedding dimension $m > 2$ is more complicated but also possible. The number of such partitions (or the alphabet length) will very quickly increase with an embedding dimension, i.e., $A_{perm} = m!$. Thus, we can naively expect that such a symbolic representation might be equivalent to the original continuous representation of the reconstructed dynamical system if $m \to \infty$. Indeed, as was shown recently the permutation symbols can be used to define a so called permutation entropy rate which asymptotically ($m \to \infty$) approaches a metric or KS-entropy for ergodic dynamical systems [BKP02, AKK05, AK07].

2.4. Kolmogorov complexity and data compression

Kolmogorov's interest in randomness and complexity has led him to the development of the notion of *algorithmic complexity* (AC) [Kol65, Kol68, LV08, CGG89, CT91]. In contrast to the notion of entropy rate of stochastic processes – invented by Shannon several years before – Kolmogorov proposed a measure for complexity which is not explicitly based on the notion of probability of an event. In this case it is possible to define a measure for complexity of a finite object (string) without the need to define the probability of this object to occur, which is needed for a rigorous definition of entropy rate of stochastic processes.

Suppose we are seeking for a description of some object by a finite string of symbols (e.g. binary string). In principle, there are numerous such descriptions possible. It is rather logical to consider the length of the shortest description of an object as a measure for its complexity. In his seminal article [Kol65], Kolmogorov defined the notion for complexity of a finite string with respect to a certain fixed universal computer[12]. Formally, the Kolmogorov (or algorithmic) complexity $K(S)$ of a string S is the length of the shortest binary program *prog* that causes the universal computer U to print (generate) S and then stop, i.e.,

$$K(S) = \min_{U(prog)=S} \{length(prog)\} \tag{2.23}$$

The main contribution of Kolmogorov was his argument that this definition of a string's complexity is independent on the choice of the computer. Various versions of this notion of complexity were independently discovered at approximately the same time by several other authors (see references in [LV08, CT91]). To demonstrate the concept of Kolmogorov complexity let us first consider a very long periodic string $S_{prog} = 000100010001000100010001\cdots$. It is obvious that this string can be generated with a short computer program $prog_s$: *use the substring* 0001 *n−times to print* S. In contrast to this a purely random string of the same length $S_r = 010101000100110101010\cdots$ can only be generated by a program $prog_r$: *use* S_r *to print* S_r which is at least as long as the length of S_r. This indicates that Kolmogorov complexity of the above mentioned periodic string has a constant value $K(S_{prog}) = 4 + c$ and is independent of the length of S_{prog}, whereas Kolmogorov complexity of the random string grows linearly with the length of the string $K(S_r) = length(S_r) + c$. Here, c is some constant that depends on details of implementation of the used computer programs. A more interesting example is the Kolmogorov complexity of a binary string that represents first n ($n \gg 1$) bits of the number π. Appearing apparently random this string can nevertheless be generated by a relatively short computer program (e.g. using a series expression of π). This sets Kolmogorov complexity of π to some constant value that is independent on the length of the given string [EFS98].

Due to the rather abstract nature of the above given theoretical definition of Kolmogorov complexity there is a need to define a more practical definition of this measure. Naively speaking we can always argue that the shortest program to generate some string S may be defined as *prˆog*: *use* $\hat{S}$ *to print* S, where $\hat{S}$ is a compressed version of the original string S.

[12] For instance, we can consider the *universal Turing machine* as a simplest form of universal computation. For further details see [CT91].

Thus, the problem of finding the Kolmogorov complexity of a string S can be reduced to the problem of finding an optimal lossless compression scheme[13] for this string.

2.4.1. Lempel-Ziv complexity

An important and popular class of techniques for string compression has been developed by Lempel and Ziv in two seminal papers in which they described two adaptive dictionary compression algorithms [LZ76, ZL77]. The use of dictionaries for data compression goes back to the invention of the telegraph. Indeed, it was more practical to produce a codebook for frequently used phrases. Thus, it was possible for the sender to compress a transmitted message and for the receiver to decompress it back without any loss of information when both parties have access to the codebook. However, the idea to use a self-adapting dictionary for string compression was not investigated until the works of Lempel and Ziv. In the first algorithm, which is now referred to as Ziv-Lempel (ZL) algorithm, the string $S \equiv \{s_i\}_{i=1}^{N}$ is sequentially parsed into distinct non-overlapping words $\{w_1, w_2, \ldots\}$ such that each shortest new word w_k does not belong to a set of previously seen words, i.e., $w_k \notin \{w_1, w_2, \ldots, w_{k-1}\}$. For instance, the binary string $\{s_i\}_{i=1}^{N} = 110101001111\cdots$ is parsed into words $(1)(10)(101)(0)(01)(11)(1\cdots$. Thus, each new word w_k $(k > 1)$ can be considered as a combination of some previously seen word w_j $(j < k)$ along with an additional symbol s_t, i.e., $w_k \equiv w_j s_t$. In this way we can encode (compress) any – even a very long – new word with just a pair of symbols, i.e., $w_k \equiv (j, s_t)$ where j is an index that refers to the word w_j. The resulting set of words defines our dictionary that we can now use to efficiently compress the original string $\{s_i\}_{i=1}^{N}$.
The second version of adaptive dictionary compression algorithms is now referred to as Lempel-Ziv (LZ) algorithm. In this version the string $S \equiv \{s_i\}_{i=1}^{N}$ is also sequentially parsed into distinct non-overlapping words $\{w_1, w_2, \ldots\}$. In contrast to the ZL algorithm, here, each new word w_k is not necessarily an extension of a previously seen word w_j $(j < k)$. It can be considered as a minimal (1 bit) extension of any substring which was seen before. For instance, the same string $\{s_i\}_{i=1}^{N} = 110101001111\cdots$ is parsed as $(1)(10)(10100)(111)(1\cdots$. Particularly, we set the first word $w_1 = \{s_i\}_{i=1}^{1} = 1$, the second word $w_2 = \{s_i\}_{i=2}^{3} = 10$ is a minimal extension of the substring $1 \in \{s_i\}_{i=1}^{2} = 11$, the third word $w_3 = \{s_i\}_{i=4}^{8} = 10100$ is a minimal extension of the substring $1010 \in \{s_i\}_{i=1}^{7} = 1101010$, and the $k^{th}-$ word $w_k = \{s_i\}_{i=j}^{t}$ is now a minimal extension of some substring $\xi \in \{s_i\}_{i=1}^{t-1}$ $(j \leq t)$. The number of distinct words defines the size of the dictionary for a given string S. The size of the dictionary constructed by the LZ algorithm is referred as *Lempel-Ziv complexity* $C_{LZ}(S)$, and for the ZL algorithm as *Ziv-Lempel complexity* $C_{ZL}(S)$.

Relation of Lempel-Ziv and Ziv-Lempel complexities to entropy rate

Let us now consider our string $S \equiv \{s_i\}_{i=1}^{N}$ as a realization of some stochastic process Z. As was already shown in section 2.2.1 the stochastic process is uniquely characterized by

[13] A lossless compression scheme is defined as a process of compression without any information loss. In this case, by using the compressed version of the string $\hat{S}$ it is always possible to reconstruct the original string S.

the entropy rate $dH(Z)$ that quantifies an average amount of entropy generated by the stochastic process per time step. According to Shannon [Sha48], the entropy rate $dH(Z)$ also quantifies to what extent the original message can be compressed without loss of information. As was shown later by Ziv, Lempel and others, both complexity measures $C_{LZ}(S)$ and $C_{ZL}(S)$ of a string S are asymptotically related to the entropy rate of an underlying stochastic process Z as

$$dH(Z) = \lim_{N\to\infty} \frac{\log(N)}{\langle C(S)\rangle} \tag{2.24}$$

where $\langle C(S)\rangle = \frac{N}{C_{LZ}(S)}$ or $\langle C(S)\rangle = \frac{N}{C_{ZL}(S)}$ denote the average word length estimated with either the LZ- or the ZL adaptive dictionary compression algorithms accordingly. It was shown that both estimators indeed converge (with $N \to \infty$) to the entropy rate, if our string is a realization of a stationary ergodic stochastic process [LZ76, ZL77, CT91, SG96]. However, the convergence in the case of the ZL-algorithm was demonstrated to be considerably slower than in the case of the LZ- algorithm [SG96]. According to these findings the LZ- algorithm provides a more superior approach to estimate the entropy rate for usually short experimental data and will be used throughout all further analyses in this thesis.

The approaches for lossless compression provide a practical means to estimate[14] the algorithmic complexity of finite strings. As originally stated by Kolmogorov [Kol68], in contrast to the notion of entropy rate of a stochastic process, that can only be rigorously defined for infinitely long strings, the notion of algorithmic complexity provides a measure for complexity of finite objects (strings).

2.4.2. Lempel-Ziv complexity for multivariate data analysis

Recent investigations have shown a potential use of the concept of LZ-complexity to characterize correlations in multivariate data [ZRB05]. Indeed, let us – for simplicity – consider the case of two binary strings[15] $S_1 \equiv \{s_i^1\}_{i=1}^N = 011010100\cdots$ and $S_2 \equiv \{s_i^2\}_{i=1}^N = 110010011\cdots$, which are the realizations of two stationary and ergodic stochastic processes Z_1 and Z_2. Both strings can now be jointly seen as a new string $S_{12} \equiv \{s_i^{12}\}_{i=1}^N = \left\{\binom{s_i^1}{s_i^2}\right\}_{i=1}^N = \binom{0}{1}\binom{1}{1}\binom{1}{0}\binom{0}{0}\binom{1}{1}\binom{0}{0}\binom{1}{0}\binom{0}{1}\binom{0}{1}\cdots = 231030122\cdots$, whose symbols are taken from the joined alphabet which was defined as $s_i^{12} = s_i^1 + 2s_i^2$. As it was shown recently all main results of the Lempel and Ziv approach remain valid for such a vectorial sequence too [ZRB05]. By using the same argumentation (i.e., by sequentially parsing the string $\{s_i^{12}\}_{i=1}^N$ into a set of distinct words which were not seen before, i.e.,

[14] The notion of Kolmogorov or algorithmic complexity is deeply related to some main results of theoretical computer science, namely: Gödel's incompleteness theorem and Turing's halting problem [LV08, CT91]. One of the main consequences of this interrelation is the theorem of the incomputability of Kolmogorov complexity. According to this theorem there is no computer program that takes a string S as an input and produces the integer $K(S)$ as an output. Thus, all approaches (e.g. LZ-complexity) to compute the algorithmic complexity of a string can only provide an estimate of this measure.

[15] The presented results remain valid for strings with an arbitrary size of the alphabet.

$\{s_i^{12}\}_{i=1}^N = (2)(3)(1)(0)(30)(12)(2\cdots)$ as in section 2.4.1 it is possible to define a *joined Lempel-Ziv* complexity $C_{LZ}(S_{12}) = C_{LZ}(S_1, S_2)$ of strings $\{s_i^1\}_{i=1}^N$ and $\{s_i^2\}_{i=1}^N$. Defined in such a way a bivariate measure for joined complexity of two strings is obviously symmetric, i.e., $C_{LZ}(S_1, S_2) = C_{LZ}(S_2, S_1)$. Moreover, as it was shown by [ZRB05] its value is related to the joined entropy rate (see section 2.15) of the underlying bivariate stochastic process $Z_b = (Z_1, Z_2)$ as

$$dH(Z_1, Z_2) = \lim_{N\to\infty} \frac{\log(N)}{\langle C_{LZ}(S_1, S_2)\rangle} \tag{2.25}$$

where $\langle C_{LZ}(S_1, S_2)\rangle = \frac{N}{C_{LZ}(S_1,S_2)}$ again denotes the average word length in the dictionary obtained by parsing the string $\{s_i^{12}\}_{i=1}^N$ with a LZ-compression algorithm (see section 2.4.1). Using the same argumentation it is possible to extent the definition of joint LZ-complexity to the case of more than two stochastic processes [ZRB05].
Let us again consider two strings $S_1 \equiv \{s_i^1\}_{i=1}^N$ and $S_2 \equiv \{s_i^2\}_{i=1}^N$ which are realizations of two stationary ergodic processes Z_1 and Z_2. By following the analogy with the definition of mutual information and mutual information rate between Z_1 and Z_2 (see section A.3) it is now rather logical to define the *mutual LZ-complexity* between two strings as

$$dI_{LZ}(S_1, S_2) = C_{LZ}(S_1) + C_{LZ}(S_2) - C_{LZ}(S_1, S_2) \tag{2.26}$$

and by using the asymptotic property of LZ-complexities (see Eq. 2.24 and Eq. 2.25) to define the normalized *mutual LZ-complexity* as

$$dI_{LZ}(S_1, S_2; N) = \frac{\log(N)}{N}(C_{LZ}(S_1) + C_{LZ}(S_2) - C_{LZ}(S_1, S_2)). \tag{2.27}$$

According to equations 2.24 and 2.25 for the case of infinitely long strings (i.e., $N \to \infty$) the normalized mutual LZ-complexity is related to the mutual information rate between two stochastic processes Z_1 and Z_2, i.e.,

$$\lim_{N\to\infty} dI_{LZ}(S_1, S_2; N) = dH(Z_1) + dH(Z_2) - dH(Z_1, Z_2) \tag{2.28}$$

and, according to equation 2.16 (see also Fig. A.2), can be rewritten as

$$\lim_{N\to\infty} dI_{LZ}(S_1, S_2; N) = TE(Z_1, Z_2) + TE(Z_2, Z_1) + dI(Z_1, Z_2). \tag{2.29}$$

$TE(Z_1, Z_2)$ and $TE(Z_2, Z_1)$ are transfer entropies (asymmetric part of interaction) and $dI(Z_1, Z_2)$ is the mutual information rate (symmetric part of interaction) between stochastic processes Z_1 and Z_2. Thus, *mutual LZ-complexity* characterizes the degree of correlation between two strings S_1 and S_2 and, in the limit of infinitely long strings, converges to the sum of transfer entropies and mutual information rate between underlying stochastic processes. As it was pointed out in [ZRB05], while the mutual information can be understood as a distance (Kullback-Leibler divergence) between two probability densities of the underlying sequences of data, the mutual LZ-complexity directly characterizes a divergence between these two sequences. Indeed, as it already was pointed out by Kolmogorov [Kol68], the

notion of mutual algorithmic complexity (in our case mutual LZ-complexity as a measure for mutual algorithmic complexity) allows one to define a measure of common information between two finite objects, e.g. finite strings. In contrast to this, the notions of mutual information (see Eq. A.21) and mutual information rate (see Eq. A.33) are defined as measures of common information between two stochastic variables and stochastic processes respectively [Sha48]. Practical applications of mutual LZ-complexity given in [ZRB05] revealed that its value can, in general, be negative, in contrast to the positively defined two-point mutual information function (see section A.3). This finding demonstrated that a direct application of mutual LZ-complexity may be rather limited and special care has to be taken for a reliable interpretation of obtained results.

3. Estimating Kolmogorov-Sinai entropy of chaotic dynamical system

Kolmogorov-Sinai or KS-entropy is one of the important characteristic invariants of deterministic dynamical systems. By definition (see Eq. 2.20) it characterizes the amount of entropy, measured in bits, generated by a dynamical system per time unit. As we could already see in section 2.3, KS-entropy can be estimated from a Shannon entropy rate of a symbol series obtained via partitioning of state space of the dynamical system (see Eq. 2.20). A complementary approach for the partitioning of a state space via permutation of state variables was introduced in [BP02, BKP02] where a real-valued observable of some dynamical system was transformed into a series of permutation symbols (for the definition of permutation symbols see section 2.3.1). In the same work the authors also introduced the notion of the permutation entropy rate which resembles the Shannon entropy (Eq. 2.6) of permutation symbols. Recent findings [AKK05, AK07] showed that the permutation entropy rate is indeed related to KS-entropy of a dynamical system. However, the question as to what extent the Shannon entropy rate (Eq. 2.14) of permutation symbols is numerically related to KS-entropy of the underlying dynamical system remains unclear. In this chapter we will numerically study the symbolic representation of a simple time-discrete deterministic dynamical system – a one-dimensional chaotic map. Such a system is usually defined by a recursive relation between the state variable at time step n and the state variable at the next time step $x_{n+1} = f(x_n)$ but can nevertheless exhibit a rather complex dynamical behavior. For example, the tent map is defined by a rather simple piece-wise function f (Eq. 3.1). However, this function is not bijective and the reverse mapping $x_n = f^{-1}(x_{n+1})$ is not uniquely defined. Thus, during each iteration of the tent map some amount of information about the current state of the system is lost. This results in a chaotic dynamical behavior and the sensitive dependence on initial conditions.

In the first part of the chapter we address the question as to what extent the entropy rate of permutation symbols obtained from real-valued time series of the tent map is related to its KS-entropy. First, we consider the tent map and briefly review its main property – the ability to generate real-valued time series with a varying and analytically defined value of KS-entropy. Then, we transform real-valued time series of the tent map into a series of permutation symbols with different values of the embedding parameters: m and τ (for definition see section 2.3.1). Finally, we numerically estimate the entropy rate of permutation symbols with the Shannon entropy rate as well as the permutation entropy rate proposed in [BP02, BKP02] for different embedding parameters and compare obtained results with the analytically given values of KS-entropy of the tent map. In the second part of this chapter we consider a binary symbolic representation of the tent and Hénon maps obtained with the threshold-crossing partitions Π_{bin} (defined below). Next, we estimate

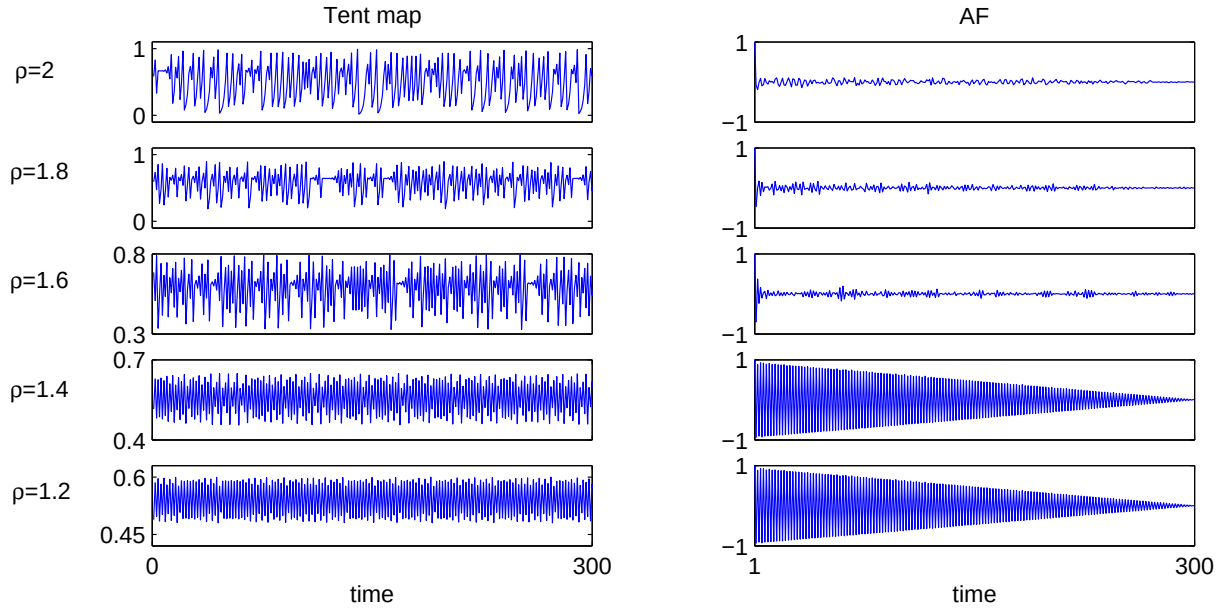


Figure 3.1.: Exemplary segments of tent map time series and corresponding autocorrelation functions (AF) for different values of the control parameter ρ (Eq. 3.1).

KS-entropy of these maps by computing the Shannon entropy rates of different orders as well as the Lempel-Ziv estimator of entropy rate of corresponding binary symbol series. By generating tent map time series with different values of KS-entropy as well as time series obtained with the Hénon map we investigate main similarities and differences between two these approaches for the estimation of KS-entropy of chaotic maps.

3.1. Symbolic representation of tent map

The tent map is widely used (see [Str01]) as a one-dimensional iterative map exhibiting chaotic behavior for $\rho \in [1, 2]$

$$x_{n+1} = \rho(1 - 2|x_n - 0.5|). \tag{3.1}$$

By iterating a randomly chosen initial condition with Eq. 3.1 it is possible to generate time series of different complexities and different degrees of temporal correlations (i.e., long-term memory effects; see Fig. 3.1). To exhibit chaotic dynamical behavior the tent map should show the sensitive dependence on initial conditions, in the sense that neighboring orbits in state space separate exponentially fast. This results in the positivity of the Lyapunov exponent (section 2.1.2).
In fact, the dynamics of the tent map can be studied analytically even in the chaotic regime. It can be shown that the Lyapunov exponent of the tent map is $\lambda_{tent} = \log(\rho)$ [Str01]. Since the value of the Lyapunov exponent is known it is now possible, by using *Pesin's identity* (see Eq. 2.21), to estimate the upper bound of KS-entropy of the tent map as $h_{KS}^{tent} \leq \log(\rho)$ for $\rho \in [1, 2]$. Moreover, it is also known that the equality is achieved for the topological

entropy of the tent map [dMvS93]. Thus, by varying the value of the control parameter ρ we can generate time series with different values of KS-entropy $h_{KS}^{tent} \in [0, 1]$. Let us consider an exemplary time series generated by a tent map with $\rho = 2$. In this case KS-entropy of this time series is analytically given as $h_{KS}^{tent} = 1$ bit per iteration. According to the definition of KS-entropy (see section 2.2.3) we are going to loose 1 bit of information about the initial state per single forward iteration. Thus, if we now specify the value of the initial state with, let us say, 32 bit precision then after 32 iterations all information about the initial state will be lost.

Several attempts have been made to estimate KS-entropy of one-dimensional chaotic maps via a symbolic representation. In [EST01] the authors used threshold-crossing partitions to derive a symbolic representation of real-valued time series. Complementary to this, other authors used permutation symbols to estimate KS-entropy of one-dimensional chaotic maps [BKP02]. Further theoretical analyses have shown that permutation symbols can be used to define the permutation entropy rate which approaches the KS-entropy of an arbitrary ergodic dynamical system [AKK05, AK07].

3.1.1. Symbolic representation with permutation partition

Following [BKP02] we start our analysis of KS-entropy of the tent map by transforming the real-valued time series generated with this map into a series of permutation symbols (for the definition of permutation symbols see section 2.3.1). Let x_n $(n = 1, \ldots, N)$ denote a time series of length $N = 10^5$ generated by a tent map and π_n $(n = 1, \ldots, \tilde{N}$; where $\tilde{N} = N - (m-1)\tau)$ denote a series of permutation symbols which form a string S. For each value of the control parameter ρ (Eq. 3.1), by using random initial conditions, we generated 20 realizations of tent map time series. Following [BP02, BKP02] let us define the permutation entropy rate of the tent map as

$$h_m = \frac{1}{m-1} H(\pi_i), \tag{3.2}$$

where $i \in [1, \tilde{N}]$ and $H(\pi_i)$ denotes the Shannon entropy (Eq. A.18) of a series of permutation symbols S. As it was shown in [BKP02] the permutation entropy rate h_m of the tent map indeed converges to KS-entropy for $m \to \infty$. However, it was also demonstrated that this estimator of KS-entropy exhibited a rather slow converging behavior (especially for $\rho < 1.5$, see Fig. 3.3 and [BKP02]). In contrast to this approach, in our work we will estimate KS-entropy of the tent map by computing the entropy rate of a series of permutation symbols S. To do this, we used (see section 2.2.1 and [Sha48]) the Shannon entropy rate estimator of order k which was defined as

$$dH^k(S) = H(\pi_i, \pi_{i-1}, \ldots, \pi_{i-k}) - H(\pi_{i-1}, \ldots, \pi_{i-k}), \tag{3.3}$$

where $i \in [k{+}1, \tilde{N}]$. In order to compute the block entropies $H(\pi_i, \ldots, \pi_{i-k})$ and $H(\pi_{i-1}, \ldots, \pi_{i-k})$ we estimated empirical joint probabilities $\hat{p}(\pi_i, \ldots, \pi_{i-k})$ and $\hat{p}(\pi_{i-1}, \ldots, \pi_{i-k})$. To do this, we counted the corresponding relative frequencies, i.e.,

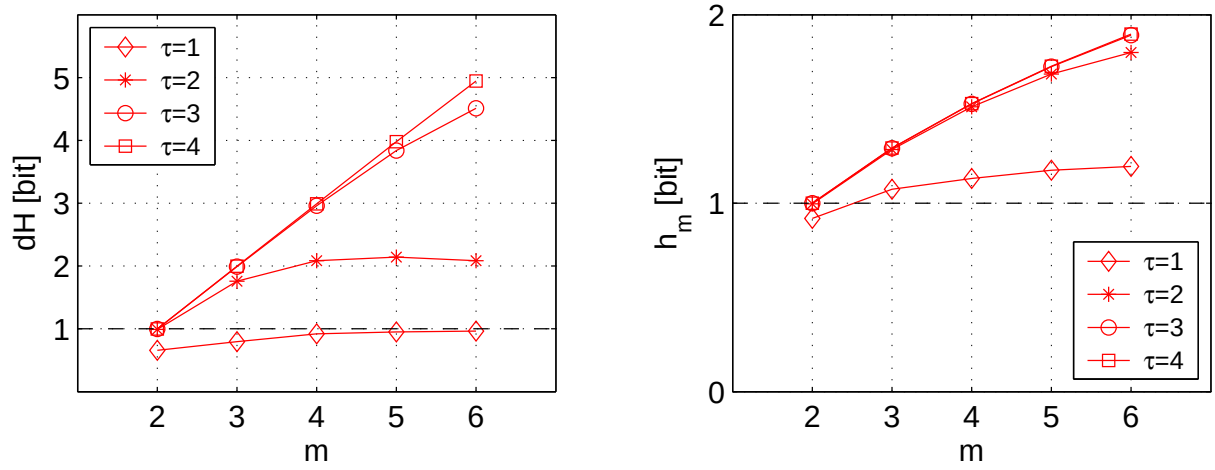


Figure 3.2.: Entropy rate of permutation symbols obtained from a real-valued time series generated with the tent map for $\rho = 2$ (Eq. 3.1). In order to estimate entropy rate we used the first order Shannon estimator dH (left) and the permutation entropy rate h_m (right). Tent map time series were transformed into series of permutation symbols by using different embedding parameters m and τ. The dashed line represents the analytical value of KS-entropy of the tent map. Each point on both plots corresponds to the mean value taken over 20 realizations. The standard deviations were too small and are not shown.

$$\hat{p}(\pi_i, \pi_{i-1}, \ldots, \pi_{i-k}) = \frac{W(\pi_i, \pi_{i-1}, \ldots, \pi_{i-k})}{\tilde{N} - k} \tag{3.4}$$

and

$$\hat{p}(\pi_{i-1}, \ldots, \pi_{i-k}) = \frac{W(\pi_{i-1}, \ldots, \pi_{i-k})}{\tilde{N} - k},$$

where $W(\pi_i, \pi_{i-1}, \ldots, \pi_{i-k})$ and $W(\pi_{i-1}, \ldots, \pi_{i-k})$ denote the number of occurrences of blocks of symbols $(\pi_i, \pi_{i-1}, \ldots, \pi_{i-k})$ and $(\pi_{i-1}, \ldots, \pi_{i-k})$ in the string S. By assuming that the symbol series S represents a realization of a stationary and ergodic stochastic process we guaranteed that these empirical joint probabilities will asymptotically converge to the true probability distributions, i.e., $\lim_{\tilde{N}\to\infty} \hat{p} \to p$. In the following we will skip the index for the first order Shannon entropy rate, i.e., $dH^k = dH$ for $k = 1$.
Let us first consider time series generated by a tent map with the control parameter $\rho = 2$ (Eq. 3.1). In this case, KS-entropy of the tent map time series is analytically given ($h_{KS}^{tent} = \log(2) = 1$ bit per iteration). Analysis of the dependency of the Shannon entropy rate $dH(S)$ on embedding dimension m revealed a different converging behavior of this estimator to KS-entropy of the tent map depending on the used time delay τ (see Fig. 3.2). Our findings indicated that for $\tau = 1$ the Shannon entropy rate showed a converging behavior

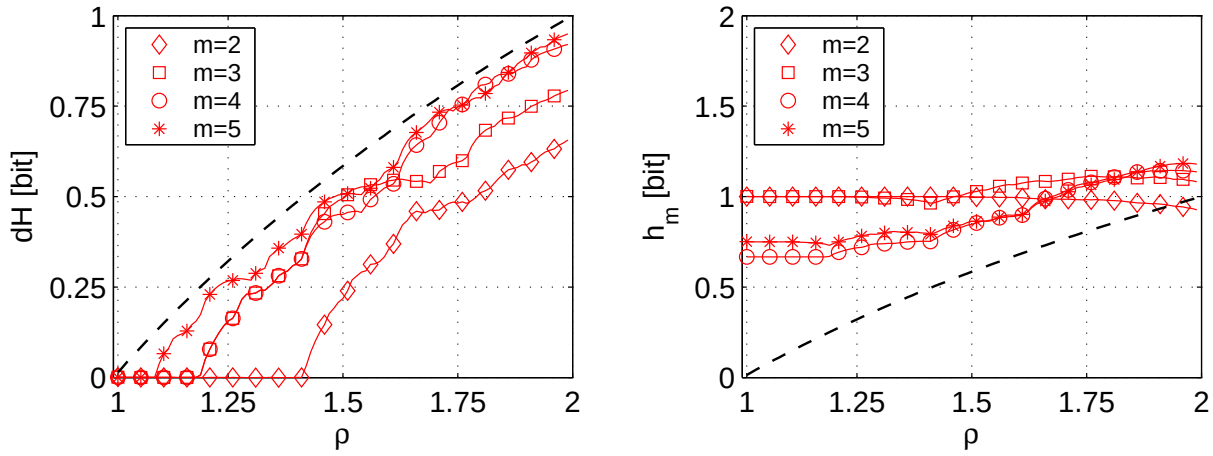


Figure 3.3.: Entropy rate of permutation symbols obtained from real-valued time series of the tent map for increasing value of control parameter $\rho \in (1, 2]$ (Eq. 3.1). In order to estimate entropy rate we used the first order Shannon estimator dH (left) and the permutation entropy rate h_m (right). Tent map time series were transformed into series of permutation symbols by using different embedding dimensions m and time delay $\tau = 1$. The dashed line represents the analytical value of KS-entropy of the tent map. Each point on both plots corresponds to the mean value taken over 20 realizations. The standard deviations were too small and are not shown.

(from below) to KS-entropy of the tent map for the increasing embedding dimension m, i.e., $dH(S) \to h_{KS}^{tent}$ for increasing values of m (Fig. 3.2). Further analysis for $\tau > 1$ indicated that obtained values of the Shannon entropy rate $dH(S)$ exceeded KS-entropy of the tent map and for $\tau = \tau_0 = 4$ (where τ_0 is the first minimum of the time-delayed mutual information function) the Shannon entropy rate was linearly growing with the embedding dimension m as $dH(S) = (m-1)h_{KS}^{tent}$. Thus, as one can see in Fig. 3.2 the Shannon entropy rate $dH(S)$ is indeed related to KS-entropy of the tent map but its value has to be normalized accordingly,

$$dH^k(S) = L_{norm}(H(\pi_i, \pi_{i-1}, \ldots, \pi_{i-k}) - H(\pi_{i-1}, \ldots, \pi_{i-k})), \tag{3.5}$$

with

$$L_{norm} = \begin{cases} 1, \text{ for } \tau = 1 \\ 1/(m-1), \text{ for } \tau \geq \tau_0 \end{cases} \tag{3.6}$$

where the time delay τ_0 can be estimated as the first minimum of the time-delayed mutual information[1]. To explain the observed dependency of the Shannon entropy rate $dH(S)$ on the embedding parameters m and τ we need to recall that, by construction, permutation

[1]Similar considerations are also used by the time delay embedding procedure in order to find an optimal embedding vector for the reconstruction of state space of dynamical systems (further details are given in section 2.1.3 and in [KS03]).

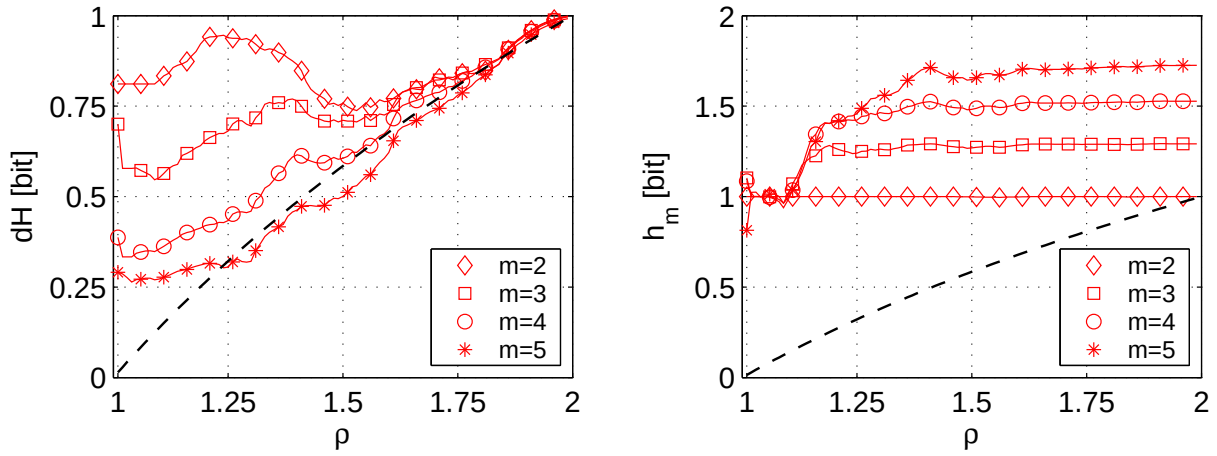


Figure 3.4.: Entropy rate of permutation symbols obtained from real-valued time series of the tent map for increasing value of control parameter $\rho \in (1, 2]$ (Eq. 3.1). In order to estimate entropy rate we used the first order Shannon estimator dH (left) and the permutation entropy rate h_m (right). Tent map time series were transformed into series of permutation symbols by using different embedding dimensions m and time delay $\tau = 4$. The dashed line represents the analytical value of KS-entropy of the tent map. Each point on both plots corresponds to the mean value taken over 20 realizations. The standard deviations were too small and are not shown.

symbols can be correlated with each other and thus contain redundant information. Let us first set $\tau = 1$ and consider a permutation symbol π_n along with the corresponding embedding vector $\vec{X}(n) = (x_n, x_{n-1}, \ldots, x_{n-(m-1)})$ (see section 2.3.1). It is easy to see that the embedding vector $\vec{X}(n)$ and the next embedding vector $\vec{X}(n+1) = (x_{n+1}, x_n, \ldots, x_{n-(m-2)})$ overlap with each other. This implies that, independent of the embedding dimension m, only the first component of the embedding vector $\vec{X}(n+1)$ will contain new information generated by the tent map during one iteration. Next, let us set $\tau > 1$ and again consider two successive embedding vectors $\vec{X}(n)$ and $\vec{X}(n+1)$. In this case, these vectors will not overlap with each other and therefore all components of the embedding vector $\vec{X}(n+1)$ will contain some amount of new information. In this case, the total amount of new information contained in the the embedding vector $\vec{X}(n+1)$ (relative to $\vec{X}(n)$) and therefore in the permutation symbol π_{n+1} (relative to π_n) should be proportional to the embedding dimension m. Indeed, as one can see in Fig. 3.2 for $\tau = \tau_0$ (for this τ the components of the embedding vector were independent from each other), the amount of new information per a permutation symbol was linearly growing with the embedding dimension m, i.e., $dH(S) = (m-1)h_{KS}^{tent}$. This allowed us to normalize Shannon entropy rate $dH(S)$ (see Eqs. 3.3 and 3.5) and thus to make it independent on the embedding dimension m. It is, however, important to point out that for $1 < \tau < \tau_0$ the components of the embedding vector will not be independent from each other anymore. In this case, as one can see in Fig. 3.2 for $\tau = 2$ and $\tau = 3$, the amount of new information per permutation symbol was non-

linearly depending on the embedding dimension m and the Shannon entropy rate $dH(S)$ could not be normalized such as to make it independent on the embedding dimension m. Analysis of the permutation entropy rate $h_m(S)$ for $m = 2$ showed that its values converged to KS-entropy of the tent map for increasing values of τ. For other embedding parameters the obtained values of the permutation entropy rate exceeded KS-entropy of the tent map and, in general, exhibited non-linear dependence on the embedding dimension m.
Let us now extent the analysis made above and consider a set of time series generated by a tent map with increasing value of the control parameter $\rho = 1, \ldots, 2$ with a step $\delta\rho = 0.05$ (see Eq. 3.1). This allowed us to generate time series with analytically given KS-entropy ($h_{KS}^{tent} = \log(\rho) = 0, \ldots, 1$) and different degrees of temporal correlations (see Fig. 3.1). The real-valued time series of the tent map were now transformed into a series of permutation symbols S by using increasing embedding dimensions $m = 2, \ldots, 5$ and two delay times $\tau = 1$ and $\tau = \tau_0 = 4$. The Shannon entropy rate $dH(S)$ (mean value taken over 20 realization) was now estimated according to Eq. 3.5. For $\tau = 1$ our analysis revealed that the Shannon entropy rate approached KS-entropy of the tent map when we increased the embedding dimension m (Fig. 3.3). However, for $\tau = \tau_0$, we observed a different converging behavior of $dH(S)$ depending on the the control parameter ρ. For $\rho > 1.5$ we found that $dH^1(S)$ was almost independent on the embedding dimension m, whereas for $\rho < 1.5$ the Shannon entropy rate was approaching to KS-entropy for increasing embedding dimension m (Fig. 3.4). To explain this finding we need to recall that the time series generated with a tent map with either $\rho > 1.5$ or $\rho < 1.5$ exhibited a different degree of temporal correlations. The analysis of temporal correlations of the tent map time series revealed that for $\rho > 1.5$ such time series exhibited a fast decaying autocorrelation function, whereas for $\rho < 1.5$ the autocorrelation function decayed very slowly indicating long-term temporal correlations (see Fig. 3.1). Estimation of KS-entropy of the tent map using the permutation entropy rate $h_m(S)$ for $\tau = 1$ (see Fig. 3.3) revealed a converging behavior (from above) for an increasing embedding dimension m. Similarly to findings presented in [BKP02] we observed a rather slow converging behavior especially for $\rho < 1.5$. Numerical analysis of KS-entropy of the tent map by using the permutation entropy rate h_m made in [BKP02] showed that the accurate estimation of KS-entropy required embedding dimensions $m > 10$. Next, we repeated the analysis of KS-entropy of the tent map but now we set the delay time to $\tau = \tau_0 = 4$. In this case, we observed that the values of the permutation entropy rate $h_m(S)$ exceeded KS-entropy for all embedding dimensions m and control parameters ρ. Moreover, oppositely to the case $\tau = 1$ considered above, the permutation entropy rate $h_m(S)$ was growing for increasing embedding dimension m and thus did not exhibit a converging behavior to KS-entropy.

3.1.2. Symbolic representation with threshold-crossing partition

In this part of the chapter we consider a binary symbolic representation of chaotic maps. In this representation the real-valued time series generated with these maps are transformed into series of binary symbols. In order to derive such a binary symbolic representation we applied a threshold-crossing partitioning used e.g. in [EST01]. Being rather simple and therefore useful in many practical applications this partitioning scheme can, in some cases,

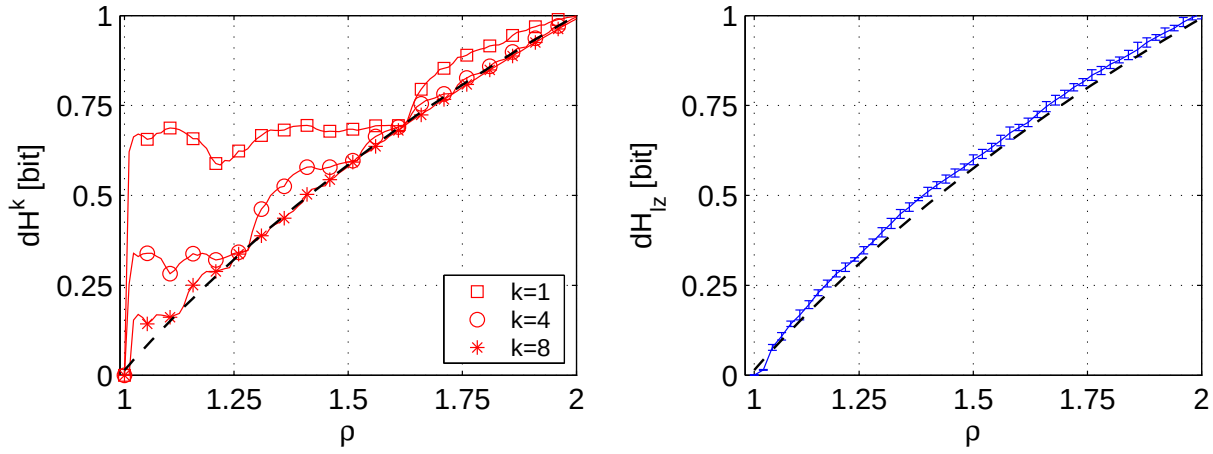


Figure 3.5.: Entropy rate of binary symbols obtained from real-valued time series of the tent map for increasing value of control parameter $\rho \in (1, 2]$ (Eq. 3.1). In order to estimate entropy rate we used the Shannon estimator dH^k of increasing orders k (left) and the LZ-based estimator dH_{LZ} (right). The tent map time series were transformed into series of binary symbols by using binary partition Π_{bin} (see text). The dashed line represents the analytical value of KS-entropy of the tent map. Each point on both plots corresponds to the mean value taken over 20 realizations. Left: Error bars were too small and are not shown. Right: Error bars denote standard deviations.

lead to the incorrect characterization of the underlying dynamical behavior [BSLZ01]. For instance, in the following analysis it will be demonstrated that the incorrect selection of a threshold can, in general, lead to an overestimation of KS-entropy of chaotic maps. For the estimation of KS-entropy of chaotic maps we will extent the analysis of the previous section and will compute it by using high order Shannon as well as Lempel-Ziv (see section 2.4.1) estimates of the entropy rate. Let us recall that the Lempel-Ziv approach is based on the notion of the Lempel-Ziv complexity C_{LZ}. Following [LZ76, ZL77] we defined the LZ-complexity of the string S as a size of the dictionary obtained with the Lempel-Ziv parsing algorithm. The dictionary was computed by sequential parsing of the string into a set of distinct non-overlapping words $\{w_1, w_2, \ldots\}$ such that each shortest new word w_k does not belong to the set of previously seen words, i.e., $w_k \notin \{w_1, w_2, \ldots, w_{k-1}\}$ (for details see section 2.4.1). Following [LZ76, ZL77, CT91, SG96] we defined an approximation of the entropy rate for the string S as

$$dH_{LZ}(S) = C_{LZ}(S)\frac{\log(\tilde{N})}{\tilde{N}}. \tag{3.7}$$

In contrast to the Shannon entropy rate $dH^k(S)$ (Eq. 3.3), which takes into account temporal correlations only over the k time steps, the Lempel-Ziv approach allows to capture long-term correlations within the string S.

Let us first consider a set of time series generated by a tent map with increasing value of the control parameter $\rho = 1, \ldots, 2$ with a step $\delta\rho = 0.05$ (Eq. 3.1). In contrast to the

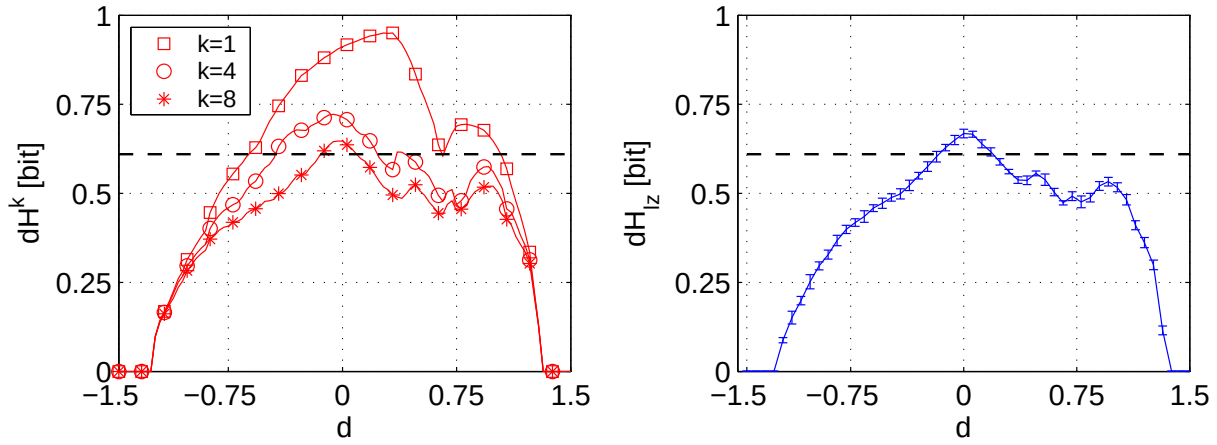


Figure 3.6.: Entropy rate of binary symbols obtained from real-valued time series of the Hénon map for increasing value of the threshold d. In order to estimate entropy rate we used the Shannon estimator dH^k of increasing orders k (left) and the LZ-based estimator dH_{LZ} (right). The Hénon map time series were transformed into series of binary symbols by using threshold crossing partition with the threshold d (see text). The dashed line represents the analytical value of KS-entropy of the Hénon map. Each point on both plots corresponds to the mean value taken over 20 realizations. Left: Error bars were too small and are not shown. Right: Error bars denote standard deviations.

previous section, we derived the coarse-grained representation B_n $(n = 1, \ldots, N = 10^5)$ of tent-map time series by using the binary partition Π_{bin} which was defined as $\Pi_{bin} = [\Pi_1; \Pi_2]$ with $\Pi_1 \equiv [0, 0.5]$ and $\Pi_2 \equiv (0.5, 1]$. For each value of the control parameter ρ we generated 20 realizations of tent map time series. By assuming that under the action of the dynamics the trajectory of the tent map (which is bounded to the interval $[0, 1]$; see Fig. 3.1) visits both elements of this partition, we assigned to each of the two elements of Π_{bin} a symbol B_n and thus encoded the time evolution of the tent map into a sequence of binary symbols S_{bin}. Next, we computed the Shannon entropy rates $dH^k(S_{bin})$ of different orders k (Eq. 3.3) as well as the LZ-based entropy rate $dH_{LZ}(S_{bin})$ (see Eq. 3.7 and Fig. 3.5). Our findings indicate that for $\rho = 2$ the obtained estimates of $dH^k(S_{bin})$ were almost independent on the order k. This implies that for $\rho = 2$ the used binary partition Π_{bin} represents a generating (or even Markov) partition of the tent map (see section 2.3 or [Str01]). However, by analyzing the tent map time series for $\rho < 2$ we found that the obtained estimates of $dH^k(S_{bin})$ were now dependent on k approaching KS-entropy of the tent map from above as the order k was increased. This finding indicates that the used binary partition Π_{bin} might not necessarily be a Markov partition for a tent map with $\rho < 2$ and binary symbols produced with this partition might exhibit long-term temporal correlations (memory effects). For instance, as one can see in Fig. 3.5, the estimation of KS-entropy of the tent map for $\rho \approx 1.2$ ($h_{KS}^{tent} = log(\rho) \approx 0.25$ bit per iteration) required at least the $8^{th}-$ order estimator of the Shannon entropy rate $dH^{(8)}(S_{bin})$. The further analysis of the tent map time series with $dH_{LZ}(S_{bin})$ indicated that the LZ-based approach

to measure entropy rate of a symbolic sequence, at least in this case, provided a more accurate approximation of KS-entropy than the Shannon estimator for all investigated values of the control parameter ρ (see Fig. 3.5).

As a next step, let us consider a two-dimensional Hénon map defined by Eq. A.2. For the canonical set of parameters (see section A.1) this map is chaotic and has one positive Lyapunov exponent $\lambda_1 = 0.61$ (see table 2.1). This allowed us, by using Pesin's identity (Eq. 2.21) to define the upper bound for KS-entropy of the Hénon map, $h_{KS} \leq 0.61$. Similar to the analysis made above, we derived the coarse-grained representation B_n ($n = 1, \ldots, N = 10^5$) by using the $x-$ component of the Hénon map (20 realizations) and then applied a family of binary partitions $\Pi_{bin}(d)$ which were parametrized with a threshold d and defined as $\Pi_{bin}(d) = [\Pi_1; \Pi_2]$ with $\Pi_1 \equiv (-\infty, d]$ and $\Pi_2 \equiv (d, \infty)$. By assuming that under the action of the dynamics the trajectory of the Hénon map visits both elements of this partition, we assigned to each of the two elements of Π_{bin} a symbol B_n and thus encoded the time evolution of the Hénon map into a sequence of binary symbols S_{bin}. Next, we computed the Shannon entropy rates $dH^k(S_{bin})$ of different orders k (Eq. 3.3) as well as the LZ-based entropy rate $dH_{LZ}(S_{bin})$ (Eq. 3.7) for increasing values of the threshold $d \in [-1.5, 1.5]$. By using this range of the threshold d we approximately covered the range of possible amplitudes of the $x-$ component of the Hénon map. Analysis of the first order Shannon entropy rate $dH(S_{bin})$ revealed its dependency on the threshold d (see Fig. 3.6). Following its definition (see section 2.2.3) we approximated KS-entropy of the Hénon map by taking a supremum over the all investigated binary partitions $\Pi_{bin}(d)$, i.e., $\sup_d \{dH(S_{bin})\} \approx 0.95$. Further analysis of KS-entropy by using higher orders ($k > 1$) Shannon entropy rates indicated a converging behavior to KS-entropy. It was found that $\sup_d \left\{dH^k(S_{bin})\right\} \approx 0.73$ for $k = 4$ and $\sup_d \left\{dH^k(S_{bin})\right\} \approx 0.63$ for $k = 8$ (see Fig. 3.6). Next, we repeated the analysis of KS-entropy of the Hénon map by using the Lempel-Ziv estimator of the entropy rate $dH_{LZ}(S_{bin})$. Our findings showed that this approach also provided an accurate approximation of KS-entropy, i.e., $\sup_d \{dH_{LZ}(S_{bin})\} \approx 0.64$, and its value was numerically very close to the estimate obtained with the 8^{th} order Shannon entropy rate. This finding agrees with the fact that the LZ-based approach is, by definition, optimized for the analysis of temporally correlated time series whereas the first order Shannon estimator better suites for the analysis of time series exhibiting short-term temporal correlations [SG96].

In this chapter we investigated two different symbolic representations of chaotic maps for which we estimated KS-entropy. In the first part of our analysis we addressed the question as to what extent the entropy rate of permutation symbols obtained from real-valued time series of the tent map is related to its KS-entropy. To estimate the entropy rate of permutation symbols we used the Shannon entropy rate dH as well as the permutation entropy rate h_m proposed in [BP02, BKP02]. It was shown that dH was approaching KS-entropy of the tent map for increasing embedding dimensions m. However, the converging behavior depended on the delay time τ. For $\tau = 1$ the Shannon entropy rate converged to KS-entropy of the tent map from below (Fig. 3.3). The analysis of entropy rate of permutation symbols obtained with $\tau > 1$ indicated that Shannon estimator was also approaching KS-entropy of the tent map (from above) when its values were normalized according to Eq. 3.5. Thus,

the obtained findings demonstrated that the entropy rate of the permutation symbols was indeed related to KS-entropy of the tent map and can be used to estimate the amount of entropy generated by the tent map per iteration. In contrast to the permutation entropy rate h_m which converges very slow to KS-entropy of the tent map (for $m > 10$ see [BKP02]), the Shannon estimator of the entropy rate dH allowed us to achieve a good approximation of KS-entropy of the tent map already for $m = 5$ ($\tau = 1$) for all values of control parameter ρ and already for $m = 2$ ($\tau = \tau_0$) for $\rho > 1.5$ (see Fig. 3.3). To answer the question whether these findings can be generalized to an arbitrary ergodic dynamical system requires additional investigations (e.g. time-continuous dynamical systems such as chaotic oscillators). In the second part of our analysis we considered a binary symbolic representation of the tent and the Hénon map. Analysis of tent map time series for increasing values of the control parameter ρ indicated that for $\rho < 1.5$ the Shannon entropy rates of high orders ($k > 4$) were required in order to obtain an accurate approximation of KS-entropy of the tent map (see Fig. 3.5). Analysis of entropy rate of a binary symbolic representation of the tent map using the Lempel-Ziv estimator showed that, in this special case, the LZ-based estimator provided a more accurate approximation of KS-entropy. Similar findings were also obtained when we estimated KS-entropy of the Hénon map (see Fig. 3.6).
Summarizing this chapter, we can conclude that the symbolic representation of the real-valued dynamics can be used to estimate an amount of entropy generated by a dynamical system per time unit. Extending these findings, in the next chapter we will characterize interactions between dynamical model systems by estimating the amount of mutual information and entropy transfer between them.

4. Characterization of interactions in dynamical systems

In this chapter we will apply information-theoretic measures (such as mutual information and transfer entropy; see sections 2.2.1 and A.3) to characterize strength and direction of interactions between dynamical model systems with a priori known properties and coupling schemes. In the following, we analyze interactions between time series from structurally identical as well as structurally non-identical chaotic maps and chaotic oscillators. In order to characterize the strength of interaction between these dynamical systems, we measure the amount of common (mutual) information shared between them by computing the so called symbolic mutual information. As a first step of our analysis the real-valued time series of used dynamical model systems are transformed into a symbolic representation. To do this we follow [SL08, SL09] and transform the time series into a series of permutation symbols. The symbolic mutual information is then estimated by measuring the zero-lagged mutual information (Eq. A.21) between symbol series. The defined in this way symbolic mutual information characterizes a degree of zero-lag (static) correlations between corresponding series of permutation symbols. To answer the questions as to how good the symbolic mutual information captures interdependencies between two time series and to what extent the estimated amount of mutual information between system observables is related to the strength of interaction between nonlinear dynamical systems we additionally estimate the mean phase coherence[1] R and qualitatively compare both measures with each other.
In order to reveal the directionality of interactions between dynamical systems we use the recently proposed symbolic transfer entropy [SL08] which allows us to estimate the amount of entropy transfer between time series. It is demonstrated that between two unidirectionally coupled dynamical systems, i.e., in the situation where, let us say, the first system (driver) is coupled into the second system (responder), there is nevertheless a certain amount of entropy transfer in the direction "responder-to-driver" despite the absence of interactions in this direction. Following and extending the concept of effective transfer entropy proposed in [MK02] we investigate a correction scheme for the symbolic transfer entropy which takes into account zero-lag (static) correlations between permutation symbols and introduce the notion of corrected symbolic transfer entropy. It is then shown that, in some cases, the observed entropy transfer in the direction "responder-to-driver" can indeed be corrected. Next, we investigate the entropy transfer in multivariate data obtained by simulating a cluster of interacting chaotic oscillators (similarly to the multivariate data

[1] The mean phase coherence R allows to characterize nonlinear interdependencies between two signals by measuring the degree of phase synchronization between them (for details see section A.2) and it is frequently used for the analysis of the strength of interactions between nonlinear dynamical systems and in field data [MLDE00, PRK01].

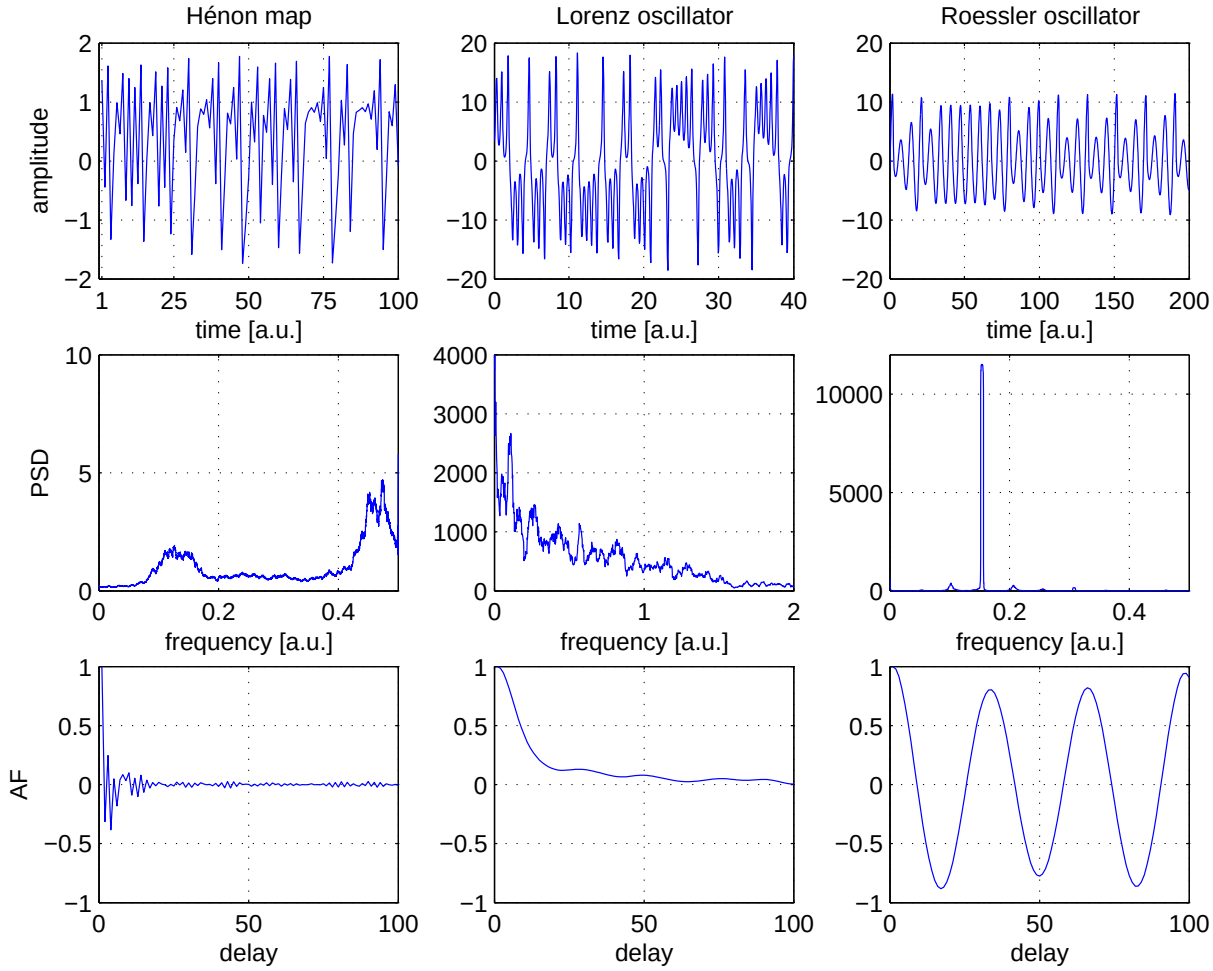


Figure 4.1.: Exemplary time series of the Hénon map, Lorenz and Rössler oscillators along with the corresponding power spectral density functions (PSD) and autocorrelation functions (AF).

investigated in [OMWL08] and [SL09]). It is demonstrated that the corrected symbolic transfer entropy allows to simplify the identification of driving structures in multivariate data.

4.1. Characterizing strength of interactions with symbolic mutual information

We begin our analysis of interactions between dynamical model systems by characterizing the strength of interactions between two diffusively coupled Lorenz oscillators in dependence on the coupling strength. To do this we estimate, for each value of the coupling strength, the symbolic mutual information MI (Eq. 4.1) along with the mean phase coherence R (Eq. A.14) between $x-$ components of the Lorenz oscillators (Eq. A.8). In the second part

of this section we investigate the ability of both measures to characterize the strength of interactions between noise-contaminated signals.

In order to generate Lorenz time series we integrated a system of differential equations (Eq. A.8) using a fourth order Runge-Kutta algorithm with an integration step $dt = 0.005$ and then downsampled the obtained data to the sampling rate $d\tilde{t} = 0.03$. The initial conditions were normally distributed with zero mean and unit variance, and in order to eliminate transients, the first 10^4 iterations were discarded. Coupling strength c_{12} was gradually increased from 0 to 10 with the step $\delta c_{12} = 0.25$. For each coupling strength 20 realizations were generated with normally distributed *Rayleigh* numbers $R_j \in \mathbf{N}(28, 0.5)$. Let x_n^j $(n = 1, \ldots, N)$ denote a pair of real-valued time series of length $N = 10^4$ representing $x-$components of the first $(j = 1)$ and second $(j = 2)$ Lorenz oscillator and π_n^j $(n = 1, \ldots, \tilde{N}$; $\tilde{N} = N - (m-1)\tau)$ a pair of series of corresponding permutation symbols which form two symbol series S_1 and S_2. Both series were formed by setting different embedding dimensions $m \in [2, 5]$. Following [SL08] we chose the embedding window such as to approximately cover a basic period of the Lorenz oscillator T_{Lor}, i.e., $(m-1)\tau \approx T_{Lor}$. With the used sampling rate $d\tilde{t} = 0.03$ the basic period of the Lorenz oscillator can be estimated as $T_{Loz} \approx 20$ integration steps (or, in absolute units as $T_{Loz}d\tilde{t} = 0.6$, see Fig 4.1). This defines a possible range of the delay time as $\tau \approx 20/(m-1)$. In the following analysis we set the delay time to $\tau = 10$.

In order to characterize the strength of interactions between Lorenz oscillators we estimated the degree of zero-lag (static) correlations between two series of permutation symbols S_1 and S_2 by computing the zero-lagged mutual information between them (see Eq. 2.11),

$$MI(S_1, S_2) = \frac{1}{m-1}(H(\pi_i^{(1)}) + H(\pi_i^{(2)}) - H(\pi_i^{(1)}, \pi_i^{(2)})), \tag{4.1}$$

where $i \in [1, \tilde{N}]$. $H(\pi_i^{(1)})$ and $H(\pi_i^{(2)})$ are the Shannon entropies of the symbol series S_1 and S_2 and $H(\pi_i^{(1)}, \pi_i^{(2)})$ is the joined Shannon entropy. Defined in this way the symbolic mutual information $MI(S_1, S_2)$ provides an amount of common information (measured in bits) between two symbol series S_1 and S_2. Similar to the analysis made in the previous chapter (see section 3.1) to compute Shannon entropies we estimated the corresponding empirical probabilities $\hat{p}(\pi_i^{(1)})$, $\hat{p}(\pi_i^{(2)})$, and $\hat{p}(\pi_i^{(1)}, \pi_i^{(2)})$ of these symbols to occur as relative frequencies, i.e.,

$$\hat{p}(\pi_i^{(j)}) = \frac{W(\pi_i^{(j)})}{\tilde{N}} \tag{4.2}$$

and

$$\hat{p}(\pi_i^{(1)}, \pi_i^{(2)}) = \frac{W(\pi_i^{(1)}, \pi_i^{(2)})}{\tilde{N}},$$

where $W(\pi_i^{(j)})$ denotes the number of occurrences of a symbol $\pi_i^{(j)}$ in string S_1 $(j = 1)$ and string S_2 $(j = 2)$, respectively. $W(\pi_i^{(1)}, \pi_i^{(2)})$ denotes the number of simultaneous occurrences of a pair of symbols $\pi_i^{(1)}$ and $\pi_i^{(2)}$ in both strings.

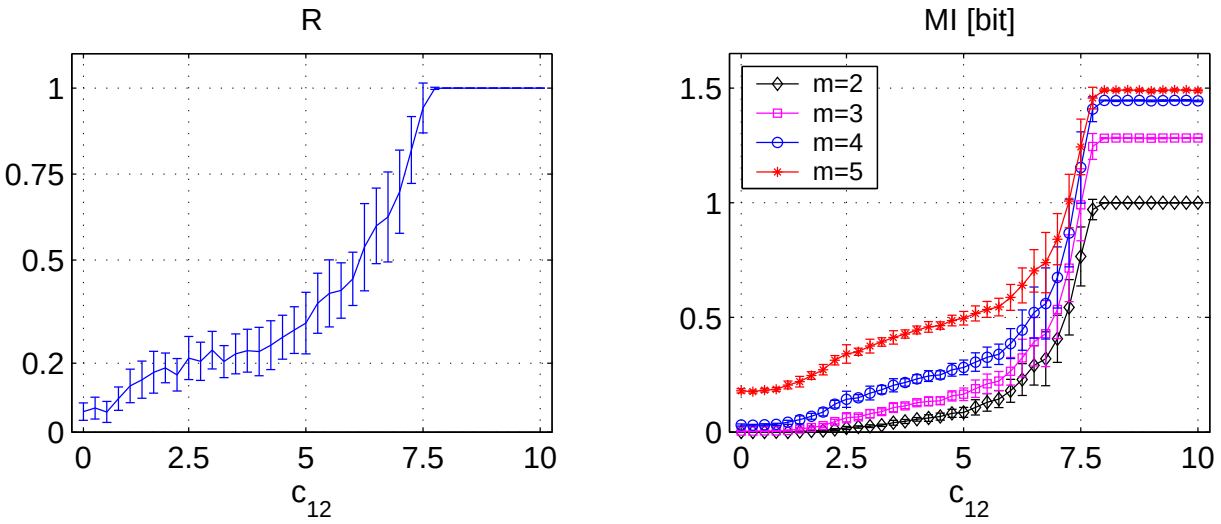


Figure 4.2.: Mean phase coherence R (left) and symbolic mutual information MI (right) between $x-$ components of two interacting Lorenz oscillators for increasing coupling strength c_{12}. Mean values and standard deviations (shown as error bars) are estimated from 20 realizations of Lorenz time series. In order to estimate mutual information the $x-$ components of both oscillators were first transformed into permutation symbols with different values of embedding dimension m and a fixed value of the delay time $\tau = 10$.

In order to characterize the strength of interactions between two coupled Lorenz oscillators we additionally estimated the degree of phase synchronization between $x-$ components of the corresponding chaotic oscillators. According to its definition, the mean phase coherence R is confined to the interval $[0, 1]$, where $R = 1$ indicates the case of fully phase-synchronized signals. Following [MLDE00] we applied a Hilbert transform to extract a phase distribution of the $x-$ components of both Lorenz oscillators and then computed the mean phase coherence R according to Eq. A.14 (for more details see section A.2).

Figure 4.2 shows the dependencies of the mean phase coherence R and the symbolic mutual information MI on the coupling strength c_{12} of two interacting Lorenz oscillators. For a zero coupling ($c_{12} = 0$) we observed that the mean phase coherence had a small positive bias despite the absence of interactions between oscillators which resulted from the finite sample effect. The estimation of the symbolic mutual information showed that its values also suffered from the finite sample effect and, moreover, for the embedding dimension $m = 5$ its values significantly deviated from zero ($MI(S_1, S_2) \approx 0.18 \pm 0.01$) even for $c_{12} = 0$ (for further details see [Rou99]). This finding indicated that the choice of the embedding dimension is limited to $m \leq 4$, at least, for the analysis of time series of length $N < 10^4$. For the case of non-zero coupling both measures showed a qualitatively similar dependence on the coupling strength and were gradually growing with an increasing coupling strength. The observed values of the symbolic mutual information were slightly higher for higher values of the embedding dimension m for the intermediate coupling strengths ($c_{12} \in [2, 7.5]$). This might imply that the permutation symbols obtained with higher values of the embed-

ding dimension m contain an additional amount of information about the dynamics. At a strong coupling ($c_{12} > 7.5$) the Lorenz oscillators were getting fully synchronized and the mean phase coherence approached its maximum ($R = 1$) whereas the symbolic mutual information approached the permutation entropy rate h_m defined in [BP02]. Indeed, for fully synchronized oscillators one obtains $S_1 = S_2$ and according to Eq. 4.1 the symbolic mutual information $MI(S_1, S_1) = \frac{1}{m-1} H(\pi_i^{(1)}) = h_m(S_1)$. As it was shown in [AKK05] for ergodic dynamical systems the permutation entropy rate h_m asymptotically converges to KS-entropy with $m \to \infty$ [2]. Thus, it is reasonable to assume that the symbolic mutual information MI asymptotically ($m \to \infty$) provides an estimate for the amount of common (mutual) information (measured in bits per a time step) shared between two interacting dynamical systems.

Analysis of noisy time series

The influence of the amount of noise in data on the ability of the symbolic mutual information MI to characterize the strength of interactions between signals is a very important factor for the analysis of field data. In the following, we generated 20 realizations (using the same integration parameters as in the previous section) of $x-$ components of two unidirectionally coupled ($c_{21} = 0$, $c_{12} > 0$) Lorenz oscillators (Eq. A.8) with a small ($c_{12} = 2$), intermediate ($c_{12} = 5$), and strong ($c_{12} = 10$) coupling strength. For each coupling strength we gradually increased the noise-to-signal ratio[3] (NSR) from 0 to 2 and computed the symbolic mutual information MI. The same analysis was also performed with the mean phase coherence R. In order to perform a quantitative comparison between R and MI computed for different values of the coupling strength c_{12} and embedding dimension m we analyzed the relative, rather than absolute, values of the mean phase coherence and symbolic mutual information which we defined as R/R_0 and MI/MI_0 (where R_0 and MI_0 denote the values of mean phase coherence and symbolic mutual information for NSR = 0). According to the given definition, R/R_0 and MI/MI_0 are confined in the interval $[0, 1]$ and asymptotically approach 1 in case of noise-free data (NSR = 0) and approach 0 in case of a white noise time series (NSR $\to \infty$). However, a preliminary investigation revealed that both measures remained positive even for the limiting case NSR $\to \infty$ and approached zero by increasing the length of the time series. This can be explained as a result of a finite sample effect. In order to investigate its influence on obtained numerical estimates of both measures we additionally computed R_∞/R_0 and MI_∞/MI_0 (where R_∞ and MI_∞ denote the mean phase coherence and symbolic mutual information for NSR $\to \infty$). Figure 4.3 shows the dependencies of R/R_0 and MI/MI_0 on the noise-to-signal ratio. We observed that both measures gradually declined with an increasing amount of noise in the time series until they approached minimal values which were given by R_∞/R_0 and MI_∞/MI_0 respectively (shown as dashed lines in Fig. 4.3). A quantitative comparison between R/R_0 and MI/MI_0 revealed that the relative symbolic mutual information approached MI_∞/MI_0 already for

[2] However, the convergence of the permutation entropy rate to KS-entropy for chaotic oscillators has not yet been shown explicitly.

[3] The noise-to-signal ratio is defined as $\sigma_{noise}/\sigma_{signal}$, where σ_{noise} and σ_{signal} denote the standard deviations of noise and signals respectively.

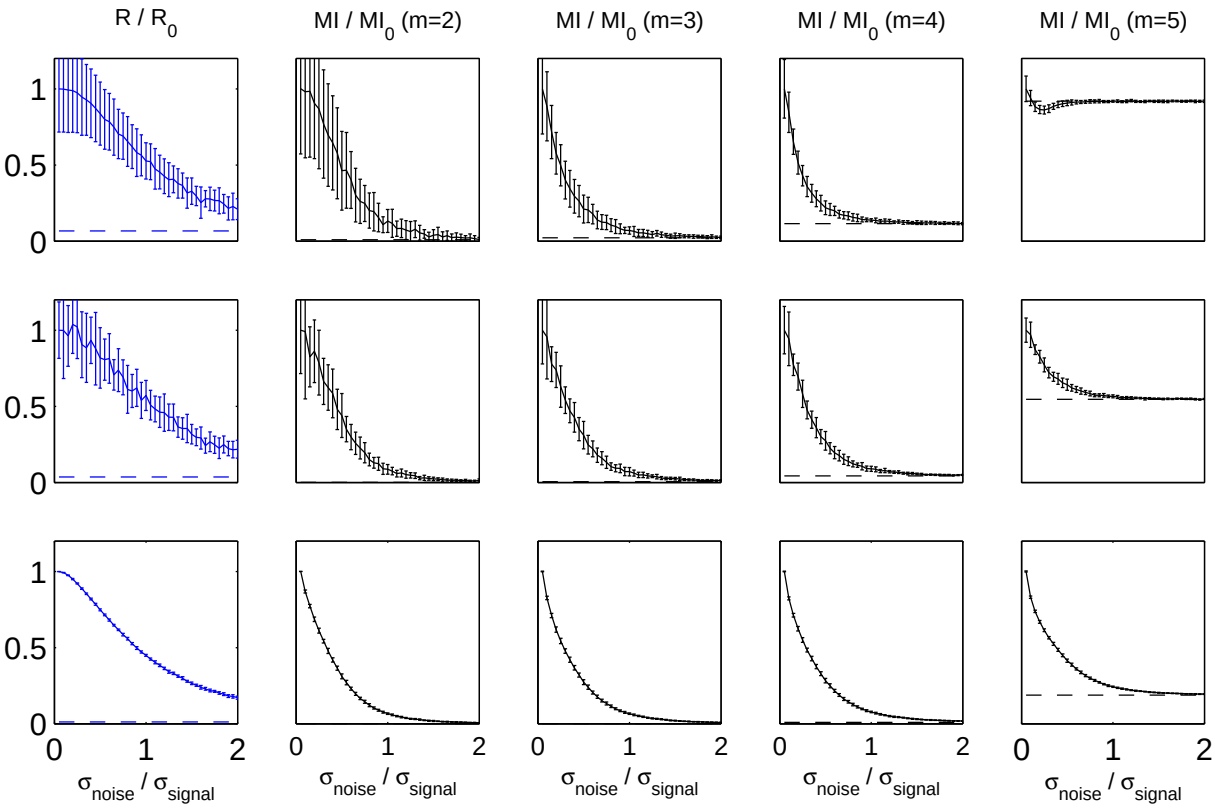


Figure 4.3.: Relative mean phase coherence R/R_0 and relative symbolic mutual information MI/MI_0 between noise-contaminated $x-$ components of two interacting Lorenz oscillators for small $c_{12} = 2$ (top), intermediate $c_{12} = 5$ (middle), and strong $c_{12} = 10$ (bottom) coupling strengths. Mean values and standard deviations (shown as error bars) are estimated from 20 realizations of Lorenz time series. Here, R_0 and MI_0 denote the mean phase coherence and the symbolic mutual information for noise-free data. The dashed line indicates the relative values of both measures computed for the limiting case NSR $\to \infty$.

NSR ≈ 1 (weak coupling), NSR ≈ 1.5 (intermediate coupling), NSR ≈ 1.9 (strong coupling) and for embedding dimensions $m \leq 4$ whereas the relative mean phase coherence remained well above R_∞/R_0 for all investigated values of NSR and coupling strengths. Additionally, we observed that statistical errors of the symbolic mutual information can, in general, be reduced by increasing the embedding dimension m especially for weak and intermediate couplings. Further analysis revealed that special care should be taken by considering the symbolic mutual information computed with the embedding dimension $m = 5$. In this case the finite sample effect appeared to have a rather significant influence on the obtained estimates of the symbolic mutual information and MI/MI_0 approached MI_∞/MI_0 already for NSR ≈ 0.1 (weak coupling), NSR ≈ 1 (intermediate coupling) and NSR ≈ 1.4 (strong coupling). Thus, already a tiny amount of noise in the data can limit the characterization of the strength of interactions in weakly coupled Lorenz oscillators by means of the symbolic

mutual information computed with $m = 5$.
In this section we analyzed the strength of interactions between two coupled Lorenz oscillators by measuring the symbolic mutual information MI and, for comparative purposes, the mean phase coherence R for different values of the coupling strength and noise-to-signal ratios. Capturing different aspects[4] of the dynamics both measures nevertheless showed a qualitatively similar dependence on the coupling strength. Analysis of the dependence of the symbolic mutual information on the embedding dimension m revealed that its estimates obtained with higher values of m can, in general, capture more information about the underlying dynamics. However, we also observed that for an embedding dimension $m = 5$ the obtained estimates of the symbolic mutual information deviated from zero even for the case of uncoupled Lorenz oscillators, at least, for time series with length $N = 10^4$. Analysis of Lorenz time series with different noise-to-signal ratio indicated that the mean phase coherence R is, in general, less sensitive to the amount noise in data than the symbolic mutual information MI which can only be reliably applied up to $NSR \approx 1$ (at least for the time series of length $N < 10^4$). This observation can be explained if we recall the definition of the symbolic mutual information (Eq. 4.1) which characterizes the amount of common (mutual) information between a pair of permutation symbols. In noise-contaminated signals with $NSR > 1$ the dynamical structure of order patterns can be significantly impaired leading to a randomization of permutation symbols and thus to a reduction of mutual information between a pair of symbol series (see section 2.3.1).
In the next section we will analyze the directionality of interactions between coupled dynamical model systems by measuring the symbolic transfer entropy (see Eq. 4.5 and [SL08]) which extends the notion of the symbolic mutual information and allows to estimate the amount of entropy transfered between dynamical systems in a unit of time. In the following we will briefly review the concept for causality given by Granger and discuss its relationship with entropy transfer between stochastic processes.

4.2. Characterizing directionality of interactions with symbolic transfer entropy

Granger causality

An operational definition for causality has been discussed in detail by Wiener and later formalized by Granger [Gra01]. Granger proposed a list of restrictions which the notion of causality should fulfill to be logically consistent. According to Granger's definition two events (let us say an event E_2 chronologically follows an event E_1) are considered as causally interconnected when the forecast error for event E_2 can be reduced with the knowledge of the outcome of E_1. Following the work of Granger, let us assume that an investigated system is isolated form the environment and is completely described by a set of discrete-time stochastic processes $\{X, Y, Z, \ldots\}$ and its state at time step n is given by $\Omega_n \equiv \{x_n, y_n, z_n, \ldots\}$. The reduced set of variables $\acute{\Omega}_n \equiv \{x_n, z_n, \ldots\}$ defines the state of the

[4]The mean phase coherence R characterizes the phase relationships between two signals whereas the symbolic mutual information MI characterizes the amount of common information between them.

system at time step n excluding y_n. Now, we can state that X causally follows Y if and only if

$$\Delta(x_{n+1}|\Omega_n) < \Delta(x_{n+1}|\acute{\Omega}_n) \tag{4.3}$$

where $\Delta(x_{n+1}|\Omega_n)$ and $\Delta(x_{n+1}|\acute{\Omega}_n)$ are the forecast errors of process X at time step $n+1$ when Ω_n and $\acute{\Omega}_n$ are given. An important contribution of Granger was his set of axioms which have to be fulfilled for the logically consistent definition of causality:

- *Axiom A.* The past and present may cause the future, but the future cannot cause the past.
- *Axiom B.* Ω_n contains no redundant information, i.e., if some variable Z is functionally related to one or more other variables in a deterministic fashion (e.g. $Z = F(X, Y, \ldots)$), then Z should be excluded from Ω_n.
- *Axiom C.* Ω_n contains complete information, i.e., if some not redundant variables W is missing in Ω_n the correct inference of causality using the definition given above (Eq. 4.3) cannot be guaranteed.

To provide an operational mathematical definition of causality Granger exploited the framework of autoregressive processes. However, using an information-theoretic approach to describe stochastic processes the definition of Granger causality can be reformulated as

$$H(x_{n+1}|\Omega_n) < H(x_{n+1}|\acute{\Omega}_n) \tag{4.4}$$

where $H(x_{n+1}|\Omega_n)$ denotes a conditional entropy of the stochastic process X at time step n when Ω_n or $\acute{\Omega}_n$ are given [HSPVB07, BBS09]. In the simplest possible case when only two stochastic processes X and Y are considered, i.e., when $\Omega_n \equiv \{x_n, y_n\}$ and $\acute{\Omega}_n \equiv \{x_n\}$ the condition (Eq. 4.4) can be reduced to $TE(X,Y) = H(x_{n+1}|x_n) - H(x_{n+1}|x_n, y_n) > 0$ (see Eq. 2.17). This implies that at least under the framework of discrete stochastic processes the positivity of the transfer entropy $TE(X,Y)$ between two stochastic processes indicates Granger causality between X and Y. However, it has to be pointed out that in case of multivariate data (i.e., when we consider more than two stochastic processes) the positivity of the original, bivariate version of the transfer entropy ($T(X,Y) > 0$) cannot guarantee Granger causality between X and Y (axiom C).

4.2.1. Corrected symbolic transfer entropy

The task of inferring causal or directional interactions between dynamical systems from experimental signals is a very challenging and important scientific problem. According to the definition of causality given above the existence of such interactions between two dynamical systems can be identified by characterizing the amount of correlations between present and/or past states of the first (second) signal and future states of the second (first) signal correspondingly [Gra01]. These correlations can be called as *dynamic correlations* because they reflect the dynamical structure (evolution) of the signals. In many real-world

applications, the investigated signals can be functionally related to each other and exhibit correlations without a time delay. In contrast to dynamic correlations defined above such correlations do not reflect the dynamical structure and only characterize the similarity between signals. Following [Sch00] such correlations can be called as *static correlations*. As we could see above (see section 4.1) by measuring the amount of static correlations between two signals by means of the symbolic mutual information MI (see Eq. 4.1) it is possible to characterize the strength of interactions between dynamical systems. In general, this approach can easily be extended to measure dynamic correlations between signals and thus to infer the directionality of interactions between underlying dynamical systems. This can be done by measuring the time-lagged rather than the zero-lag (as we did in section 4.1) mutual information between a pair of series of permutation symbols. The obtained measure can be called time-lagged symbolic mutual information and it characterizes the amount of information shared between a present state of the first (second) signal and a future state of the second (first) signal respectively. However, as we just mentioned above, the investigated signals can additionally exhibit a high degree of static correlations. This violates second axiom in the definition of causality given by Granger. For instance, in the limiting case of two identical signals, the time-lagged symbolic mutual information can be positive and thus can indicate the existence of directional interactions despite the fact that both signals do not contain any additional information about each other. The notion of transfer entropy proposed in [Sch00] as well as of the conditional mutual information given in [PV07] extend the notion of time-lagged mutual information allowing to characterize dynamic correlations and thus to infer the directionality of interactions in signals exhibiting a high degree of static correlations.

In this section we will analyze the directionality of interactions between coupled nonlinear dynamical systems by measuring entropy transfer between their components. By comparing the obtained values of entropy transfer in both directions allows us to infer the directionality of interactions between two dynamical systems. In the beginning of the section, by following [SL08], we provide an algorithm for the estimation of symbolic transfer entropy between two real-valued time series. Next, we will introduce a correction scheme for the symbolic transfer entropy that takes into account static correlations between time series. We start the numerical part of our analysis by considering two structurally identical as well as structurally non-identical unidirectionally interacting Hénon maps (Eqs. A.6 and A.7). We repeat our analysis by considering two unidirectionally interacting identical Lorenz oscillators (Eq. A.8) as well as a system of unidirectionally coupled Rössler (driver) and Lorenz (responder) oscillators (see Eq. A.10).

Symbolic transfer entropy

Let x_n^j $(n = 1, \ldots, N)$ denote a pair $(j = 1, 2)$ of time series of length N and $\pi_n^{(j)}$ $(n = 1, \ldots, \tilde{N};\ \tilde{N} = N - (m-1)\tau)$ a pair of series of corresponding permutation symbols which form strings $S_1 \equiv \left\{\pi_n^{(1)}\right\}_{n=1}^{\tilde{N}}$ and $S_2 \equiv \left\{\pi_n^{(2)}\right\}_{n=1}^{\tilde{N}}$. Both strings are formed by setting an embedding dimension m and delay time τ. Assuming that the resulting series of permutation symbols exhibit realizations of two interacting stochastic processes and ac-

cording to the definitions given in [Sch00, SL08] (also see Eqs. A.31 and A.32) the symbolic transfer entropy from S_1 to S_2 of order k ($k \geq 1$) is defined as

$$\begin{aligned} TE^k(S_1,S_2) = \frac{1}{m-1}[H(\pi^{(2)}_{i+1},\ldots,\pi^{(2)}_{i-k+1}) - H(\pi^{(2)}_{i},\ldots,\pi^{(2)}_{i-k+1}) \\ -H(\pi^{(2)}_{i+1},\pi^{(2)}_{i},\pi^{(1)}_{i},\ldots,\pi^{(2)}_{i-k+1},\pi^{(1)}_{i-k+1}) + H(\pi^{(2)}_{i},\pi^{(1)}_{i},\ldots,\pi^{(2)}_{i-k+1},\pi^{(1)}_{i-k+1})], \end{aligned} \tag{4.5}$$

where $i \in [1, \tilde{N}-k]$. It is known that the transfer entropy represents a combination of four different block-entropies[5] [KS02]. In order to compute the symbolic transfer entropy of order k the corresponding $k+1$, k, $2k+1$, and $2k$ dimensional probability distributions: $p(\pi^{(2)}_{i+1},\ldots,\pi^{(2)}_{i-k+1})$, $p(\pi^{(2)}_{i},\ldots,\pi^{(2)}_{i-k+1})$, $p(\pi^{(2)}_{i+1},\pi^{(2)}_{i},\pi^{(1)}_{i},\ldots,\pi^{(2)}_{i-k+1},\pi^{(1)}_{i-k+1})$, and $p(\pi^{(2)}_{i},\pi^{(1)}_{i},\ldots,\pi^{(2)}_{i-k+1},\pi^{(1)}_{i-k+1})$ have to be known. In our work, by following [SL08], we used the *plug-in* estimators of block-entropies [HSPVB07]. These estimators are obtained by using the empirical probability distributions which are defined as relative frequencies of occurrence of corresponding blocks of permutation symbols (words) in the strings S_1 and S_2 (for more details see section A.4).

It is known that for a finite symbol series of length $\tilde{N}$ the *plug-in* estimator of the block-entropy may suffer from systematic and statistical (random) errors (for details see [Gra88, HSE94, SG96, Rou99]). This results from the fact that for small $\tilde{N}$ the empirical probability distribution can, in general, be undersampled and thus cannot be independent on the length $\tilde{N}$, by only asymptotically converging to the underlying probability distribution when $\tilde{N} \to \infty$. In some cases, when the form of the underlying probability distribution is known, the resulting systematic errors can be corrected even for a very short symbol series [Gra88, HSE94, Rou99]. However, for field applications where one usually deals with an arbitrary symbol series with an unknown underlying probability distribution the corrections schemes cannot be applied. In this case the length of symbol series has to be large enough to reduce systematic and statistical errors of the *plug-in* estimator to a satisfactory level where the effects of the undersampling of the empirical probability distribution can be neglected. To proceed, we follow [Kre99] and define the so called statistical quotient $SQ = \tilde{N}/W$, where W denotes the number of different symbols (or different blocks of symbols) which appear in symbol series of length $\tilde{N}$. In this work, by following [Kre99], we will consider that the systematic and statistical errors of the *plug-in* estimator of the block-entropy of the symbol series are reduced to a satisfactory level when the statistical quotient $SQ \geq 10$. Under this condition, the number of occurrence of each symbol (or block of symbols) is, on average, set to 10 and the empirical probability distribution is reasonably good filled.

According to the definition (Eq. 4.5), in order to compute the symbolic transfer of order k one needs to estimate $(2k+1)-$dimensional probability distributions of permutations symbols. Since the maximum possible number of different permutation symbols is $m!$ the number of possible words of length $2k+1$ can be analytically computed as $W = (m!)^{2k+1}$. Thus, for the symbol series of length $\tilde{N}$ the statistical quotient of the symbolic transfer entropy of order k can be defined as

[5]The symbolic transfer entropy in the opposite direction $TE^k(S_2,S_1)$, i.e., from the string S_2 to S_1, is defined in the analogous way.

$$SQ = \frac{\tilde{N}}{(m!)^{2k+1}}. \tag{4.6}$$

For the symbol series of length $\tilde{N}$ the number of possible symbols (or words) is alway limited, i.e., $W \leq \tilde{N}$. This implies that the statistical quotient, by definition, $SQ \geq 1$. Thus, in the case when $(m!)^{2k+1} > N$, the number of different symbols (words) W has to be numerically estimated (see e.g. section 6.1).

Corrected symbolic transfer entropy

Let us assume that the series of permutation symbols S_1 and S_2 are realizations of two interacting stochastic Markov processes of first[6] order Z_1 and Z_2. By assuming that the permutation symbols are chronologically-ordered we can denote $\pi_{i-1}^{(j)}$ as a *past*, $\pi_i^{(j)}$ as a *present*, and $\pi_{i+1}^{(j)}$ as a *future* state of processes Z_j ($j = 1, 2$ and $\forall i \in [2, \tilde{N} - 1]$). According to the definition of a first order Markov process given in section 2.2.1 the future states $\pi_{i+1}^{(j)}$ are uniquely defined by present states $\pi_i^{(j)}$ and independent on the complete history of both processes Z_j ($j = 1, 2$). In this case the definition of the symbolic transfer entropy given by Eq. 4.5 can be reduced to the symbolic transfer entropy of first order which is defined as

$$TE(S_1, S_2) = \frac{1}{m-1}[H(\pi_{i+1}^{(2)}, \pi_i^{(2)}) - H(\pi_i^{(2)}) - H(\pi_{i+1}^{(2)}, \pi_i^{(2)}, \pi_i^{(1)}) + H(\pi_i^{(2)}, \pi_i^{(1)})]. \tag{4.7}$$

It is more convenient now to rewrite[7] this definition of the symbolic transfer entropy of first order as

$$TE(S_1, S_2) = \frac{1}{m-1}[MI(\pi_i^{(1)}, \pi_{i+1}^{(2)}) - MI(\pi_i^{(1)}, \pi_i^{(2)}, \pi_{i+1}^{(2)})]. \tag{4.8}$$

In this form symbolic transfer entropy can be seen as a two-point (time-lagged) mutual information $MI(\pi_i^{(1)}, \pi_{i+1}^{(2)})$ with subtracted three-point mutual information $MI(\pi_i^{(1)}, \pi_i^{(2)}, \pi_{i+1}^{(2)})$. In the case where both series S_1 and S_2 do not have static correlations (i.e., the present states of both systems are not functionally related and axiom B defined by Granger is fulfilled) we have $MI(\pi_i^{(1)}, \pi_i^{(2)}) = 0$. This immediately implies that $MI(\pi_i^{(1)}, \pi_i^{(2)}, \pi_{i+1}^{(2)}) = 0$ (see definition of three-point mutual information in section A.3) and thus $TE(S_1, S_2) = MI(\pi_i^{(1)}, \pi_{i+1}^{(2)})$. The symbolic transfer entropy just resembles the two-point mutual information between the present state of Z_1 ($\pi_i^{(1)}$) and the future state of Z_2 ($\pi_{i+1}^{(2)}$) (causal or directional correlations; see Fig.4.4). However, in field applications, Granger's axiom B (see section 4.2) can, in general, not be fulfilled and time series are very often functionally

[6] In this chapter we will only consider the symbolic transfer entropies of first order ($k = 1$) and, for simplicity, the index k will be omitted, i.e., $TE^1(S_1, S_2) \equiv TE(S_1, S_2)$.

[7] Indeed, according to the definitions of two- and three-point mutual information functions (Eqs. A.21 and A.24; see also Fig. A.2) we have $TE(S_1, S_2) = MI(\pi_i^{(1)}, \pi_{i+1}^{(2)}) - MI(\pi_i^{(1)}, \pi_i^{(2)}, \pi_{i+1}^{(2)}) = [H(\pi_i^{(1)}) + H(\pi_{i+1}^{(2)}) - H(\pi_i^{(1)}, \pi_{i+1}^{(2)})] - [H(\pi_i^{(1)}) + H(\pi_i^{(2)}) + H(\pi_{i+1}^{(2)}) - H(\pi_i^{(1)}, \pi_i^{(2)}) - H(\pi_{i+1}^{(2)}, \pi_i^{(2)}) - H(\pi_i^{(2)}, \pi_i^{(1)}) + H(\pi_{i+1}^{(2)}, \pi_i^{(2)}, \pi_i^{(1)})] = H(\pi_{i+1}^{(2)}, \pi_i^{(2)}) - H(\pi_i^{(2)}) - H(\pi_{i+1}^{(2)}, \pi_i^{(2)}, \pi_i^{(1)}) + H(\pi_i^{(2)}, \pi_i^{(1)})$.

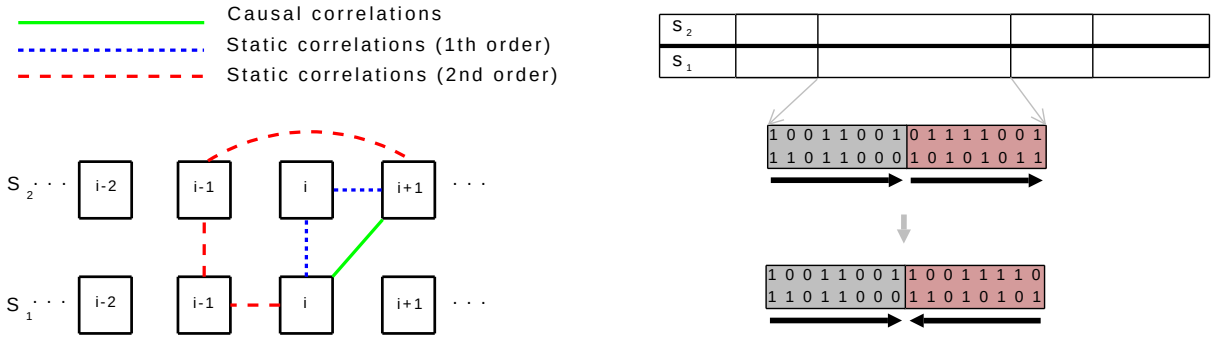


Figure 4.4.: Left: Schematic representation of causal (direct) correlations along with static (indirect) correlations of first and second order between symbol series S_1 and S_2. Right: Schematic representation of a time-inversion procedure which is used to generate MI-preserving surrogates.

related[8]. This was a main motivation by Schreiber to extend the notion of time-lagged mutual information to the notion of transfer entropy. Indeed, according to the original definition of transfer entropy " ... *transfer entropy is able to detect the directed exchange of information between two systems. Unlike [time-lagged] mutual information, it is designed to ignore static correlations due to the common history* ..." [Sch00]. Due to the static correlations between S_1 and S_2 (i.e., $MI(\pi_i^{(1)}, \pi_i^{(2)}) > 0$) the present state of Z_1 ($\pi_i^{(1)}$) and the future state of Z_2 ($\pi_{i+1}^{(2)}$) can be indirectly interrelated (Fig. 4.4). These indirect correlations can be quantified by measuring many-point mutual information functions (see section A.3). For example, to characterize static correlations of first order a three-point mutual information $MI(\pi_i^{(1)}, \pi_i^{(2)}, \pi_{i+1}^{(2)})$ has to be taken into account. Thus, in order to measure the amount of causal or directional correlations from S_1 to S_2 we have to compute the time-lagged mutual information $MI(\pi_i^{(1)}, \pi_{i+1}^{(2)})$ and then subtract the three-point mutual information $MI(\pi_i^{(1)}, \pi_i^{(2)}, \pi_{i+1}^{(2)})$. This provides us a definition of symbolic transfer entropy of first order from S_1 to S_2 (Eq. 4.8). In order to correct the influence of static correlations of higher orders the symbolic transfer entropy of corresponding orders have to computed (Eq. 4.5).

As was originally pointed out by Schreiber the estimation of transfer entropies of higher orders requires a large amount of data which is not always available in field applications [Sch00]. An exponential growth of the number of possible symbols, from one side, and a limited amount of data in most practical applications, from the other side, limits the applicability of transfer entropies of higher orders. Thus, for most of practical applications the transfer entropy of first order (i.e., $k = 1$, see Eq. 4.5) is the only choice and thus the effect of high order static correlations cannot be avoided (see Fig. 4.4). Analysis of the influence of the finite sample effect on numerical estimates of transfer entropy led to the concept of *effective* transfer entropy (ETE) proposed in [MK02]. The authors defined TE_e as a difference of the usual transfer entropy $TE(S_1, S_2)$ calculated for the series S_1 and S_2 and transfer entropy $TE_{shuf}(S_1, S_2)$ calculated for surrogate series obtained by a random

[8]For example, two interacting chaotic oscillators exhibiting generalized synchronization [PRK01].

shuffling of series S_1 (series S_2 is kept unchanged). A random shuffling of S_1 destroys correlations between both symbol series and the positivity of $TE_{shuf}(S_1, S_2)$ characterizes the amount of transfer entropy that is a result of the finite sample effect. In our work we attempt to extend the notion of effective transfer entropy and develop a correction scheme that additionally takes into account the influences of mutual static correlations in symbol series. To do this, we introduce a so called time-inversion procedure (see Fig. 4.4) that is used to generate a MI-preserving surrogate symbol series. In order to obtain a MI-preserving surrogate a bivariate symbol series (consisting of two series of permutation symbols S_1 and S_2) is first randomly divided into $N_{seg} = 5$ non-overlapping segments of different and randomly chosen lengths. Each segment is then divided into two equal parts. The first part of the segment is kept unchanged whereas the order of symbols in the second part is reversed. By definition, this transformation does not change the empirical probabilities $\hat{p}(\pi_i^{(1)})$, $\hat{p}(\pi_i^{(2)})$ and $\hat{p}(\pi_i^{(1)}, \pi_i^{(2)})$ (defined by Eq. 4.2) and therefore does not change the estimates for the Shannon entropies $H(\pi_i^{(1)})$ and $H(\pi_i^{(2)})$ as well as the joint Shannon entropy $H(\pi_i^{(1)}, \pi_i^{(2)})$. From this follows that the symbolic mutual information MI defined in Eq. 4.1 is also preserved. In contrast to surrogate series used in [MK02] (where all correlations between two symbol series are destroyed), the time-inversion procedure preserves static correlations between symbol series while destroying dynamic correlations between them. We define a *corrected* symbolic transfer entropy $TE_c(S_1, S_2)$ as the difference of the usual symbolic transfer entropy $TE(S_1, S_2)$ calculated for the series S_1 and S_2 and the symbolic transfer entropy $TE_s(S_1, S_2)$ calculated for MI-preserving symbolic surrogate series. Following the definition of effective transfer entropy TE_e given in [MK02] the corrected symbolic transfer entropy $TE_c(S_1, S_2)$ is defined as

$$TE_c(S_1, S_2) = \begin{cases} TE(S_1, S_2) - TE_s(S_1, S_2), \text{ if } TE(S_1, S_2) \geq TE_s(S_1, S_2) \\ 0, \text{ otherwise.} \end{cases} \tag{4.9}$$

Thus, the positivity of symbolic transfer entropy computed between MI-preserving surrogate symbol series (i.e., $TE_s > 0$) characterizes the amount of transfer entropy that is a result of high order static correlations. In the next section we will analyze entropy transfer between coupled model dynamical systems. Additionally, in order to investigate the influence of the finite sample effect and high order static correlations we will estimate TE_{shuf} and TE_s and finally compute the corrected symbolic transfer entropy TE_c defined in Eq. 4.9.

4.2.2. Entropy transfer between time series of dynamical model systems

In this section we analyze the directionality of an interaction by measuring entropy transfer between structurally identical and structurally non-identical unidirectionally coupled chaotic maps and chaotic oscillators for different values of the coupling strength. In order to estimate entropy transfer we computed the symbolic transfer entropy TE as well as corrected symbolic transfer entropy TE_c of first order (Eqs. 4.7 and 4.9; see section 4.2.1) between $x-$ components of investigated dynamical systems for different values of the embedding parameters.

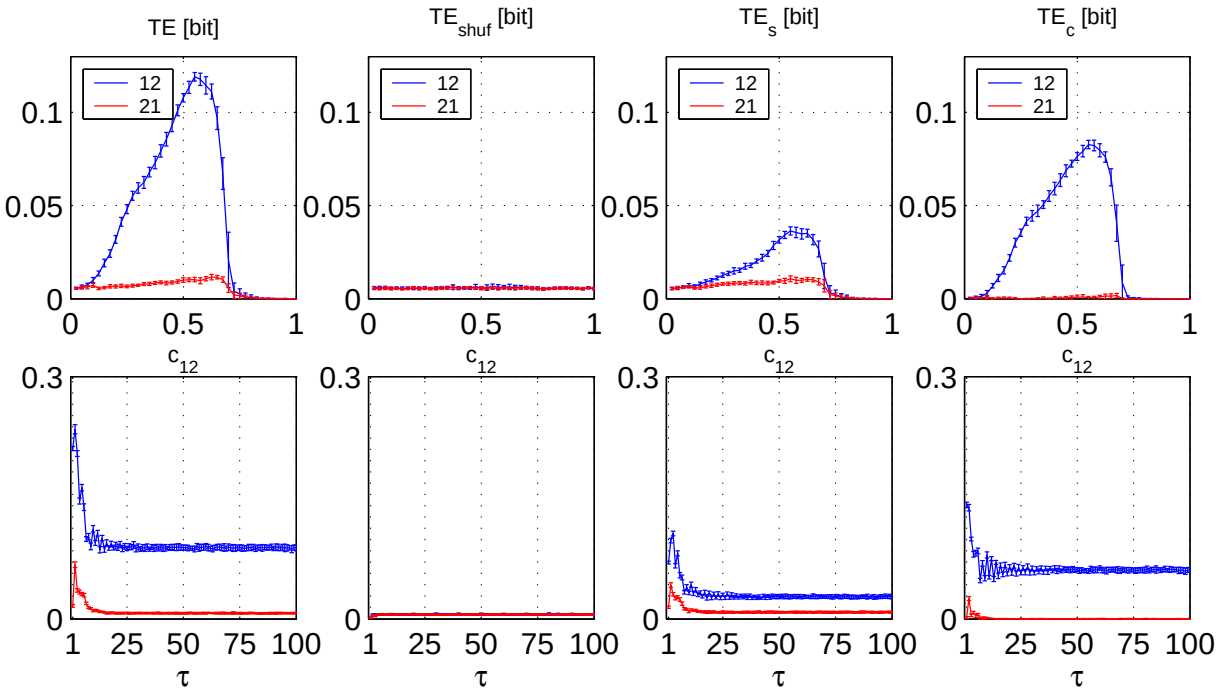


Figure 4.5.: Upper row: Symbolic transfer entropies TE and corrected symbolic transfer entropies TE_c between $x-$ components of two unidirectionally coupled Hénon maps for increasing coupling strength c_{12} and with fixed embedding parameters: $m = 3$, $\tau = 10$. Lower row: The same but now for the fixed coupling strength $c_{12} = 0.5$ and with varying embedding parameters: $m = 3$, $\tau \in [1, 100]$. TE_{shuf} and TE_s represent symbolic transfer entropies between surrogate symbol series obtained by either a random shuffling or time-inversion transformation of data. Each point on all plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

First, we considered a discrete time dynamical system such as two unidirectionally coupled, structurally identical ($b_1 = 0.3$, $b_2 = 0.3$) and non-identical ($b_1 = 0.3$, $b_2 = 0.1$) Hénon maps defined by Eqs. A.6 and A.7. With the used coupling scheme the first Hénon map (driver) was coupled into the second Hénon map (responder) with the coupling strength c_{12}. By using randomly chosen initial conditions (taken from the unit interval) and iterating them with Eqs. A.6 and A.7 we generated 20 realizations of real valued time series of length $N = 10^4$ for each value of the coupling strength c_{12}. The coupling strength c_{12} was gradually increased from 0 to 1 with a step $\delta c_{12} = 0.025$. Next, we considered continuous time dynamical systems such as two unidirectionally coupled chaotic oscillators. For the analysis of structurally identical dynamical systems we used a system of two unidirectionally coupled Lorenz oscillators defined by Eq. A.8 ($R_{1,2} \in \mathbf{N}(28, 0.5)$). According to this equation the first Lorenz oscillator (driver) was diffusively coupled into the second Lorenz oscillator (responder) with the coupling strength c_{12}. For the case of structurally non-identical dynamical systems we used a system of unidirectionally coupled Rössler (driver)

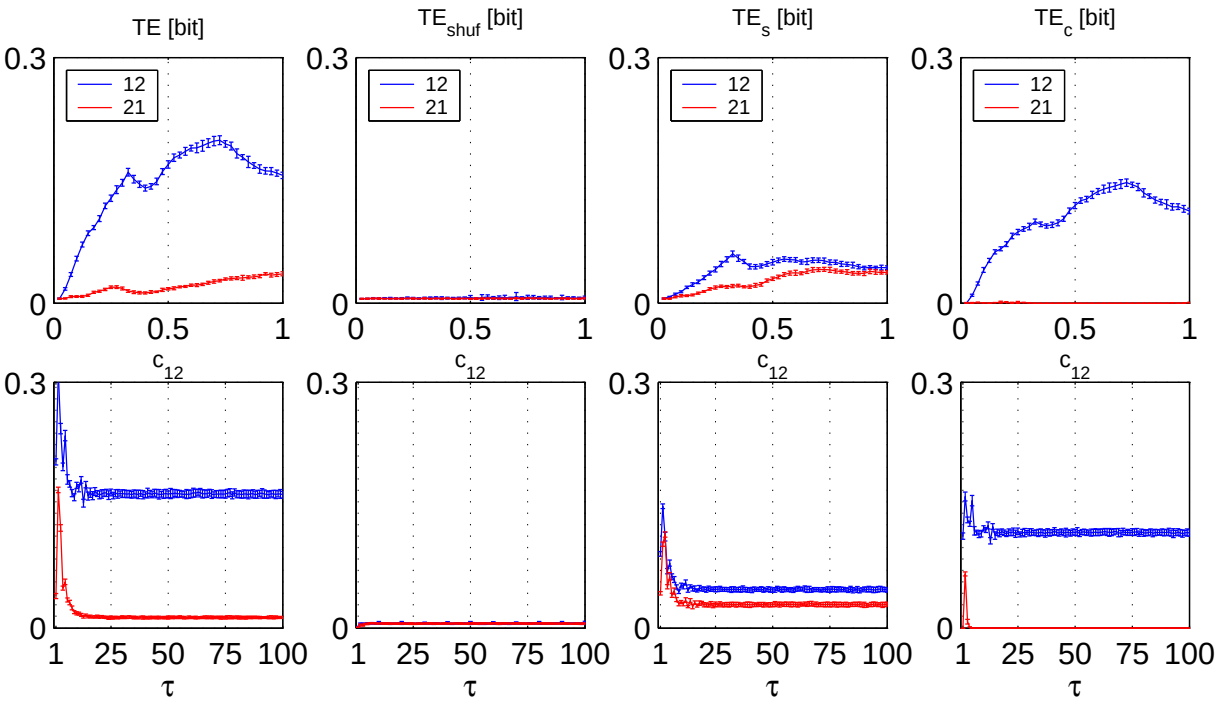


Figure 4.6.: Upper row: Symbolic transfer entropies TE and corrected symbolic transfer entropies TE_c between $x-$ components of two unidirectionally coupled, structurally different Hénon maps for increasing coupling strength c_{12} and with fixed embedding parameters: $m = 3$, $\tau = 10$. Lower row: The same but now for the fixed coupling strength $c_{12} = 0.5$ and with varying embedding parameters: $m = 3$, $\tau \in [1, 100]$. TE_{shuf} and TE_s represent symbolic transfer entropies between surrogate symbol series obtained by either a random shuffling or time-inversion transformation of data. Each point on all plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

and Lorenz (responder) oscillators defined by Eq. A.10 for different values of the coupling strength c_{rl}. In both cases the differential equations (Eq. A.8 and Eq. A.10) were integrated using a fourth order Runge-Kutta algorithm with integration step $dt = 0.005$ and then downsampled to $d\tilde{t} = 0.03$. The initial conditions were normally distributed with zero mean and unit variance. In order to eliminate transients, the first 10^4 iterations were discarded. For each coupling strength c_{12} (c_{rl}) we again generated 20 realizations of real valued time series of length $N = 10^4$. The coupling strengths c_{12} and c_{rl} were here gradually varied from 0 to 10 with a step $\delta c_{12} = \delta c_{rl} = 0.25$. Thus, with all given above coupling schemes one can expect to find a positive value of the entropy transfer in the direction "driver-to-responder" and a zero value of transfer entropy in the opposite direction.

According to the definition of the permutation entropy rate given in [BP02, BKP02, AKK05, AK07] its value converges to KS-entropy of ergodic dynamical systems in the limit $m \to \infty$ (see section 3). Thus, it is rather logical to assume that by increasing the embedding dimension m the symbolic transfer entropy will also converge to the actual entropy transfer

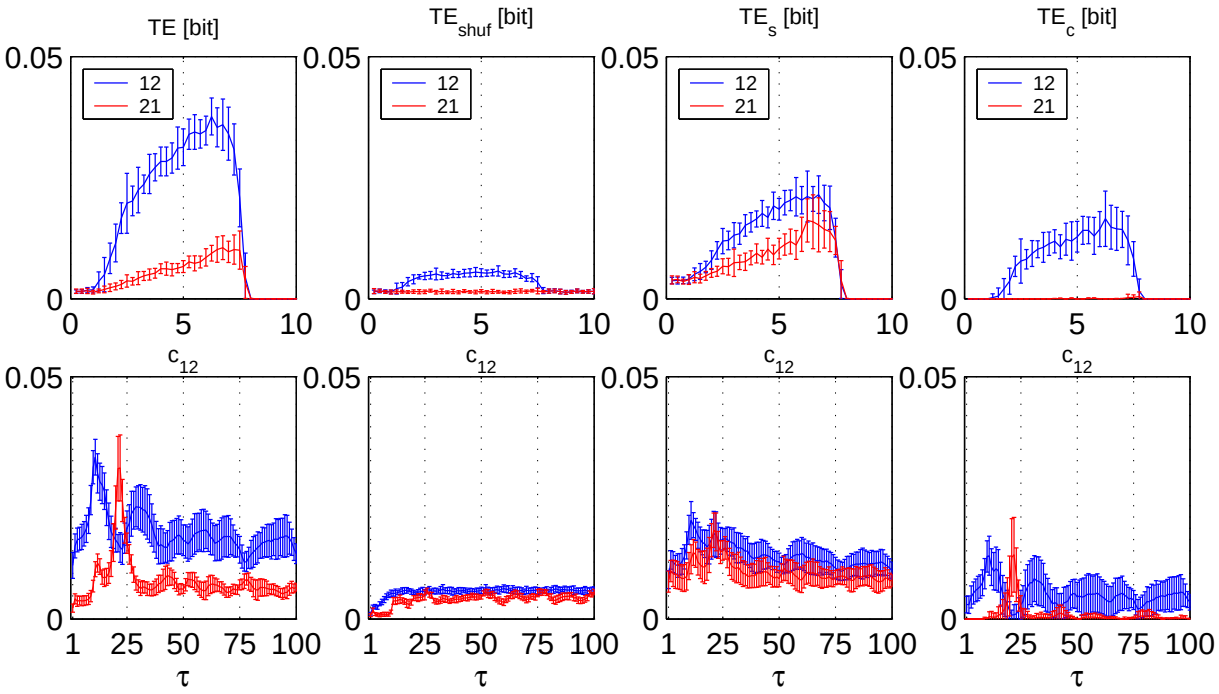


Figure 4.7.: Upper row: Symbolic transfer entropies TE and corrected symbolic transfer entropies TE_c between $x-$ components of two unidirectionally coupled Lorenz oscillators for increasing coupling strength c_{12} and with fixed embedding parameters: $m = 3$, $\tau = 10$. Lower row: The same but now for the fixed coupling strength $c_{12} = 5$ and with varying embedding parameters: $m = 3$, $\tau \in [1, 100]$. TE_{shuf} and TE_s represent symbolic transfer entropies between surrogate symbol series obtained by either a random shuffling or time-inversion transformation of data. Each point on all plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

between two dynamical systems. However, in many practical applications the amount of data is limited and the obtained estimates of the symbolic transfer entropy might be affected by systematic or statistical errors due to the finite length of the data. In order to minimize the influence of the systematic and statistical errors on the estimator of the symbolic transfer entropy of the first order ($k = 1$) we only considered the embedding dimension $m = 3$. For this embedding dimension the statistical quotient of the symbolic transfer entropy computed for time series of length $N = 10^4$ can be estimated as $SQ \approx 46$ that fulfills the condition $SQ \geq 10$, whereas already for $m = 4$ one obtains $SQ \approx 1$ (for details see section 4.2.1 and Eq. 4.6). This allows us to investigate the influence of static correlations between symbol series on obtained estimates of the symbolic transfer entropy TE by comparing it with estimates of the corrected symbolic transfer entropy TE_c and minimize the possible influences of a finite sample effect on numerical estimates of both measures. To choose the appropriate delay time τ we followed [SL08] and set the embedding window such as to approximately cover a basic period T of investigated dynamical systems.,

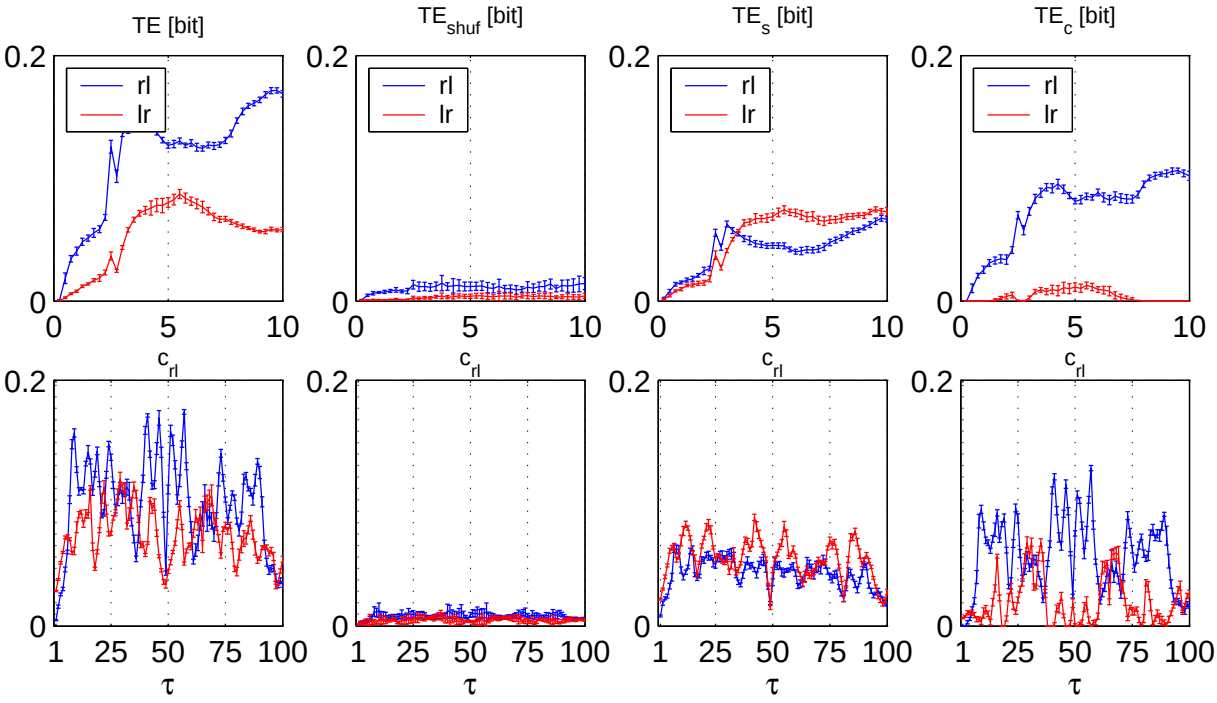


Figure 4.8.: Upper row: Symbolic transfer entropies TE and corrected symbolic transfer entropies TE_c between $x-$ components of unidirectionally coupled Rössler (driver) and Lorenz (responder) oscillators for increasing coupling strength c_{rl} and with fixed embedding parameters: $m = 3$, $\tau = 10$. Lower row: The same but now for the fixed coupling strength $c_{12} = 5$ and with varying embedding parameters: $m = 3$, $\tau \in [1, 100]$. TE_{shuf} and TE_s represent symbolic transfer entropies between surrogate symbol series obtained by either a random shuffling or time-inversion transformation of data. Each point on all plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

i.e., $(m-1)\tau \approx T$. Thus, for $m = 3$ we set $\tau \approx T/2$. With the used sampling rate $d\tilde{t} = 0.03$ the basic period of the Lorenz oscillator can be estimated as $T_{Loz} \approx 20$ integration steps (or, in absolute units as $T_{Loz}d\tilde{t} = 0.6$, see Fig 4.1). Since the Hénon map time series has a broad band spectrum and does not possess a clearly defined periodic structure (Fig 4.1) the choice of delay time τ should not have a real influence on the obtained values of the symbolic transfer entropy. Figures 4.5 and 4.6 show the dependencies of symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ between two unidirectionally coupled ($c_{12} = 0.5$) structurally identical and structurally non-identical Hénon maps on the delay time τ. These figures show that the inference of the directionality of an interaction can be successfully performed (i.e., $TE(S_1, S_2) > TE(S_2, S_1)$) for all τ and, moreover, obtained values of the symbolic transfer entropies are independent on the delay time for $\tau \geq 10$. For small delay times ($\tau < 10$) the components of the embedding vector are not independent from each other[9] and this leads to

[9]This can be determined my measuring the autocorrelation function of the Hénon map time series which

additional correlations between permutation symbols and thus to higher values of symbolic transfer entropies. The analysis of entropy transfer between two unidirectionally coupled ($c_{12} = 5$) Lorenz oscillators revealed a more complex dependency of the symbolic transfer entropy on the delay time τ. As can be seen from Fig. 4.7 the inference of the directionality of an interaction can be successfully performed for $\tau = T_{Lor}/2 = 10$. However, for some values of τ (e.g. $\tau \approx 20$) – in contrast to the case of interacting Hénon maps – the evaluation of the symbolic transfer entropies between Lorenz oscillators led to the incorrect inference of directionality when $TE(S_1, S_2) < TE(S_2, S_1)$. The analysis of entropy transfer between a system of unidirectionally coupled ($c_{rl} = 5$) Rössler (driver) and Lorenz (responder) oscillators revealed a more stronger dependency of the symbolic transfer entropies on the delay time τ (Fig. 4.8). We observed that for all delay times $\tau \approx nT_{Roes}$ ($n = 1, 2, 3, \ldots$) the evaluation of entropy transfer with the symbolic transfer entropy led to an incorrect inference of directionality of an interaction. Here, $T_{Roes} \approx 32$ denotes a basic period of the Rössler oscillator. Thus, obtained values of the transfer entropy between oscillating dynamical systems such as chaotic oscillators revealed a strong dependence of the symbolic transfer entropies on delay time τ. Similar to [SL08] we also found that the optimal choice of τ is achieved by using the embedding window $(m-1)\tau$ which approximately covers a basic period T of investigated dynamical systems. In case when investigated systems exhibited two different basic periods (in our case coupled Lorenz and Rössler oscillators) we observed that each of either periods can be used to estimate the optimal time delay, i.e., $\tau = T_{Lor}/2 = 10$ or $\tau = T_{Roes}/2 = 16$ (see Fig. 4.8). Based on obtained findings, the delay time for all investigated dynamical model dynamical systems, in this section, was set to $\tau = 10$.

Figure 4.5 shows the dependencies of symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ between two unidirectionally coupled structurally identical Hénon maps as the coupling strength c_{12} is increased. Comparing to findings obtained in [SL08] we observed a qualitatively similar behavior of the symbolic transfer entropy $TE(S_1, S_2)$. Its value is growing with the coupling strength c_{12}, reaching its maximum ($TE(S_1, S_2) \approx 0.11$) at $c_{12} \approx 0.6$ and then approaching zero for $c_{12} > 0.75$ where the Hénon maps get fully synchronized [QAG00]. This indicates the directional interaction from the first to the second Hénon map as it is expected from the used coupling scheme (Eqs. A.6 and A.7). Analysis of dependencies of the symbolic transfer entropies between unidirectionally coupled structurally non-identical Hénon maps on the coupling strength c_{12} revealed a qualitatively similar behavior (Fig. 4.6). The value of symbolic transfer entropy in the direction "driver-to-responder" $TE(S_1, S_2)$ is also growing with the coupling strength c_{12} reaching its maximum ($TE(S_1, S_2) \approx 0.2$) at $c_{12} \approx 0.7$. However, in comparison to the case of structurally identical maps, now we observed higher values of $TE(S_1, S_2)$ in the same range of coupling strengths. This observation might indicate that the entropy transfer from a driver to a responder is stronger for structurally non-identical than for structurally identical dynamical systems. Similar findings were also obtained in [QAG00] by using the state space-based approach to characterize interactions between dynamical systems. Due to the fact that with this choice of control parameters ($b_1 = 0.3$, $b_2 = 0.1$) the complete synchronization of Hénon maps cannot be

vanishes as $\tau \geq 10$ (Fig. 4.1).

achieved (even for higher values of coupling strengths; see [QAG00]) we observed that the symbolic transfer entropy $TE(S_1, S_2)$ remains positive for all analyzed coupling strengths $c_{12} \in [0, 1]$.

It is important to point out that the symbolic transfer entropy $TE(S_2, S_1)$ has also positive values despite the absence of coupling in the direction "responder-to-driver" (Fig. 4.5). This effect appears to be more pronounced in case of structurally different Hénon maps (Fig. 4.6). To test whether the observed positive values of $TE(S_2, S_1)$ are a result of systematic errors due to the finite length of symbol series we additionally estimated $TE_{shuf}(S_1, S_2)$ and $TE_{shuf}(S_2, S_1)$ which were calculated for surrogate series obtained by a random shuffling of either the series S_1 or S_2. At a zero coupling strength ($c_{12} = 0$) the obtained values of $TE_{shuf}(S_2, S_1)$ and $TE(S_2, S_1)$ are almost equal to each other indicating that the positive values of the symbolic transfer entropy in the direction "responder-to-driver" can be explained as a result of the finite length of the symbol series S_1 and S_2 for both structurally identical and non-identical Hénon maps. With the positive coupling strengths ($c_{12} > 0$) the obtained values of $TE(S_2, S_1)$ exceed $TE_{shuf}(S_2, S_1)$ indicating that in this case the observed entropy transfer in the direction "responder-to-driver" cannot be explained as a result of finite sample effects alone. To test whether obtained values of $TE(S_2, S_1)$ are a result of high-order static correlations between S_1 and S_2 we also estimated $TE_s(S_1, S_2)$ (and $TE_s(S_2, S_1)$) which were calculated between surrogate series obtained by a time-reversion transformation of the original symbol series S_1, S_2 (see Fig. 4.4). The observed dependence of $TE_s(S_2, S_1)$ on the coupling strength c_{12} shows a qualitatively similar behavior as $TE(S_2, S_1)$ for both structurally identical and non-identical Hénon maps. This allows us to conclude that an observed "spurious" entropy transfer in the direction "responder-to-driver" (i.e.,$TE(S_2, S_1) > 0$) might be a result of high-order static correlations between S_1 and S_2 (see Fig. 4.4). The positive value of $TE_s(S_1, S_2) > 0$ indicates that a part of the entropy transfer in this direction might also be a result of high order static correlations between S_1 and S_2. As a final step, by using relation 4.9 we computed the corrected symbolic transfer entropies $TE_c(S_1, S_2)$ and $TE_c(S_2, S_1)$ in dependence on the coupling strength c_{12}. The observed values of the corrected symbolic transfer entropy indicate that in the case of two structurally identical as well as structurally non-identical unidirectionally coupled Hénon maps it is, in general, possible to minimize the influence of static correlations in data and thus to correct the observed positive values of the symbolic transfer entropy in the direction "responder-to-driver".

We continue our analysis of directional interactions in model dynamical systems by measuring the symbolic transfer entropy and corrected symbolic transfer entropy between two unidirectionally coupled chaotic oscillators. The analysis of symbolic transfer entropy between a pair of identical Lorenz oscillators in the direction "driver-to-responder" $TE(S_1, S_2)$ showed a similar behavior as in the above considered case of interacting Hénon maps (see Fig. 4.7). The value of $TE(S_1, S_2)$ was growing with the coupling strength c_{12} reaching its maximum ($TE(S_1, S_2) \approx 0.04$) at $c_{12} \approx 6$ and then approaching zero for $c_{12} > 8$ where Lorenz oscillators reached a regime of complete synchronization [PRK01]. This indicates the directional interactions from the first to the second Lorenz oscillator as it is expected from the used coupling scheme (Eq. A.8). We observed that the values of the symbolic transfer entropies between Lorenz oscillators have been reduced in comparison to the values

obtained for interacting Hénon maps. This can be explained if we recall the fact that the Lorenz oscillator has a smaller value of the entropy rate comparing to the entropy rate of the Hénon map[10]. The analysis of symbolic transfer entropies between unidirectionally coupled Rössler (driver) and Lorenz (responder) oscillators showed that the directionality of an interaction can also be successfully identified, i.e., $TE(S_r, S_l) > TE(S_l, S_r)$, for all analyzed values of the coupling strength c_{rl} (Fig. 4.8). Due to the fact that with the used values of the system's parameters the complete synchronization between Rössler and Lorenz oscillators cannot be achieved (see [QAG00]) we observed that the symbolic transfer entropy $TE(S_1, S_2)$ remains positive for all analyzed coupling strengths $c_{rl} \in [0, 10]$.
The analysis of the directionality of interactions between Lorenz oscillators in the direction "responder-to-driver" showed an existence of entropy transfer ($TE(S_2, S_1) > 0$) despite the absence of coupling in this direction. The same observation was also obtained by analyzing the symbolic transfer entropy from Lorenz to Rössler oscillators $TE(S_l, S_r)$. To test whether these observations can be explained as a result of systematic errors due to the finite length of symbol series we again estimated the symbolic transfer entropy TE_{shuf} between surrogate series which were obtained by a random shuffling of the original symbol series. Similarly to the observations obtained above (see the case of interacting Hénon maps) it was found that the absolute values of the symbolic transfer entropies $TE(S_2, S_1)$ ($TE(S_l, S_r)$) exceeded the values of $TE_{shuf}(S_2, S_1)$ ($TE_{shuf}(S_l, S_r)$) for all coupling strengths c_{21} (c_{rl}). This indicates that the observed entropy transfer in the direction "responder-to-driver" cannot be explained as a result of finite sample effects alone. To test whether these observations are a result of high-order static correlations between S_1 and S_2 (or S_l and S_r) we estimated the symbolic transfer entropy $TE_s(S_2, S_1)$ ($TE_s(S_l, S_r)$) between surrogate series obtained by a time-reversion transformation (see Fig. 4.4) of the original symbol series. This step of analysis revealed a qualitatively similar but not exact relationship between $TE_s(S_2, S_1)$ ($TE_s(S_l, S_r)$) and $TE(S_2, S_1)$ ($TE(S_l, S_r)$). Finally, by using relation 4.9 we again computed the corrected symbolic transfer entropies TE_c as a function of the coupling strength c_{12} (c_{rl}). The observed values of the corrected symbolic transfer entropy indicate that it is possible to correct the observed "spurious" entropy transfer in the direction "responder-to-driver". However, in contrast to the case of time discrete dynamical systems (interacting Hénon maps considered above) where the positive bias of entropy transfer in the direction "responder-to-driver" was efficiently corrected the analysis of continuous time dynamical systems revealed that the corrected symbolic transfer entropy $TE_c(S_l, S_r)$ can remain positive for some values of the coupling strength c_{rl} (Fig. 4.8).

4.2.3. Entropy transfer between noise-contaminated time series

In this section, we analyze the influence of noise in time series on numerical estimates of the symbolic transfer entropy TE as well as of the corrected symbolic transfer entropy TE_c. To test the robustness of TE and TE_c against the amount of noise in time series we analyzed a pair of time series generated by two unidirectionally coupled Hénon maps and two unidirectionally coupled Lorenz oscillators with increasing coupling strengths and a

[10]The values of entropy rates for a Hénon map and Lorenz oscillator can be estimated by using Pesin's identity (Eq. 2.21) and values of corresponding Lyapunov exponents (see Table 2.1)

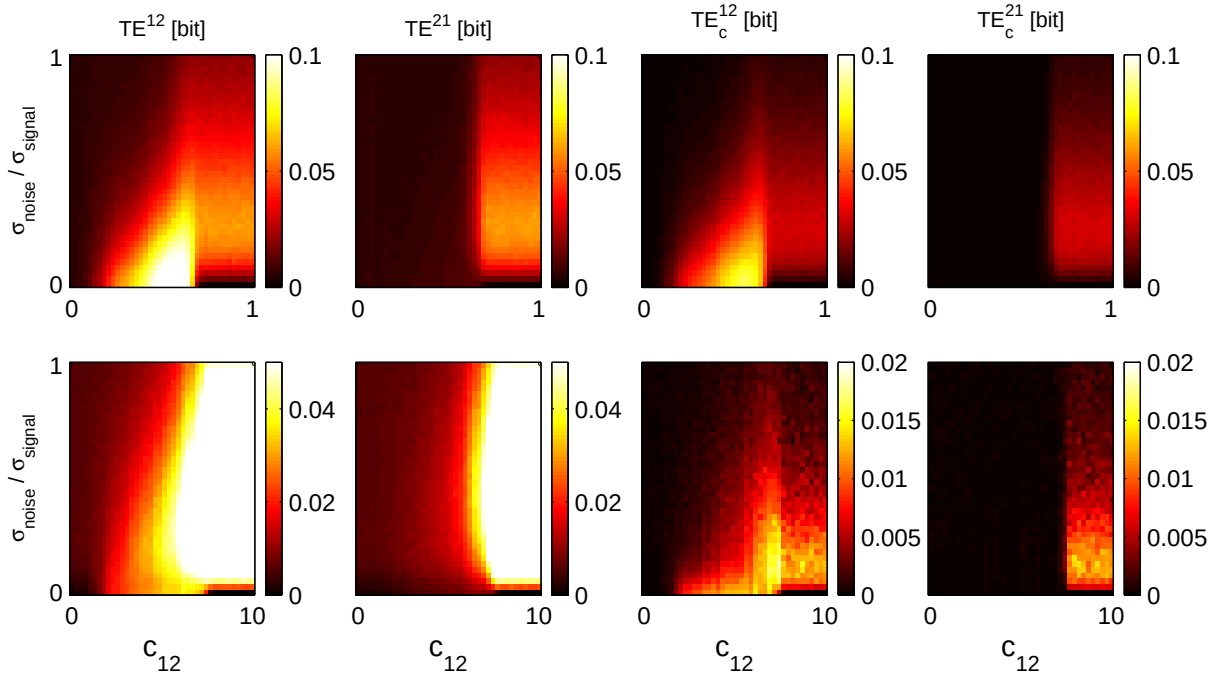


Figure 4.9.: The dependencies of the symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ as well as the corrected symbolic transfer entropies $TE_c(S_1, S_2)$ and $TE_c(S_2, S_1)$ between two Hénon maps (upper row) or two Lorenz oscillators (lower row) on the coupling strength c_{12} for different noise-to-signal ratios. Each point on all figures corresponds to the mean value of the transfer entropy taken over 20 realizations.

different amount of additive noise. This allows us to compare the influence of the amount of noise in time series to the abilities of these measures to infer the directionality of interaction between dynamical systems and also allows us to demonstrate some limitations and pitfalls which can arise in real-world situations.

Let us consider two unidirectionally coupled, structurally identical Hénon maps (Eqs. A.6 and A.7; $b_1 = 0.3$ $b_2 = 0.3$) and Lorenz oscillators (Eq. A.8; $R_{1,2} \in \mathbf{N}(28, 0.5)$) with coupling strengths $c_{12} \in [0, 1]$ (for Hénon maps) and $c_{12} \in [0, 10]$ (for Lorenz oscillators). In both cases the first system (driver) is coupled into the second system (responder) and it is expected to find a positive value of entropy transfer in the direction "driver-to-responder" and zero entropy transfer in the opposite direction. In the same way as we did in the previous section the differential equations defining two interacting Lorenz oscillators (Eq. A.8) were integrated by using a fourth order Runge-Kutta algorithm with integration step $dt = 0.005$ and then downsampled to $d\tilde{t} = 0.03$. We generated 20 realizations of Hénon (Lorenz) time series of length $N = 10^4$ for each coupling strength c_{12}. The initial conditions were normally distributed with zero mean and unit variance. In order to eliminate transients, the first 10^4 iterations were discarded. Next, the $x-$ components of each Hénon map (Lorenz oscillator) were superimposed with additive white noise using different noise-to-signal ratios and transformed into a pair of series of permutation symbols S_1 and S_2 ($m = 3$, $\tau = 10$).

For each coupling strength we increased the standard deviation of the additive white noise $\sigma_{noise} = 0, \ldots, \sigma_{signal}$ and thus increased the noise-to-signal ratio from 0 to 1.
Figure 4.9 shows the dependencies of the symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ along with the corrected symbolic transfer entropies $TE_c(S_1, S_2)$ and $TE_c(S_2, S_1)$ on the coupling strength c_{12} and on the noise-to-signal ratio. In the regime of weak and intermediate synchronization ($c_{12} < 0.75$ for Hénon maps and $c_{12} < 7$ for Lorenz oscillators) the absolute values of the symbolic transfer entropy as well as of the corrected symbolic transfer entropy gradually declined when the amount of noise in the time series was increased. The qualitative comparison of these dependencies for both measures revealed that the corrected symbolic transfer entropy declined faster for increasing noise-to-signal ratio than the symbolic transfer entropy which appeared to be more robust against noise in time series. As it was shown in the previous section (see Fig. 4.5 and Fig. 4.7), in the regime of complete synchronization ($c_{12} > 0.75$ for Hénon maps and $c_{12} > 7$ for Lorenz oscillators) there is no entropy transfer in both directions, i.e., $TE(S_1, S_2) = TE(S_2, S_1) \approx 0$. However, already with a small amount of noise the symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ became positive, indicating an entropy transfer even in the regime of complete synchronization. This spurious (noise-induced) entropy transfer can be reduced by estimating it with the corrected symbolic transfer entropy. It is important to point out that in the regime of complete synchronization of Hénon maps ($c_{12} > 0.75$) and Lorenz oscillators ($c_{12} > 7$) the influence of noise on estimates of the symbolic transfer entropy was symmetric, i.e., $TE(S_1, S_2) \approx TE(S_2, S_1) > 0$. In this case – despite the fact that already a small amount of noise results in positive values of the symbolic transfer entropy – the influence of noise did not led to an incorrect inference of the directionality of interactions between dynamical systems.
To investigate the influence of the asymmetry in the amount of noise between two signals on the symbolic transfer entropy along with the corrected symbolic transfer entropy we again measured entropy transfer between two unidirectionally coupled Hénon maps (Lorenz oscillators) as we did above. But now for a fixed coupling strength ($c_{12} = 0.5$ for Hénon maps and to $c_{12} = 5$ for Lorenz oscillators) and gradually increased noise-to-signal ratios ($\sigma^{(j)}_{noise}/\sigma^{(j)}_{signal}$) which were set for driver ($j = 1$) and responder ($j = 2$) time series separately. Figure 4.10 shows the symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ along with the corrected symbolic transfer entropies $TE_c(S_1, S_2)$ and $TE_c(S_2, S_1)$ for different noise-to-signal ratios. Our findings showed that, as expected, both measures indicated an entropy transfer in the direction "driver-to-responder" ($TE(S_1, S_2) > 0$ and $TE_c(S_1, S_2) > 0$) and an almost zero entropy transfer in the opposite direction ($TE(S_1, S_2) \approx 0$ and $TE_c(S_1, S_2) \approx 0$). However, the values of $TE(S_1, S_2)$ and $TE_c(S_1, S_2)$ gradually declined when the amount of noise in the time series was increased. In contrast to the case considered above, where the entropy transfer was measured between time series superimposed with an equal amount of noise, now we observed that both measures of entropy transfer declined more faster when the noise was added to the driver (first system) rather than to the responder (second system) time series (see Fig. 4.10).
Next, we studied the influence of noise on entropy transfer between two interacting Lorenz oscillators. As one can see in Fig. 4.10 the symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ showed more complex dependencies on the amount of noise added to either

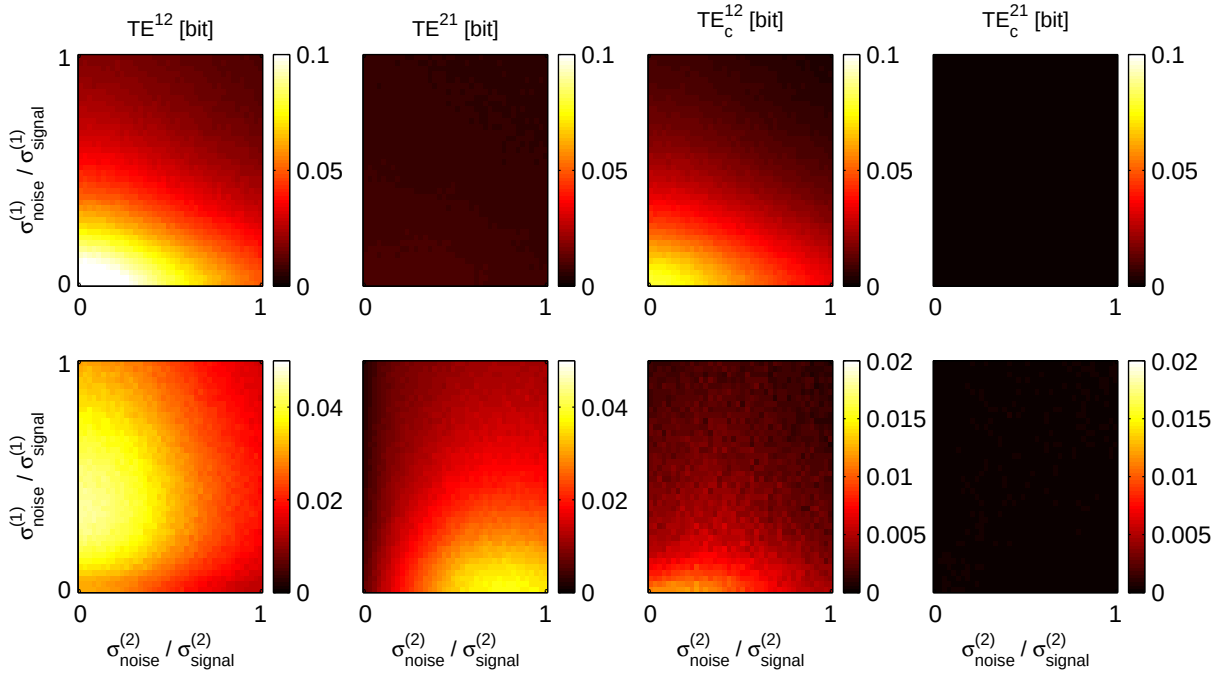


Figure 4.10.: The dependencies of the symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ as well as the corrected symbolic transfer entropies $TE_c(S_1, S_2)$ and $TE_c(S_2, S_1)$ between two Hénon maps (upper row) or two Lorenz oscillators (lower row) on different noise-to-signal ratios and fixed coupling strengths $c_{12} = 0.5$ (Hénon) and $c_{12} = 5$ (Lorenz). The $x-$ component of the first (second) system was superimposed with Gaussian white noise with increasing standard deviations $\sigma^{(1)}_{noise}$ ($\sigma^{(2)}_{noise}$). Each point on all figures corresponds to the mean value of the transfer entropy taken over 20 realizations.

driver or responder time series than in case of two interacting Hénon maps. By increasing the amount of noise added to the responder (second system) we observed a gradual decline of the symbolic transfer entropy in the direction "driver-to-responder" $TE(S_1, S_2)$. In contrast to the analysis of Hénon time series, here, we observed a noise-induced entropy transfer in the direction "responder-to-driver" $TE(S_2, S_1)$. This spurious entropy transfer exhibited a resonance-like dependence[11] on the amount of noise added to the responder and was characterized by an increase and a following gradual decline of the symbolic transfer entropy $TE(S_2, S_1)$ when the amount of noise in the responder was increased. By increasing the amount of noise added to the driver (first system) we observed an almost opposite behavior, i.e., a decline of the symbolic transfer entropy in the direction "responder-to-driver" $TE(S_2, S_1)$ and a resonance-like behavior of the symbolic transfer entropy in the direction "driver-to-responder" $TE(S_1, S_2)$. Summarizing obtained findings we can conclude that in case of two unidirectionally coupled Lorenz oscillators a less noisy oscillator

[11] A similar resonance-like phenomenon was already observed in [SL08] for the symbolic transfer entropy between interacting Lorenz oscillators.

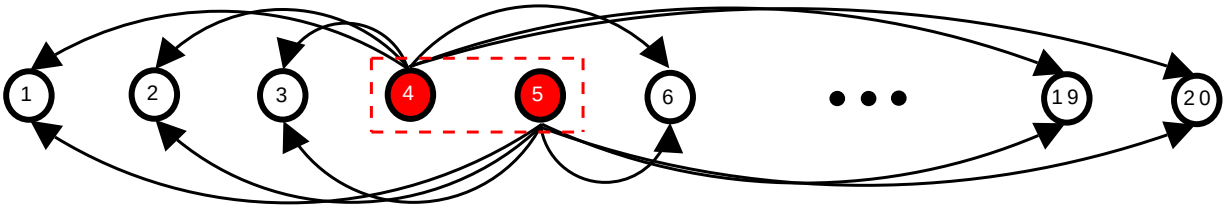


Figure 4.11.: Coupling scheme of 20 interacting Lorenz oscillators: responder (white circles) and drivers (red circles). Black arrows represent unidirectional diffusive coupling strength c_d.

appears to be driven by a more noisy one. Moreover, it was also found that when e.g. $\sigma^{(1)}_{noise}/\sigma^{(1)}_{signal} < 0.5$ and $\sigma^{(2)}_{noise}/\sigma^{(2)}_{signal} \approx 1$ (see Fig 4.10) the symbolic transfer entropy in the direction "responder-to-driver" $TE(S_2, S_1)$ exceeded the symbolic transfer entropy in the direction "driver-to-responder" $TE(S_1, S_2)$ indicating an incorrect directionality of interactions. By repeating the analysis of entropy transfer between Lorenz oscillators with the corrected transfer entropy TE_c it was found that the spurious entropy transfer in the direction "responder-to-driver" was almost zero ($TE_c(S_2, S_1) \approx 0$) and thus the effect of noise-induced asymmetry of entropy transfer was now not observed. However, our findings again indicated that the corrected symbolic transfer entropy TE_c appeared to be less robust against noise than the symbolic transfer entropy TE.

Analysis of entropy transfer between dynamical model systems, made in this section, revealed a destructive influence of noise on the estimates of the symbolic transfer entropy TE as well as corrected symbolic transfer entropy TE_c. It was shown that by adding an equal amount of noise to the driver and responder time series of two interacting Hénon maps (Lorenz oscillators) resulted in a gradual decline of absolute values of both measures. We found that the inference of the directionality of interactions between Hénon maps (Lorenz oscillators) was still possible for noise contaminated time series with $\sigma_{noise}/\sigma_{signal} \approx 0.5$ ($\sigma_{noise}/\sigma_{signal} \approx 1$), as one can see in Fig 4.9. However, these findings also indicated that the range of coupling strengths c_{12}, where the inference of directionality was still possible, shrinked when increasing the noise-to-signal ratio. A more complex influence of noise on the estimate of the symbolic transfer entropy was observed when the amount of noise added to either driver or responder was different. It was observed that a less noisy system appeared to be driven by a more noisy one. Interestingly, such a noise-induced asymmetry of entropy transfer was not observed by estimating the symbolic transfer entropy between Hénon maps (see Fig 4.10). Analysis of entropy transfer by using the corrected symbolic transfer entropy TE_c indicated that this measure is less sensitive to the difference in the amount of noise between two time series than the symbolic transfer entropy TE. In next section we will demonstrate that this property of TE_c can be very useful for the estimation of entropy transfer in multivariate data exhibiting a different amount of noise per time series.

4.3. Directional interactions in multivariate time series

Identifying directionality of interactions in multivariate data is a widely studied area of research [BKK04, MBG$^+$05, SWD$^+$06, HSPVB07, WST07, FP07, MBRS08, RMS08, OMWL08, SL09]. The question whether bivariate analysis techniques suit for the analysis of directional interactions in multivariate data was recently addressed in [OMWL08] and under the information-theoretic framework in [FP07, SL09]. It was demonstrated that in multivariate data a high degree of synchronization between signals can considerably limit a reliable detection of directional interactions [OMWL08]. In [OMWL08] authors applied a phase modeling approach to measure strength and direction of interactions within a cluster of interacting chaotic oscillators. It was demonstrated that depending on the degree of local bidirectional coupling between oscillators some subclusters of oscillators can spuriously appear to be driven by others and this has to be taken into account when analyzing field data with unknown dynamics.

It is known that the degree of synchronization between several (even uncoupled) chaotic oscillators can be increased by a common force [PRK01]. This phenomenon needs to be taken into account when the data is generated by, for example, a network of uncoupled oscillators driven by a common force or by other oscillators (drivers). In this case the driven oscillators (responders) can synchronize with each other even without being coupled. With a high degree of synchronization between a pair of responders the symbol series obtained from these oscillators may exhibit high-order static correlations which can lead to a positive value of the symbolic transfer entropy between them. Indeed, as we could see in section 4.2.2 – by measuring entropy transfer between a pair of unidirectionally coupled chaotic oscillators – the symbolic transfer entropy TE can attain positive values in the direction "responder-to-driver". These findings indicated that a spurious entropy transfer was a result of high-order static correlations between corresponding permutation symbols. Moreover, this effect was amplified when the oscillators got more synchronized with each other (see Figs. 4.7 and 4.8). It was also shown that estimating the entropy transfer with corrected symbolic transfer entropy TE_c allowed us to reduce the influence of high-order static correlations and thus to reduce spurious entropy transfer in the direction "responder-to-driver".

In this section we follow the information-theoretic approach to measure directional interactions in multivariate data as used in [SL09]. In order to generate multivariate data we used a network of 20 coupled Lorenz oscillators (Eq. A.11). Directional interactions in data are then investigated by measuring entropy transfer with the symbolic transfer entropy TE as well as corrected symbolic transfer entropy TE_c (see section 4.2.2). With the used coupling scheme, two drivers (fourth and fifth oscillators see Fig. 4.11) were unidirectionally coupled to the other 18 oscillators (responders) with a fixed coupling strength c_d (Eq. A.11). According to the fact that the drivers are uncoupled among each other they cannot synchronize with the responders for any coupling strength c_d. The differential equations (Eq. A.11) for each oscillator were integrated by using a fourth order Runge-Kutta algorithm with an integration step $dt = 0.005$ and then downsampled to $d\tilde{t} = 0.03$. The initial conditions were normally distributed with zero mean and unit variance. In order to eliminate transients, the first 10^4 integration steps were discarded and the next $N = 10^4$ (in order to compute the mean phase coherence R we used $N = 8192$) integration steps were used

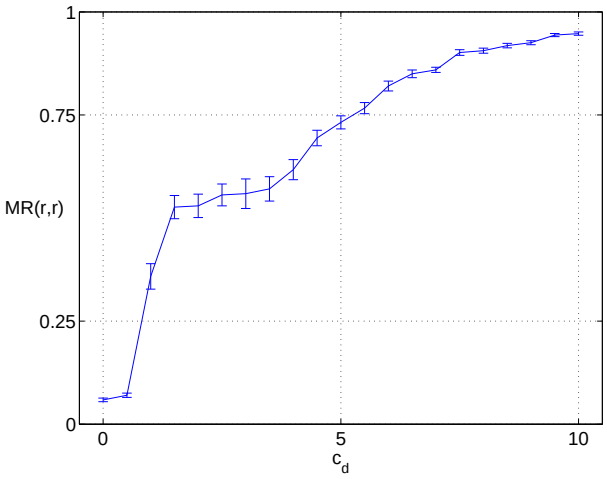


Figure 4.12.: Global synchronization $MR(r,r)$ as a function of coupling strength c_d (Fig. 4.11). Error bars denote standard deviations.

for further analysis. The $x-$ component of each oscillator i was transformed into a series of permutation symbols S_i $(i = 1, \ldots, 20)$. Following considerations made in sections 4.2.2 we used the following embedding parameters: $m = 3$, $\tau = 10$.
In order to investigate the influence of a common driving and to estimate the degree of synchronization between oscillators we computed the mean phase coherence R (Eq. A.14 in section A.2). For each pair of oscillators $(i,j) \in [1, \ldots, 20]$ and for each coupling strength c_d we generated $L = 50$ realizations of Lorenz time series and estimated $R^l_{i,j}$ $(l = 1, \ldots, L)$. Next, we computed the mean value of the mean phase coherence $R_{i,j} = (1/L) \sum_{l=1}^{L} R^l_{i,j}$ and finally computed the *global synchronization* $MR(r,r)$ between responders

$$MR(r,r) = \frac{1}{K} \sum_{(i,j) \neq \{4,5\}} R_{i,j} \tag{4.10}$$

where K denotes the number of summands. Thus, the global synchronization $MR(r,r)$ is close to zero for unsynchronized responders and approaches one when all responded get fully synchronized. We observed that the global synchronization $MR(r,r)$ was nearly zero for $c_d = 0$ and then was growing with increasing coupling strength approaching one for $c_d = 10$ (Fig. 4.12). This finding is in agreement with the fact that chaotic oscillators can be synchronized by a common external force even without being directly coupled with each other [PRK01].
As a next step of our analysis we set a moderate coupling strength $c_d = 2$ and estimated entropy transfer by using either the symbolic transfer entropy TE or the corrected symbolic transfer entropy TE_c. For each pair of oscillators $(i,j) \in [1, \ldots, 20]$ we again generated 50 realizations of Lorenz time series and for each realization estimated $TE^l(S_i, S_j)$ and $TE^l_c(S_i, S_j)$ $(l = 1, \ldots, 50)$. Next, we computed the mean values of both measures $TE(S_i, S_j) = (1/50) \sum_{l=1}^{50} TE^l(S_i, S_j)$ and $TE_c(S_i, S_j) = (1/50) \sum_{l=1}^{50} TE^l_c(S_i, S_j)$ (see Fig. 4.13). Our findings indicate a positive entropy transfer from drivers to responders ($TE(S_i, S_j) \approx 0.025$ and $TE_c(S_i, S_j) \approx 0.008$ for $i \in [4,5]$ and $j \in [1,2,3,6,\ldots,20]$) and

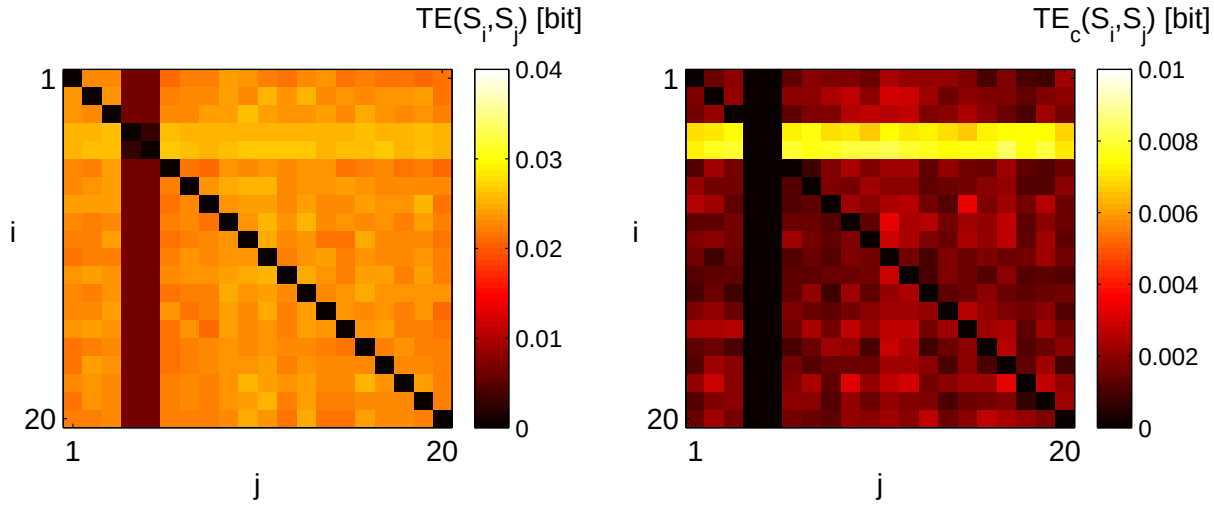


Figure 4.13.: Symbolic transfer entropy TE and corrected symbolic transfer entropy TE_c computed between 20 coupled Lorenz oscillators (see Eq. A.11 and Fig. 4.11). Each entry of both matrices represents entropy transfer from oscillator i to oscillator j for $(i, j) \in [1, \ldots, 20]$. Fourth and fifth oscillators (drivers) are unidirectionally coupled to all other oscillators (responders) with coupling strength $c_d = 2$.

nearly zero entropy transfer in the opposite direction. However, we also found a considerable entropy transfer between responders ($TE(S_i, S_j) \approx 0.023$ for $(i, j) \in [1, 2, 3, 6, \ldots, 20]$). To explain this we need to recall our results obtained in section 4.2.2. As we already observed there, two unidirectionally coupled Lorenz oscillators exhibited non-zero entropy transfer (measured with symbolic transfer entropy) in the direction "responder-to-driver". Moreover, the symbolic transfer entropy in the direction "responder-to-driver" was growing when we increased the coupling strength and oscillators were getting more synchronized (see Figs. 4.7 and 4.8). This was interpreted as a result of high order static correlations between series of permutation symbols. Analysis of entropy transfer between responders by means of the corrected symbolic transfer entropy (see Fig. 4.13) indicated positive but reduced values of transfer entropy in comparison to values obtained with the symbolic transfer entropy ($TE_c(S_i, S_j) \approx 0.002$ for $(i, j) \in [1, 2, 3, 6, \ldots, 20]$). To answer the question whether the corrected transfer entropy indeed allows us to better identify driving structures (i.e., fourth and fifth oscillators; see Fig. 4.11) we performed the following qualitative analysis. We defined a *mean entropy flow* from drivers (d) to responders (r) as

$$MTE(d, r) = \frac{1}{K} \sum_{(j) \neq \{4,5\}} TE(S_4, S_j) + TE(S_5, S_j), \tag{4.11}$$

and a *mean entropy flow* between responders as

$$MTE(r, r) = \frac{1}{K} \sum_{(i,j) \neq \{4,5\}} TE(S_i, S_j) \tag{4.12}$$

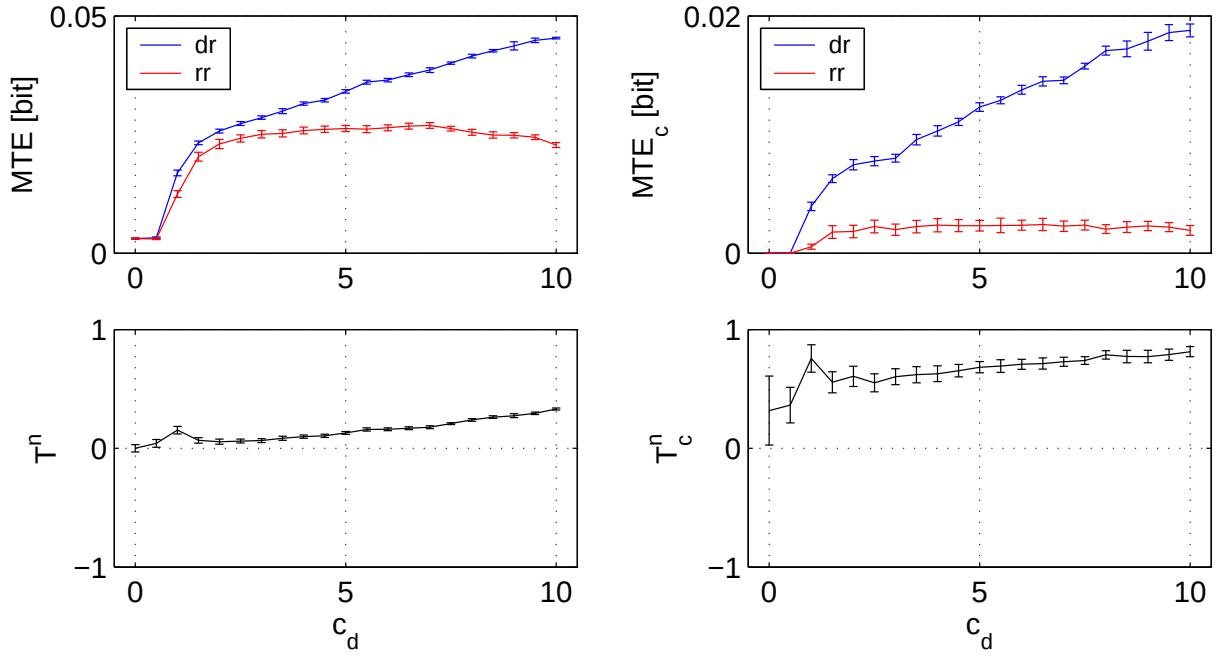


Figure 4.14.: Left: Mean entropy flows $MTE(d,r)$ (from drivers to responders) and $MTE(r,r)$ (between responders) as well as the normalized asymmetry index T^n computed with the symbolic transfer entropy as a function of the coupling strength c_d. Right: Mean entropy flows $MTE_c(d,r)$ (from drivers to responders) and $MTE_c(r,r)$ (between responders) as well as the normalized asymmetry index T_c^n computed with the corrected symbolic transfer entropy as a function of the coupling strength c_d. Error bars denote standard deviation over different pairs of oscillators.

where K denotes the number of summands. Next, we computed the mean entropy flows $MTE(d,r)$ and $MTE(r,r)$ ($MTE_c(d,r)$ and $MTE_c(r,r)$) for $c_d = 0,\ldots,10$ with a step $\delta c_d = 0.5$. Obtained dependency of $MTE(d,r)$ on c_d indicates that the mean entropy flow from drivers to responders was growing with an increasing coupling strength c_d correctly indicating the directionality of interactions in our network (see Fig. 4.14). Analysis of entropy transfer between responders revealed that $MTE(r,r) > 0$ indicating an existence of directional interactions between them. Moreover, the values of $MTE(r,r)$ were also growing with an increasing coupling strength c_d. Analysis of entropy transfer by means of corrected symbolic transfer entropy indicated that the values of the mean entropy transfer between responders $MTE_c(r,r)$ were reduced, i.e., $MTE_c(r,r) < MTE(r,r)$ for all c_d. However, it was also observed that the mean entropy flow from drivers to responders was reduced as well, i.e., $MTE_c(d,r) < MTE(d,r)$ for all c_d. In order to quantitatively investigate whether our correction scheme of the symbolic transfer entropy indeed allows us to enhance the contrast of driving structures we used the *normalized asymmetry index* T^n which was defined as

Noise schemes	Noise-to-signal ratios ($\sigma^i_{noise}/\sigma^i_{signal}$)
Constant noise	$0.05(\delta_{i,16} + \delta_{i,17})$
Noise gradient	$0.025i$

Table 4.1.: Amount of additive Gaussian white noise superimposed an i^{th} Lorenz time series ($i = 1, \ldots, 20$; see Fig. 4.11). Here, $\delta_{i,j}$ denotes Kronecker delta.

$$T^n = (MTE(d,r) - MTE(r,r))/(MTE(d,r) + MTE(r,r)). \tag{4.13}$$

The normalized asymmetry index T^n_c was defined analogously by using the mean entropy flows $MTE_c(d,r)$ and $MTE_c(r,r)$. According to the definition, the positivity of both indices ($T^n > 0$ and $T^n_c > 0$) implies a correct identification of driving structures. In cases when $T^n = 1$ and $T^n_c = 1$ the driving structures in our network exhibit a maximum contrast ($MTE(r,r) = 0$ and $MTE(d,r) > 0$) and when $T^n = 0$ and $T^n_c = 0$ the driving structures cannot be identified ($MTE(d,r) = MTE(r,r)$). Obtained values of T^n for the increasing coupling strength c_d indicated that these values were positive for all $c_d > 0$. This indicates that the driving structures in our network (i.e., fourth and fifth oscillators; see Fig. 4.11) were correctly identified (see Fig. 4.14). By estimating entropy transfer with the corrected symbolic transfer entropy we found that the contrast of the driving structures was indeed improved, i.e., $T^n_c > T^n$ for all $c_d > 0$.

The findings obtained so far demonstrate that a driving force (in our case two drivers) applied to a network of uncoupled Lorenz oscillators (responders) leads to a synchronization of these oscillators and, as a consequence, to a positive entropy transfer between them (see Fig. 4.11). However, despite the fact that this spurious entropy transfer between responders led to the reduction of contrast of driving structures the symbolic transfer entropy nevertheless allowed to identify the drivers for all coupling strengths ($T^n > 0$ for all $c_d > 0$). Moreover, entropy transfer between responders was symmetric and thus did not have influence on inference of the directionality of interactions between oscillators.

Influence of noise on entropy transfer in multivariate data

Extending our findings obtained in section 4.2.3, we consider a network of Lorenz oscillators but now by selectively adding a different amount of noise to all oscillators. As it was shown in section 4.2.3 the asymmetry in the amount of noise between two time series can lead to the spurious asymmetry in entropy transfer (see Fig. 4.10) and thus to the incorrect inference of the directionality of interactions.

First, let us consider two sets of multivariate data generated with the network of coupled Lorenz oscillators by applying two different noise schemes (Table 4.1). In the first noise scheme which we called *constant noise*, we selectively added a moderate amount of additive Gaussian white noise to the time series of 16^{th} and 17^{th} Lorenz oscillators and left the remaining 18 time series noise free (see Fig. 4.11). In the second noise scheme which we called *noise gradient*, we first added a small amount of noise to the time series of the first oscillator ($\sigma^1_{noise}/\sigma^1_{signal} = 0.025$) and then linearly increased the amount of noise added to the time series of all remaining oscillators ($\sigma^i_{noise}/\sigma^i_{signal} = 0.025i$ for $i = 2, \ldots, 20$).

To proceed, let us fix the coupling strength $c_d = 2$ and apply two noise schemes defined above. Analysis of entropy transfer between Lorenz oscillators by adding a moderate amount of noise to the time series of 16^{th} and 17^{th} oscillators (noise scheme *constant noise*) indicated that values of the symbolic transfer entropy from 16^{th} and 17^{th} to all remaining oscillators ($TE(S_i, S_j) \approx 0.04$, $i \in [16, 17]$ and $j \in [1, 2, \ldots, 15, 18, 19, 20]$) exceeded values of the symbolic transfer entropy in the opposite direction ($TE(S_j, S_i) \approx 0.025$) indicating the existence of spurious (noise-induced) driving structures in our network (see Fig. 4.15). This observation is in agreement with our findings obtained in section 4.2.3 where we observed that a less noisy Lorenz oscillator appeared to be driven by a more noisy one. Moreover, with the used coupling strength $c_d = 2$ this noise-induced entropy transfer appeared to be more pronounced than the entropy transfer from drivers to responders ($TE(S_i, S_j) \approx 0.025$ for $i \in [4, 5]$ and $j \in [1, 2, 3, 6, \ldots, 20]$). As a next step of our analysis, we again estimated entropy transfer in our network of Lorenz oscillators, but now by applying a noise scheme *noise gradient* (in this noise scheme, the time series of each Lorenz oscillator i was superimposed with an increasing amount of noise $\sigma^i_{noise}/\sigma^i_{signal} = 0.025i$; $i = 1, \ldots, 20$). Analysis of entropy transfer between Lorenz oscillators with the symbolic transfer entropy indicated that such a gradient of noise-to-signal ratios in our multivariate date set was reflected in obtained entropy transfer between all oscillators (see Fig. 4.15). It was again observed that the less noisy first oscillator oscillator ($i = 1$) was driven by the more noisy oscillators ($i = 2, \ldots, 20$). The same held for the second oscillator ($i = 2$) which appeared to be driven by the more noisy oscillators ($i = 3, \ldots, 20$), etc. Similar with findings obtained by analyzing a system of only two interacting Lorenz oscillators (see section 4.2.3) the analysis of the network of Lorenz oscillators with either noise schemes, made in this section, revealed that the effect of noise-induced entropy transfer was reduced by estimating entropy transfer with the corrected symbolic transfer entropy (see Fig. 4.15). In order to quantitatively investigate the question whether the corrected transfer entropy indeed allows us to enhance the contrast of driving structures and to reduce a destructive influence of noise-induced entropy transfer we again estimated a mean entropy flow from drivers (d) to responders (r) $MTE(d, r)$ ($MTE_c(d, r)$) and between responders $MTE(r, r)$ ($MTE_c(r, r)$) as well as both normalized asymmetry indices T^n and T^n_c introduced above (see Eqs. 4.11, 4.12, and 4.13) for the investigated range of coupling strengths $c_d \in [0, 10]$. Obtained quantitative findings demonstrated that the ability to identify the real driving structures in our network of oscillators (i.e., fourth and fifth oscillators) was indeed impaired when the entropy transfer was estimated by using the symbolic transfer entropy for both noise schemes. We observed that for the noise scheme *constant noise* the normalized asymmetry index was still positive (correctly indicating the driving structures) but not significantly deviating from zero $T^n \approx 0.1 \pm 0.15$ for all coupling strength $c_d > 0$ (see Fig. 4.16). Moreover, by applying the noise scheme *noise gradient* we observed that the identification of driving structures was not anymore possible ($T^n < 0$) for almost all coupling strengths c_d (except $c_d = 1$, see Fig. 4.17). Analysis of the dependencies of a mean entropy flow from drivers (d) to responders (r) $MTE_c(d, r)$ and between responders $MTE_c(r, r)$ as well as the normalized asymmetry indices T^n_c computed using the symbolic transfer entropy revealed that the ability to identify the driving structures for both noise schemes can be recovered for almost all coupling strength $c_d > 0.5$ ($T^n_c > 0$, see Figs. 4.16 and 4.17).

Obtained in this section findings demonstrated that by selectively adding a different amount of noise to time series generated by the network of coupled Lorenz oscillators can lead to the appearance of spurious driving structures such that less noisy oscillators appeared to be driven by a more noisy ones. This led to the situation when the contrast of the real driving structures has been reduced. It was then also demonstrated that the ability of the symbolic transfer entropy TE to identify real driving structures can be indeed recovered by correcting it by Eq. 4.9. Despite the fact that our findings, obtained in this section, appeared to be promising, a further comparison between TE and TE_c is certainly needed as, for example, by considering a more complex networks of coupled oscillators (see [OMWL08, SL09]) or by applying different types of noise (e.g. $1/f$ etc.). Moreover, in order to explain our findings that the noise-induced entropy transfer was only observed in a system of Lorenz oscillators and was not observed for Hénon maps (see Fig. 4.10) require the analysis of different types of the dynamics, or in other words, the analysis of time series exhibiting different entropy rates[12].

[12]Indeed, the fact that the noise-induced entropy transfer was only observed when analyzing Lorenz time series and not observed when analyzing Hénon times series might be a result of different KS-entropies of the Hénon map and Lorenz oscillator. As it was shown in section 3.1.2, KS-entropy of the Hénon map can be estimated as $h_{KS} \approx 0.61$ whereas KS-entropy of the Lorenz oscillator can be estimated in a similar way (see Eq. 2.21 and table 2.1) and with the used integration step $d\tilde{t} = 0.03$ ammounts to $h_{KS} \approx 0.06$.

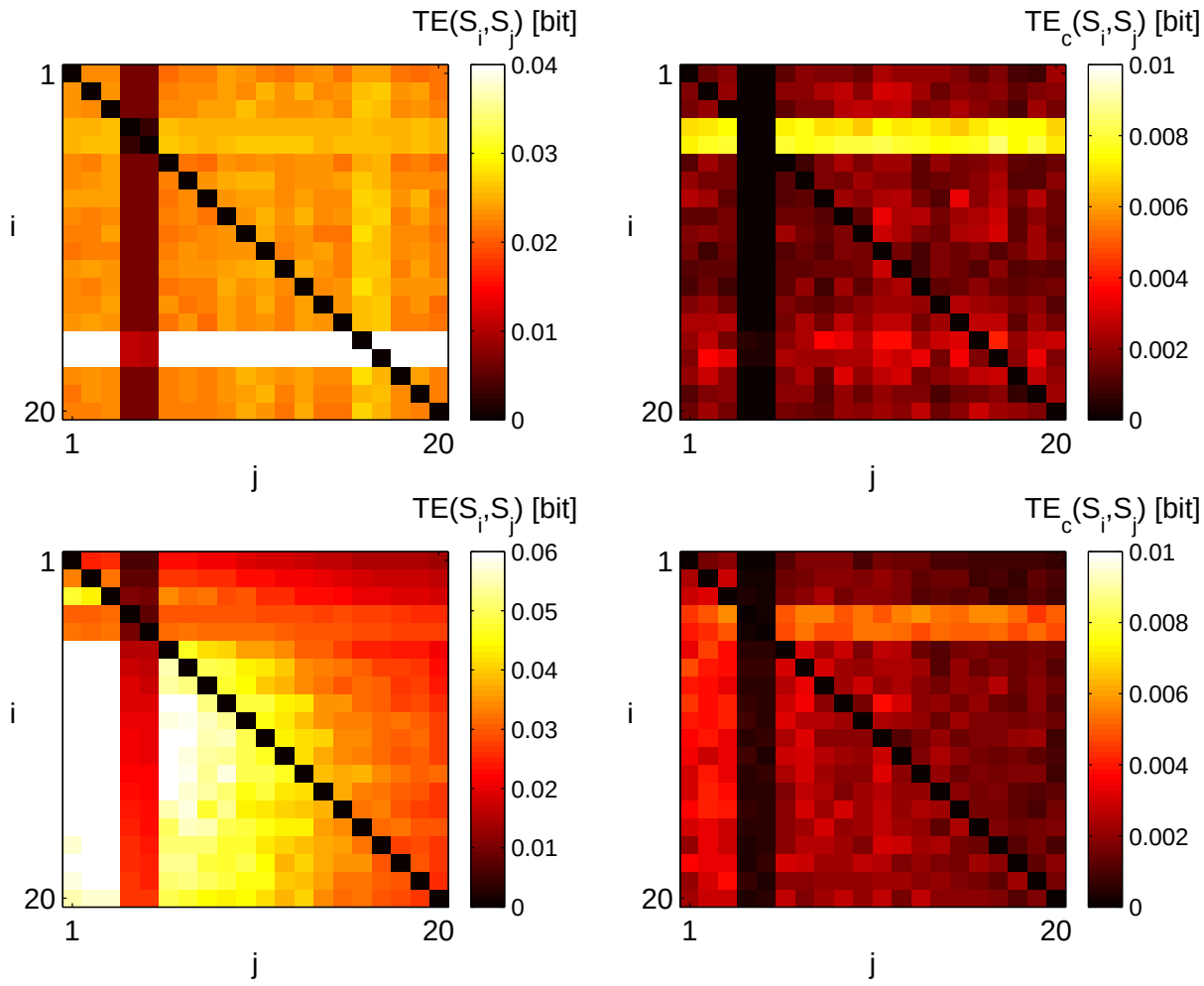


Figure 4.15.: Symbolic transfer entropy TE (left) and corrected symbolic transfer entropy TE_c (right) computed between 20 coupled Lorenz oscillators (see Eq. A.11 and Fig. 4.11). Each entry of all four matrices represents entropy transfer from an oscillator i to oscillator j for $(i, j) \in [1, \ldots, 20]$. Fourth and fifth oscillators (drivers) are unidirectionally coupled to all other oscillators (responders) with the fixed unidirectionally coupling strength $c_d = 2$. Upper row: The 16^{th} and 17^{th} oscillators are superimposed with additive Gausian white noise with noise-to-signal ratio $\sigma_{noise}/\sigma_{sigmal} = 0.05$. Lower row: Each i^{th} oscillator ($i \in [1, \ldots, 20]$) is superimposed with additive Gausian white noise of increasing noise-to-signal ratios $\sigma_{noise}/\sigma_{sigmal} = 0.025i$.

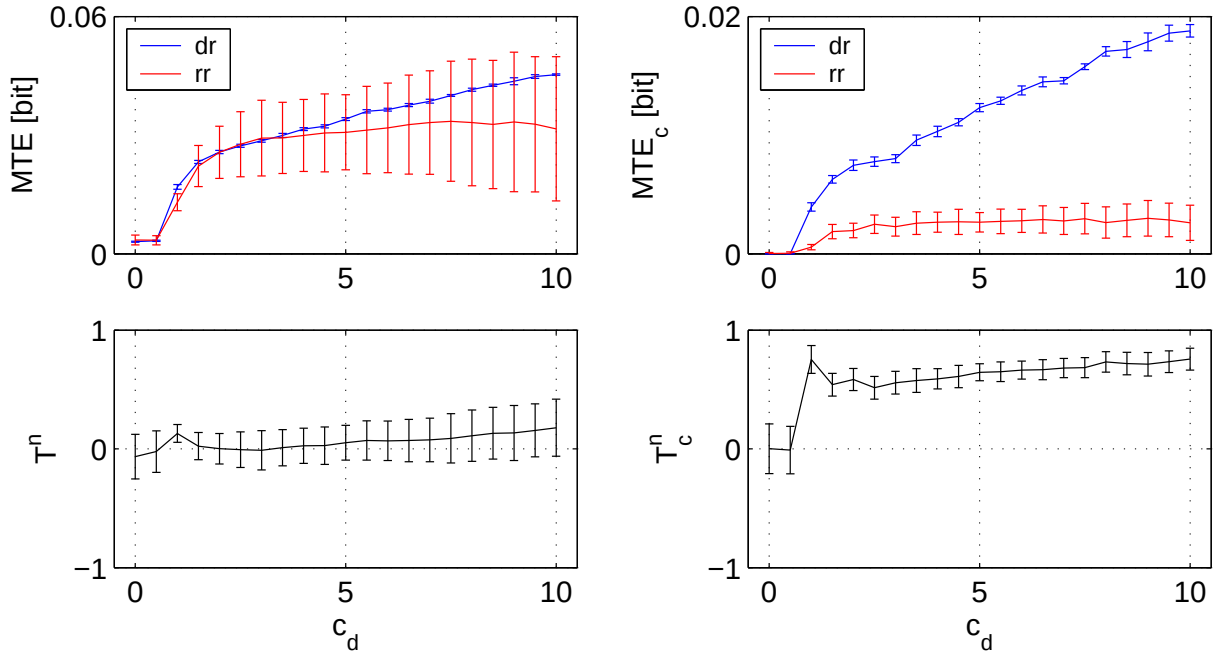


Figure 4.16.: Left: Mean entropy flows $MTE(d,r)$ (from drivers to responders) and $MTE(r,r)$ (between responders) as well as the normalized asymmetry index T^n computed with the symbolic transfer entropy as a function of the coupling strength c_d. Right: Mean entropy flows $MTE_c(d,r)$ (from drivers to responders) and $MTE_c(r,r)$ (between responders) as well as the normalized asymmetry index T_c^n computed with the corrected symbolic transfer entropy as a function of the coupling strength c_d. The 16^{th} and 17^{th} oscillators are superimposed with additive Gausian white noise with noise-to-signal ratio $\sigma_{noise}/\sigma_{sigmal} = 0.05$. Error bars denote standard deviation over pairs of oscillators.

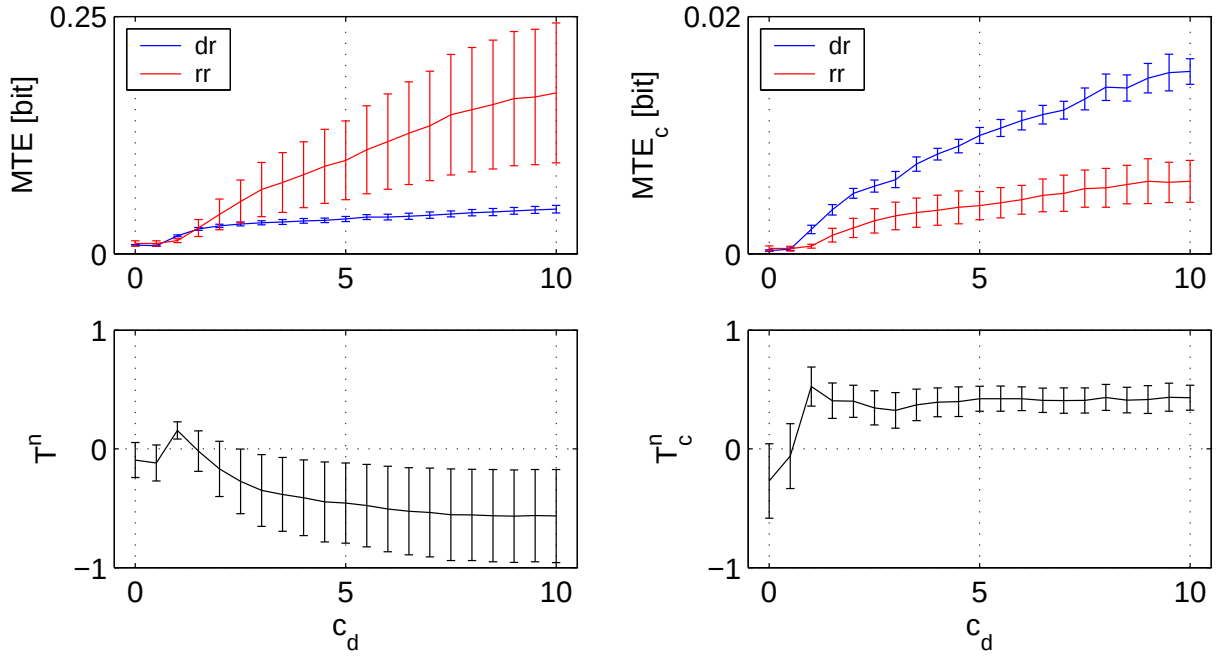


Figure 4.17.: Left: Mean entropy flows $MTE(d,r)$ (from drivers to responders) and $MTE(r,r)$ (between responders) as well as the normalized asymmetry index T^n computed with the symbolic transfer entropy as a function of the coupling strength c_d. Right: Mean entropy flows $MTE_c(d,r)$ (from drivers to responders) and $MTE_c(r,r)$ (between responders) as well as the normalized asymmetry index T^n_c computed with the corrected symbolic transfer entropy as a function of the coupling strength c_d. Each i^{th} oscillator ($i \in [1,\ldots,20]$) is superimposed with additive Gausian white noise of increasing noise-to-signal ratios $\sigma_{noise}/\sigma_{sigmal} = 0.025i$. Error bars denote standard deviation over pairs of oscillators.

5. Characterizing interactions in electroencephalograms of epilepsy patients

In this chapter we will apply the information-theoretic measures (such as mutual information and transfer entropy) to characterize strength and directionality of interactions in the epileptic brain. To do this, we estimate the symbolic mutual information and the symbolic transfer entropy between multi-channel electroencephalographic recordings obtained for diagnostic purposes or for the presurgical evaluation of epilepsy patients. Since it is known that different rhythmic patterns of electrical activity can manifest different physiological and pathophysiological aspects of the human brain [EP97, Buz06], we additionally preprocess our data by filtering it in different physiologically relevant frequency bands. First, by following [Ost08, MOA$^+$08] we investigate the problem of the functional differentiation of brain structures by measuring the strength of interactions between multi-channel electroencephalographic recordings. Next, by estimating the symbolic transfer entropy between these time series we investigate the problem of inferring of the directionality of interactions along the longitudinal axis of the hippocampal formation. Finally, we repeat the analysis of the directionality of interactions in the hippocampus but now by using the corrected symbolic transfer entropy and compare obtained values of both measures of transfer entropy with each other.

5.1. Epilepsy and electrical activity of the epileptic brain

The word *epilepsy* originates from the Greek word *epilepsia* which can be translated as "to seize" or "to attack" [EP97]. Since ancient times epilepsy was considered as a mental disorder related to some malfunction of the human brain. Nowadays, epilepsy is defined as a chronic neurological disorder characterized by a sudden and recurrent malfunction of the brain that is termed epileptic seizure. Epileptic seizures are transient symptoms of an excessive and hypersynchronous activity of neurons in the brain and are divided into two main classes: generalized and partial (focal) seizures. In case of a generalized seizure an excessive hypersynchronous activity of neurons involves the whole brain whereas during a partial seizure this activity originates from a circumscribed brain region that is called *epileptic focus*. During seizures epilepsy patients may suffer a temporal impairment or even loss of consciousness, hallucinations, excessive motor activity and other pathophysiological symptoms depending on the location of the epileptic focus. Partial seizures are further divided – depending on the extent to which the state of consciousness is affected – into simple

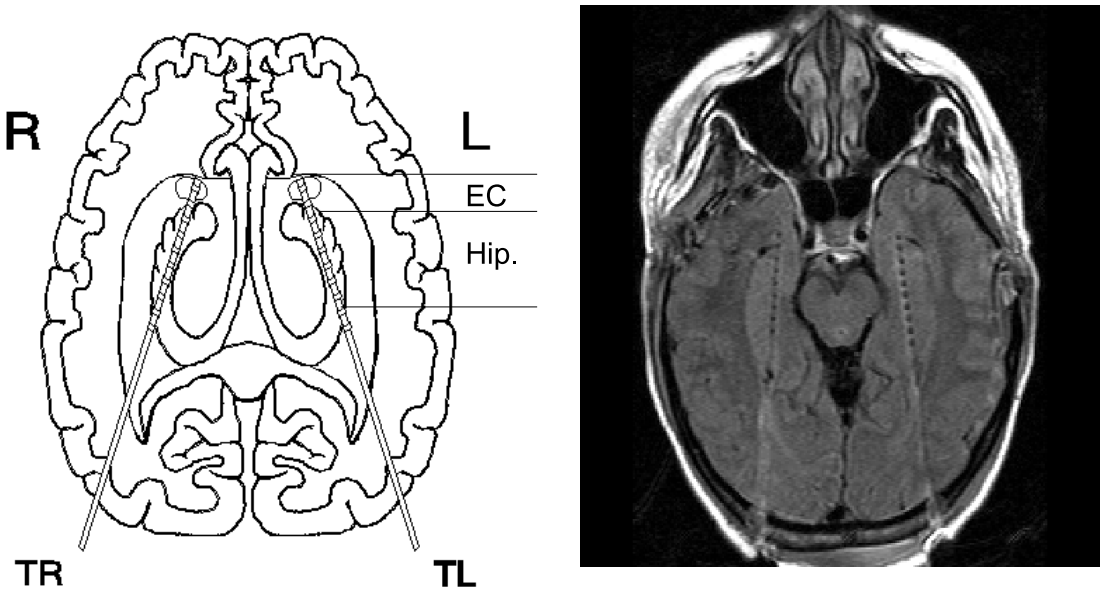


Figure 5.1.: Left: Schematic (axial view) of intracranially implanted depth electrodes: TR and TL. Each electrode contains 10 cylindrical, nickel-chromium-alloy contacts (2.5 mm). Contacts are separated by 4 mm. Anterior (frontal) contacts of both electrodes are located in the entorhinal cortex (EC) whereas posterior (backmost) ones are located in the hippocampus (H) of left (L) or right (R) brain hemispheres. Right: Exemplary axial view of an MRI-scan of an epilepsy patient with implanted depth electrodes.

Frequency bands	raw data	δ	θ	α	β	γ
Frequency range [Hz]	$0.5-85$	$0.5-4$	$4-8$	$8-13$	$13-30$	$30-49$

Table 5.1.: Frequency ranges of physiologically relevant frequency bands.

and complex partial seizures. In case of simple partial seizures the state of consciousness remains unaffected in contrast to complex partial seizures.

Epilepsy affects up to 0.8% of population worldwide [DSSW06]. Approximately every third epilepsy patient cannot be efficiently treated with today's available antiepileptic drugs. Nevertheless, 8% of patients may profit from epilepsy surgery. Successful surgical treatment requires a detailed presurgical evaluation. Localization of the epileptic focus and its delineation from functionally relevant brain areas is one of the main goals of a presurgical evaluation. Detailed neurological and neuropsychological examinations are usually complemented by various noninvasive imaging techniques such as magnetic resonance imaging (MRI) or/and single photon emission computer tomography (SPECT). An exact localization of the epileptic focus requires a direct investigation of electrical activity of impaired brain regions. This is accomplished by invasive recordings of the electrocorticogram (ECoG) and the stereo-electroencephalogram (SEEG) via chronically implanted electrodes. ECoG

Figure 5.2.: Exemplary segment of an SEEG recording from an epilepsy patient. Two depth electrodes (TL and TR) with 10 contacts each are placed within the entorhinal cortex (EC) and the hippocampus (H) of left and right brain hemispheres. The exact locations of the last contacts of both depth electrodes (TL10 and TR10) can not be identified.

represents long-term, multi-channel recordings (using grid of electrodes) of electrical activities of the cortex whereas SEEG is recorded with two depth electrodes implanted in mesial-temporal lobe of the brain. Mesial-temporal lobe epilepsy (MTLE) is a very common epilepsy in adults and frequently with poorly medication-controlled seizures [EP97]. In most cases the epileptic focus is located in the mesial-temporal lobe structures (e.g. hippocampus, amygdala, rhinal cortex). An atrophy (i.e., partial or complete loss of neurons) of the hippocampus is often considered as a main symptom of MTLE. The hippocampus is a neuroanatomical structure in the mesial-temporal lobe of the human brain [KSJ00]. It plays an important role in long-term memory and spatial navigation (see [Eic00] and references therein). A clinical practice indicates that damage of the hippocampus can result in anterograde amnesia, or in other words, in a loss of the ability to create new memories. In humans, this neuroanatomical structure along with the rhinal cortex supports declarative memory formation [FEG$^+$99, FKL$^+$01, MFA$^+$05]. The dominant theory of memory consolidation proposes an active communication between cortex and hippocampus [JW07, WAL$^+$10]. From neuroanatomy it is known that all sensory cortices are connected with the rhinal cortex which is then further connected with the hippocampus [KSJ00]. Thus, the rhinal cortex represents an intermediate structure in the neural pathway connecting cortex and hippocampus. The investigation of electrophysiological correlates from rhinal cortex and hippocampus using the surface electroencephalograms is rather limited. Thus, the analysis of SEEG recordings of epilepsy patients can be very important for the understanding of mechanisms of long-term memory formation.
In this work we will analyze SEEG recordings from a group of 26 epilepsy patients suffering from unilateral MTLE. All patient underwent presurgical evaluation and during this procedure two intracranial depth electrodes had been implanted in the left and right mesial-temporal lobe structures (see Fig. 5.1). These electrodes were positioned along the longitudinal axis of the hippocampal formation such that the anterior contacts of each electrode were placed in the entorhinal cortex and the posterior ones were placed in the hippocampus. The exact location of the electrodes was verified by using a post-operative MRI-scan of the brain. For all patients the presurgical workup indicated an epileptic focus located in either the left (17 patients) or the right (9 patients) brain hemisphere and after surgical resection of the seizure generating structures all patients became seizure-free. The SEEG recordings were sampled at 173.61 Hz using a $12-$bit analog-to-digital converter and band-pass filtered within the frequency band of $0.5 - 85$ Hz. For all patients, the SEEG recordings were performed during seizure-free intervals of approximately two hours duration.

5.2. Characterizing the strength of interactions in SEEG recordings from epilepsy patients

In this section we address the problem of functional differentiation between the entorhinal cortex and hippocampus by analyzing SEEG recordings from epilepsy patients. Analyses made in [Ost08, MOA$^+$08] revealed that SEEG recordings corresponding to intra-regional contacts (i.e., contacts located within the entorhinal cortex or the hippocampus, see Figs.

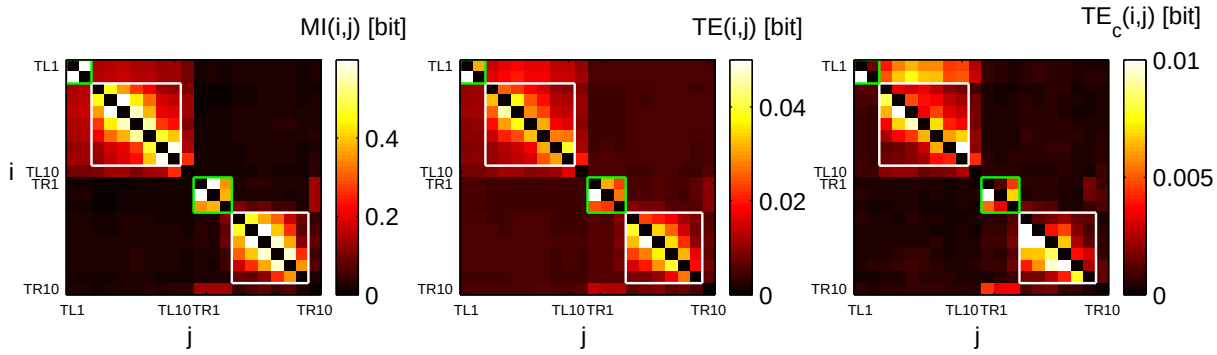


Figure 5.3.: Symbolic mutual information **MI**, symbolic transfer entropy **TE**, and corrected symbolic transfer entropy $\mathbf{TE}_c$ computed between 20 SEEG recordings from an epilepsy patient (see Fig. 5.2). The presurgical evaluation of this patient indicated an epileptic focus located in the left brain hemisphere. Green and white quadrants denote contact-combinations positioned in the entorhinal cortex (TL1-TL2 for left and TR1-TR3 for right hemisphere) and in the hippocampus (TL3-TL9 for left and TR4-TR9 for right hemisphere).

5.1 and 5.2) exhibited a higher degree of synchronization as compared to SEEG recordings corresponding to inter-regional contacts. This reflects the fact that both neuroanatomical structures perform different physiological functions and pairs of intra-regional SEEG recordings are, in general, more correlated (or, in other words, are sharing more common information with each other) than pairs of SEEG recordings obtained from inter-regional contacts.

As a first step of our analysis, we estimated mutual information between pairs of SEEG recordings from epilepsy patients by using the symbolic mutual information MI introduced in section 4.1. As we could see there the symbolic mutual information characterizes the amount of common information shared between two time series and can be used to characterize the strength of interactions between coupled dynamical systems. Following the findings obtained in [Ost08, MOA+08], we now examine the hypothesis that intra-regional SEEG recordings exhibit higher values of mutual information than SEEG recordings corresponding to inter-regional contacts. This hypothesis is motivated by the assumption that two intracranial electroencephalograms recorded from a single neuroanatomical structures (e.g. hippocampus or entorhinal cortex) should share more common information than a pair of recordings corresponding to two different and spatially separated neuroanatomical structures. To perform a moving-window analysis, the continuous SEEG recordings, for each patient, were divided into Λ consecutive windows of approximately 58 seconds ($N = 10^4$ data points) durations[1]. For each window $n = 1, \ldots, \Lambda$ and each pair of contacts $(i, j) = \{\mathrm{TL1}, \ldots, \mathrm{TL10}, \mathrm{TR1}, \ldots, \mathrm{TR10}\}$ (see the implantation scheme in Fig. 5.1) we computed the symbolic mutual information $MI^n(i, j)$ with embedding parameters $m = 3$ and $\tau = 10$ (Eq.4.1). The embedding window $w_{emb} = (m - 1)\tau$ was set to approximately cover a period of oscillatory activity belonging to the $\theta-$ frequency band (see Table 5.1).

[1]In order to select the window length N, we followed the analysis made in section 4.1

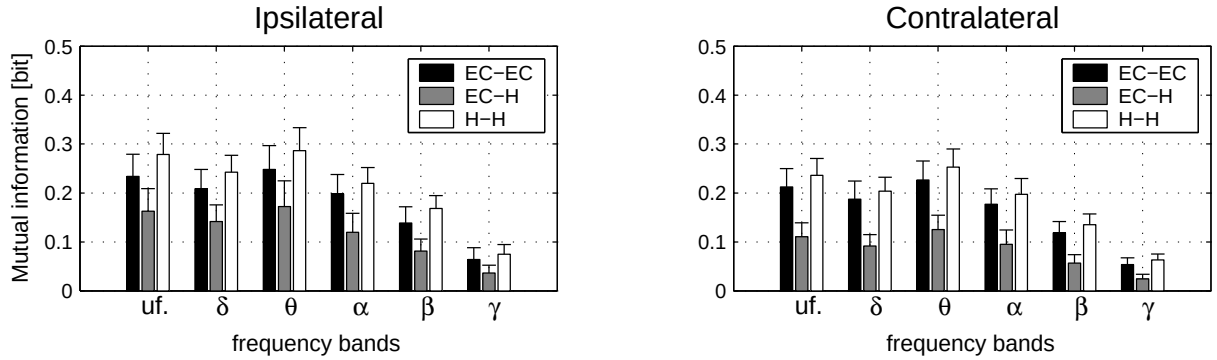


Figure 5.4.: Intra- and inter-regional mutual information (mean values) for a group of 21 epilepsy patients. Mutual information is estimated for intra- and inter-regional contact-combinations within (EC-EC,H-H) or between (EC-H) the entorhinal cortex (EC) and the hippocampus (H) of ipsilateral or contralateral brain hemispheres for different frequency bands. The error bars denote standard errors over a group of 21 epilepsy patients.

This was motivated by findings obtained in [MOA+08] where the authors provided evidence for the existence of dominant rhythmic activity in this frequency range. Since we only concentrate on spatial aspects of SEEG dynamics, we computed the averaged symbolic mutual information $\mathbf{MI}(i,j) = (1/\Lambda)\sum_{n=1}^{\Lambda} MI^n(i,j)$. An exemplary mutual information matrix $\mathbf{MI}$ for one patient is shown in Fig.5.3. Obtained values of $\mathbf{MI}$ for this patient indicated, on average, higher values of mutual information between SEEG recordings corresponding to intra-regional contacts than values obtained for inter-regional contact-combinations for both brain hemispheres. This finding supports our hypothesis that intra-regional SEEG recordings, on average, share more common information than SEEG recordings corresponding to inter-regional contact-combinations. Thus, both neuroanatomical structures can be identified as clusters of higher values of mutual information as it is shown in Fig.5.3. We observed that this effect was, in general, less pronounced in the *ipsilateral* (i.e., the brain hemisphere containing an epileptic focus) than in the *contralateral* brain hemisphere. This finding is in agreement with results obtained in [Ost08].
As a next step, in order to investigate the influence of interindividual variability between epilepsy patients we estimated mutual information $\mathbf{MI}$ between SEEG recordings for a group of 21 epilepsy patients[2]. For each patient we then computed the intra- and inter-regional mutual informations which were defined as

$$\begin{aligned} \mathbf{MI}_{intra}^{h,ec} &= (1/K) \sum_{(i,j)\in\Omega_{h,ec}} \mathbf{MI}(i,j) \\ \mathbf{MI}_{inter} &= (1/K) \sum_{i\in\Omega_h, j\in\Omega_{ec}} \mathbf{MI}(i,j) \end{aligned} \tag{5.1}$$

[2]Five patients (from a group of 26 patients), with SEEG recordings containing less than two contacts located in the entorhinal cortex were excluded from the analysis.

Frequency bands	raw data	δ	θ	α	β	γ
ipsilateral						
RNP	18/21	19/21	18/21	18/21	18/21	19/21
Significance	$8.4 \cdot 10^{-8}$	$4.1 \cdot 10^{-9}$	$8.4 \cdot 10^{-8}$	$8.4 \cdot 10^{-8}$	$8.4 \cdot 10^{-8}$	$4.1 \cdot 10^{-9}$
contralateral						
RNP	16/21	18/21	15/21	16/21	18/21	19/21
Significance	$1.0 \cdot 10^{-5}$	$8.4 \cdot 10^{-8}$	$7.2 \cdot 10^{-5}$	$1.0 \cdot 10^{-5}$	$8.4 \cdot 10^{-8}$	$4.1 \cdot 10^{-9}$

Table 5.2.: The relative number of patients (RNP) showing higher values of the symbolic mutual information between two channels located within either entorhinal cortex or hippocampus than values obtained for inter-regional channel-combinations. Additionally, the corresponding significance levels are shown.

where K denotes the number of summands in Eqs. 5.1, whereas Ω_{ec} and Ω_h are the sets of contacts located in the entorhinal cortex and hippocampus, respectively. It is known that rhythmic patterns of activity of the human brain reflect different physiological and pathophysiological aspects of brain functioning [EP97]. Following [Ost08, MOA+08] we – additionally to the analysis of mutual information of raw SEEG recordings – repeated our analysis for SEEG recordings which were now filtered in different physiologically relevant frequency bands: δ, θ, α, β, and $\gamma-$ bands (Table 5.1). In order to filter the data, we applied a $5^{th}-$order Butterworth zero-phase filter (for further details see chapter 12 in [Ste75]). We found that for a majority of patients obtained values of inter-regional mutual informations ($\mathbf{MI}_{inter}$) were smaller than values of intra-regional mutual informations ($textbf MI^{h,ec}_{intra}$) for all frequency bands. Despite the interindividual variability the mean values of intra-regional mutual informations were indeed smaller than the mean values of inter-regional mutual informations for both brain hemispheres and all frequency bands, i.e., $\bar{\mathbf{MI}}^{hip}_{intra} > \bar{\mathbf{MI}}_{inter}$ and $\bar{\mathbf{MI}}^{ec}_{intra} > \bar{\mathbf{MI}}_{inter}$ (where the bar denotes an average value over 21 patients; see Fig. 5.4). Our findings also indicated that the mean intra- and inter-regional mutual informations on the ipsilateral brain hamisphere slightly exceeded values obtained on the contralateral side of the brain. This observation is in agreement with findings shown in [Mor03, Ost08] where the authors observed, on average, a higher degree of synchronization between SEEG recordings on the ipsilateral brain hemisphere. By comparing our findings for SEEG recordings filtered in different frequency bands we observed that the mean intra- and inter-regional mutual informations had maximal values when our data was filtered in $\theta-$band. This might indicate the presence of a dominant rhythmic activity in this frequency range. Indeed, it is known that theta-oscillations (i.e., rhythmic activity in the $\theta-$band) of mammals are involved in various cognitive functions [KSJ00]. This activity is believed to be critical for the coordination of neuronal networks and for the modification of synaptic connections in the hippocampus [Buz02]. It is believed that theta-oscillations are important for normal functioning of the hippocampus. Clinical evidence indicated that manipulations with the hippocampus which disturb theta-oscillations produce behavioral impairments that mimic hippocampal lesions (for more details see [Lub09] and references therein).

It is important to point out that obtained values of the intra- and inter-regional mutual

information exhibited a relatively high variability in our group of 21 patients (see error bars in Fig. 5.4). This variability was, in general, higher for the ipsilateral brain hemisphere. In order to test our hypothesis H_1 that the intra-regional mutual information $\mathbf{MI}^h_{intra}$ ($\mathbf{MI}^{ec}_{intra}$) are indeed higher than the inter-regional mutual information $\mathbf{MI}_{inter}$ we applied a statistical test which was used in [Ost08]. To do this, we considered the null-hypothesis H_0 that our findings have random nature and therefore the intra-regional mutual information is not higher than the inter-regional mutual information. First, we recall that if our null-hypothesis is true then the a priori probabilities for the case $\mathbf{MI}^{ec}_{intra} > \mathbf{MI}_{inter}$ and for $\mathbf{MI}^h_{intra} > \mathbf{MI}_{inter}$ are 1/3 each. By counting the relative frequency of patients which fulfill H_1 it was possible to estimate the probabilities for the incorrect rejection of the null-hypothesis and thus to estimate a statistical significance level α for our results. To estimate α, we used the binomial distribution[3] with probability $p = 1/3$. Table 5.2 shows the relative numbers of patients which fulfill H_1. In this work we consider results to be statistically significant for $\alpha \leq 0.05$. The obtained significance levels exceeded 0.05 for all frequency bands as well as for the unfiltered data. This indicated that our hypothesis (i.e., the fact that the intra-regional mutual information is higher than the inter-regional mutual information) cannot be explained by random fluctuations of the symbolic mutual information among patients, despite the relatively high interindividual variability.

5.3. Characterizing directions of interactions in SEEG recordings from epilepsy patients

As it was already mentioned above theta-oscillations clock hippocampal activity during awake behavior and are critical for the modification of synaptic connections in the hippocampus [KSJ00]. These oscillations play an important role for the functioning of the hippocampus by grouping and segregating neural assemblies with each other [Buz02, Buz06]. It is believed that theta-oscillations are synchronized throughout this neuroanatomical structure. However, recent findings obtained from freely behaving rats indicated that theta-oscillations in the local field potential are traveling waves which propagate along the septotemporal axis[4] of the hippocampus [Lub09]. These findings demonstrated that theta-oscillations might not only clock the hippocampal activity but also pattern it across anatomical space. The existence of traveling waves might indicate the presence of a preferable direction of information flow in the hippocampus. Thus, the investigation of this phenomenon can be very important for our understanding of the nature of information processing in the hippocampus.
In this section we apply the symbolic transfer entropy TE as well as the corrected symbolic transfer entropy TE_c to characterize the directionality of interactions in the hippocampus. In order to obtain the symbolic representation of real-valued SEEG data, we first apply the permutation symbols approach. In addition, we will repeat our analysis by using a

[3]The significance level was estimated as $\alpha = \sum_{i=N_{pat}+1}^{21} B(N_{pat}, p)$ where $B(N_{pat}, p)$ denotes the binomial distribution with probability $p = 1/3$ and N_{pat} is the number of patients fulfilling H_1.

[4]The long axis of the hippocampal formation is referred to as the Septotemporal axis. For further details see [KSJ00].

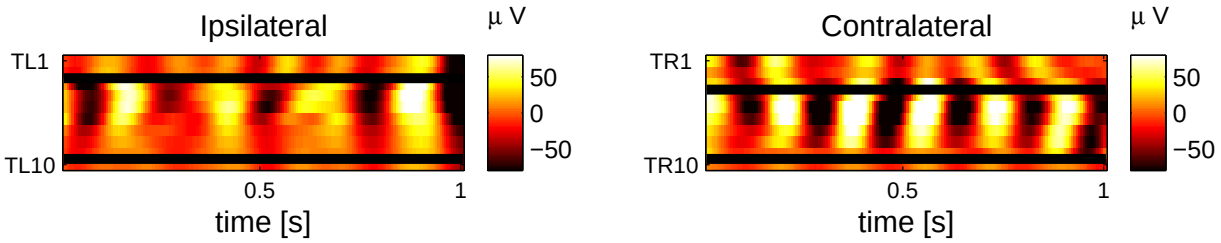


Figure 5.5.: A short segment of SEEG recordings from ten electrode contacts located in either ipsilateral (TL1-TL10) or contralateral (TR1-TR10) brain hemispheres of an epilepsy patient. The data was additionally filtered in the $\theta-$band by using a $5^{th}-$order Butterworth zero-phase filter. Black horizontal lines indicate the anterior (frontal) and posterior (backmost) boarders of left and right hippocampi (see Fig. 5.1).

conceptually different symbolization approach (the so called binning symbols approach). This will allow us to address the question as to what extend the obtained findings reflect our choice of symbolization parameters.

Exemplary analysis of a wave-traveling phenomenon in the hippocampus

In this section we address the question as to whether a wave-traveling phenomenon can take place along the longitudinal axis of the hippocampal formation in humans (see Fig. 5.1). Analysis of the degree of synchronization in SEEG recordings from epilepsy patients revealed independent delta and theta rhythms in the hippocampus [MOA$^+$08]. Analysis of mutual information between SEEG recordings of epilepsy patients made above (see Fig. 5.4) also indicated the presence of dominant rhythmic activity in the $\theta-$band. For some patients, SEEG recordings contain segments exhibiting theta-oscillations which can be better identified by filtering SEEG data in the $\theta-$band. For instance, in Fig. 5.5 one can see a short segment of band-pass filtered SEEG recordings exhibiting theta-oscillations in the hippocampus (contacts: TR4-TR9) of the non-focal (contralateral) brain hemisphere. Moreover, a qualitative comparison of theta-oscillations for different electrode contacts allowed us to identify that these oscillations represent traveling waves propagating from the posterior to the anterior border of the hippocampus (i.e., from TR9 to TR4; see Fig. 5.1). One of the approaches to characterize traveling waves in electroencephalographic recordings is to calculate the cross-correlation coefficients between them [EP97]. In order to test our assumption that SEEG recordings of this patient indeed exhibit a traveling-wave phenomenon in the posterior-anterior direction of the hippocampus, we estimated cross-correlation coefficients between the most posterior hippocampal contact (TL9 for left hemisphere and TR9 for right hemisphere) and all other contacts for different time delays between SEEG recordings. To do this we defined a cross-correlation coefficient $\Xi(i,j,\Delta t)$ as

$$\Xi(i,j,\Delta t) = (1/N)\sum_{n=1}^{N-\Delta t} \bar{x}_i(n)\bar{x}_j(n+\Delta t) \tag{5.2}$$

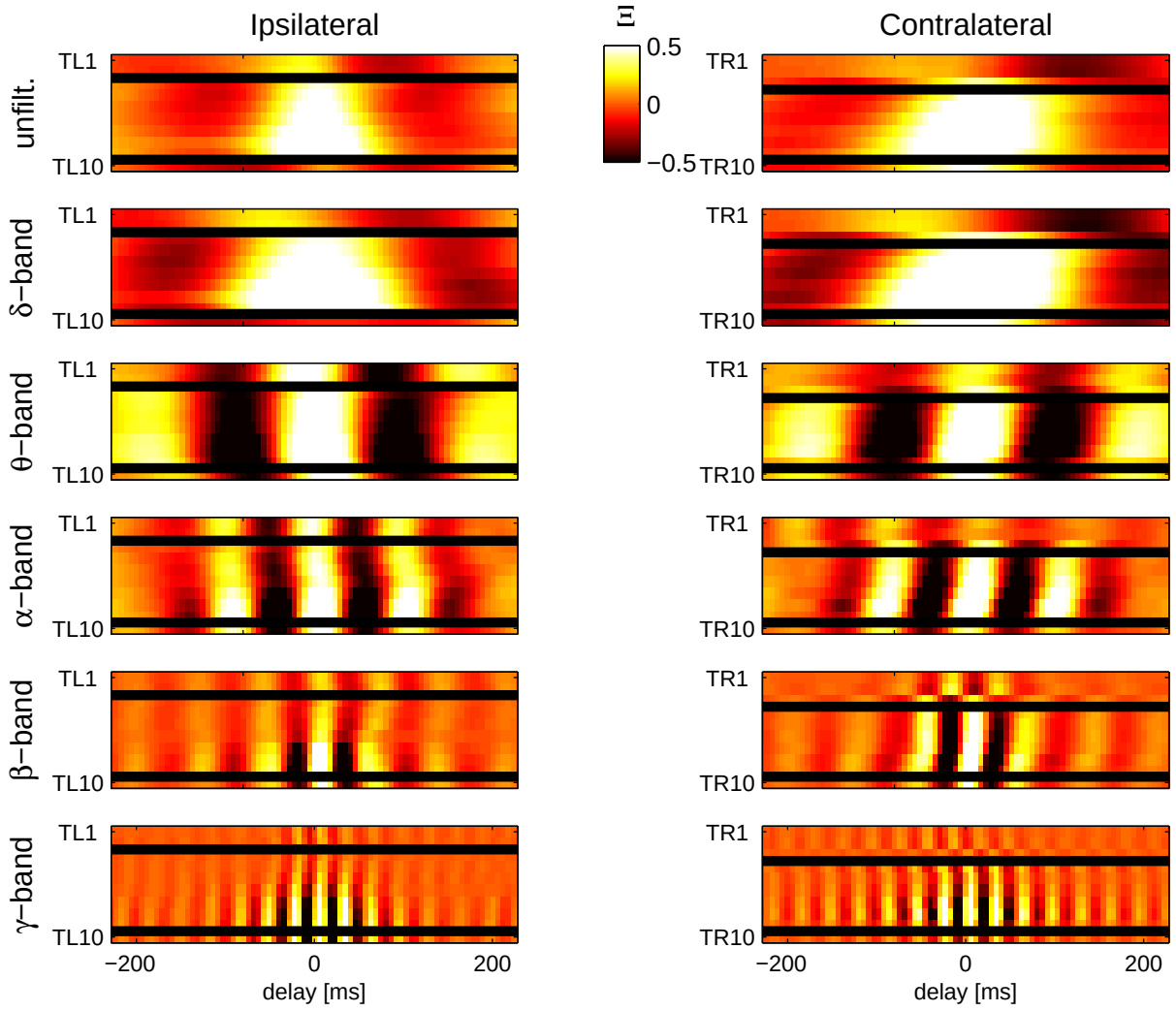


Figure 5.6.: Exemplary analysis of a wave-traveling phenomenon in the hippocampus. Cross-correlation coefficients between SEEG recordings of the most posterior contact located in the left (right) hippocampus TL9 (TR9) and all other contacts TL1-TL10 (TR1-TR10) for ipsilateral (contralateral) brain hemispheres for different frequency bands. Black horizontal lines indicate the anterior and posterior boarders of left and right hippocampi.

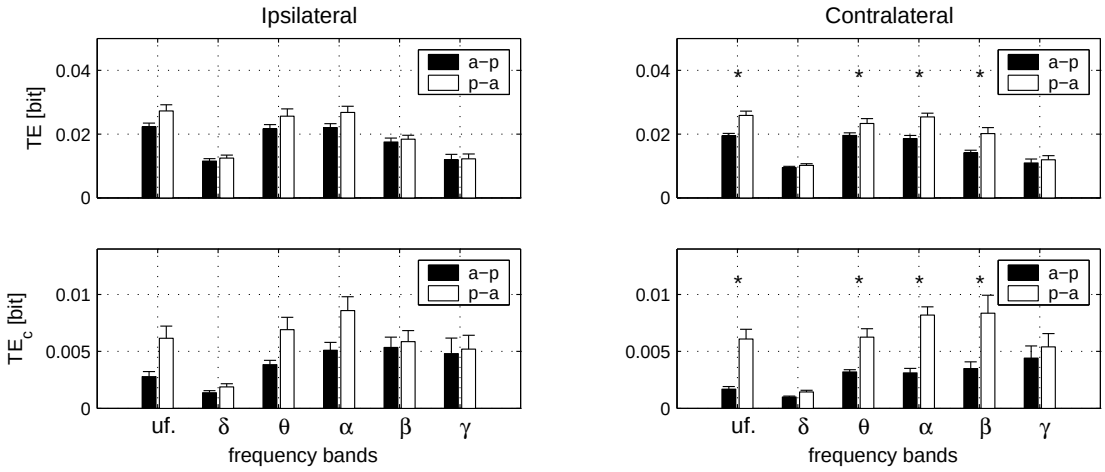


Figure 5.7.: Anterior-posterior (a-p) and posterior-anterior (p-a) transfer entropy flows in the hippocampus $\bar{\mathbf{TE}}^{pa}$ and $\bar{\mathbf{TE}}^{ap}$ of ipsilateral and contralateral brain hemispheres for raw (uf.) and band-pass filtered SEEG recordings. The analysis of entropy transfer was performed by using either the symbolic transfer entropy TE or corrected symbolic transfer entropy TE_c and permutation symbols approach. The error bars denote standard errors over a group of 26 epilepsy patients. The stars indicate statistically significant results.

where $\bar{x}_i$ and $\bar{x}_j$ denote a pair of normalized SEEG recordings (with zero mean and unit variance) of length N and Δt is a time delay. The cross-correlation coefficient approaches 1 (-1) if a pair of SEEG recordings (i, j) are completely correlated (anti-correlated) and approaches 0 if signals are uncorrelated. For the analysis of wave-traveling phenomena in the posterior-anterior direction of the right hippocampus, we computed cross-correlation coefficients $\Xi(i, j, \Delta t)$ between the most posterior hippocampal contact ($i = \mathrm{TR9}$) and all other contacts ($j = \mathrm{TR1}, \ldots, \mathrm{TR10}$) for different time delays $\Delta t = -40, \ldots, 40$. Fig. 5.6 shows cross-correlation coefficients for unfiltered as well as for band-pass filtered data in different physiologically relevant frequency bands as a function of time delay Δt. All cross-correlation coefficients were estimated for a segment of SEEG recordings of approximately 10 minutes duration ($N = 10^5$). Our analysis showed that $\Xi(\mathrm{TR9}, j, \Delta t) \neq \Xi(\mathrm{TR9}, j, -\Delta t)$ for $j = \mathrm{TR4}, \ldots, \mathrm{TR9}$. The observed time-asymmetry of cross-correlation coefficients indicates the presence of traveling waves in the posterior-anterior direction of the right hippocampus. This finding extends our observation made for a short segment of SEEG recordings (see Fig. 5.5) and indicates that a wave-traveling phenomenon may exist on time intervals of at least 10 minutes duration. It is interesting to point out that such a wave-traveling phenomenon appeared to be less pronounced in $\beta-$ and almost did not take place in $\gamma-$band. Moreover, analysis of a traveling-wave phenomenon in the left hippocampus (i.e. in the ipsilateral or focal brain hemisphere) revealed that $\Xi(\mathrm{TL9}, j, \Delta t) \approx \Xi(\mathrm{TL9}, j, -\Delta t)$ for $j = \mathrm{TL3}, \ldots, \mathrm{TL9}$ in all frequency bands. The fact that the traveling waves appeared to be considerably less pronounced in the focal (ipsilateral) brain hemisphere indicates that mechanisms underlying this phenomenon might be impaired by epilepsy.

Measuring entropy transfer in SEEG recordings using permutation symbols approach

The analysis of wave propagation phenomena in SEEG recordings using the cross-correlation coefficient can be limited by only providing us a linear relationship between signals. By assuming that wave-traveling phenomena in spatiotemporal systems may result in the asymmetry of information transport the analysis based on cross-correlation coefficient can be extended by using time-delayed mutual information. Application of time-delayed mutual information allows us to capture nonlinear relationships between signals and was already applied for the analysis of spatiotemporal phenomena in electroencephalograms [EP97]. However, as it was already point out in [Sch00] the application of time-delayed mutual information to capture the velocity of wave propagation in spatiotemporal systems can sometimes provide misleading results. It was also pointed out that in these cases the application of transfer entropy can provide us a better alternative for the characterization and correct inference of information transfer in spatiotemporal systems. As we could see in section 4.2.2 the estimation of entropy transfer between two dynamical systems allowed us to infer the directionality of interactions between them. Moreover, analysis of entropy transfer in multivariate data generated by a network of coupled chaotic oscillators allowed us to identify driving structures in our network and thus to infer the directionality of interactions between oscillators (see section 4.3). In this section we investigate the problem of asymmetry of interactions between SEEG recordings from epilepsy patients by estimating entropy transfer between these signals using the symbolic transfer entropy TE as well as corrected symbolic transfer entropy TE_c (see section 4.2.1). By comparing (averaged over all hippocampal contact-combinations) transfer entropy flows in the hippocampus we characterize directionality of interactions in the hippocampus.

Let us continue our analysis of SEEG recordings from the epilepsy patient considered above (see Figs. 5.5 and 5.6). To proceed, we first divided SEEG recordings into Λ consecutive windows of approximately 58 seconds ($N = 10^4$ data points) duration. For each window $n = 1, \ldots, \Lambda$ and each pair of contacts $(i, j) = \{\mathrm{TL1}, \ldots, \mathrm{TL10}, \mathrm{TR1}, \ldots, \mathrm{TR10}\}$ (see the implantation scheme in Fig. 5.1, and also Fig. 5.2) we computed the symbolic transfer entropy $TE^n(i, j)$ as well as corrected symbolic transfer entropy $TE_c^n(i, j)$ (Eqs. 4.5 and 4.9) with embedding parameters $m = 3$ and $\tau = 10$. The embedding window $w_{emb} = (m - 1)\tau$ was again set to approximately cover a period of oscillatory activity belonging to the $\theta-$ frequency band (see section 5.3). Since we only interest here on spatial aspects of SEEG dynamics, we computed the averaged symbolic transfer entropy $\mathbf{TE}(i, j) = (1/\Lambda) \sum_{n=1}^{\Lambda} TE^n(i, j)$ and averaged corrected symbolic transfer entropy $\mathbf{TE}_c(i, j) = (1/\Lambda) \sum_{n=1}^{\Lambda} TE_c^n(i, j)$. Fig.5.3 shows $\mathbf{TE}(i, j)$ and $\mathbf{TE}_c(i, j)$ as well as the averaged mutual information $\mathbf{MI}(i, j)$ computed between all contact-combinations of SEEG recordings. According to its definition (Eq.4.1) the symbolic mutual information characterizes the amount of common information between a pair of SEEG recordings and thus is symmetric (i.e., $\mathbf{MI}(i, j) = \mathbf{MI}(j, i)$). A qualitative comparison of entropy transfer between hippocampal contact-combinations ($(i, j) = \{\mathrm{TL3}, \mathrm{TL9}\}$ or $(i, j) = \{\mathrm{TR4}, \mathrm{TR9}\}$) indicated, on average, higher values of the (corrected) symbolic transfer entropy $\mathbf{TE}(i, j)$ ($\mathbf{TE}_c(i, j)$) in posterior-anterior ($i > j$) than in anterior-posterior ($i < j$) directions. The asymmetry of interactions between hippocampal contacts appeared to be even more pronounced in the

non-focal (contralateral) hemisphere (TR4-TR9) (see Fig.5.3). Recalling that the traveling waves in the posterior-anterior direction observed above (Figs. 5.5 and 5.6) were also more pronounced in the contralateral brain hemisphere we can conclude that, at least for this patient, the asymmetry of the entropy transfer in the hippocampus might be related to the wave-traveling phenomenon.
In order to quantitatively characterize the asymmetry of entropy transfer in the hippocampus we defined the posterior-anterior (p-a) and anterior-posterior (a-p) transfer entropy flows as

$$\begin{aligned} \mathbf{TE}^{pa} &= (1/K) \sum_{(i>j)\in\Omega_h} \mathbf{TE}(i,j) \\ \mathbf{TE}^{ap} &= (1/K) \sum_{(i<j)\in\Omega_h} \mathbf{TE}(i,j) \end{aligned} \tag{5.3}$$

where K denotes the number of summands in Eqs. 5.3, and Ω_h is a set of electrode contacts located in the hippocampus of either ipsilateral or contralateral brain hemispheres. In order to investigate the influence of interindividual variability between different patients, we estimated the posterior-anterior and anterior-posterior transfer entropy flows $\mathbf{TE}^{pa}$ and $\mathbf{TE}^{ap}$ from a group of 26 epilepsy patients. In addition to the analysis of entropy transfer for raw SEEG recordings, we repeated our analysis for SEEG recordings which were filtered in different physiologically relevant frequency bands: δ, θ, α, β, and $\gamma-$ bands (Table 5.1). Analysis of entropy transfer in raw SEEG recordings revealed that – despite an interindividual variability between patients – the mean value of the hippocampal transfer entropy flow in the posterior-anterior direction exceeded the hippocampal transfer entropy flow in the opposite direction for both brain hemispheres (i.e., $\bar{\mathbf{TE}}^{pa} > \bar{\mathbf{TE}}^{ap}$; the bar denotes the average value over 26 patients). By performing the same analysis for SEEG recordings filtered in different frequency bands, we observed that the asymmetry of the mean transfer entropy flows can be identified in $\theta-$, $\alpha-$, and $\beta-$ bands for the non-focal (contralateral) and only in $\theta-$, $\alpha-$ bands for focal (ipsilateral) hemispheres. This poses a question as how and to what extent the observed difference between brain hemispheres related to epilepsy.
Obtained values of $\mathbf{TE}^{pa}$ and $\mathbf{TE}^{ap}$ exhibited a relatively high variability between all 26 patients (see error bars in Fig. 5.7). The observed variability was, in general, slightly higher for the ipsilateral brain hemisphere. In order to determine the statistical significance of our results, we followed analysis made in section 5.2 (see also Table 5.2) and examined the hypothesis H_2 which states that the hippocampal transfer entropy flow is higher in posterior-anterior than in the opposite direction. To test this hypothesis, we again additionally considered a null-hypothesis H_0 that our findings have random nature and therefore the hippocampal transfer entropy flow in posterior-anterior does not exceed the entropy flow in the opposite direction. First, we recall that if our null-hypothesis is true then the a priori probability for the case $\mathbf{TE}^{pa} > \mathbf{TE}^{ap}$ is $1/2$. By counting the relative frequency of patients which fulfill H_2 it was possible to estimate the probabilities for the incorrect rejection of the null-hypothesis. To do this, we again used a binomial distribution and computed the significance level as $\alpha = \sum_{i=N_{pat}+1}^{26} B(N_{pat}, p)$ (where N_{pat} is the number of patients fulfilling

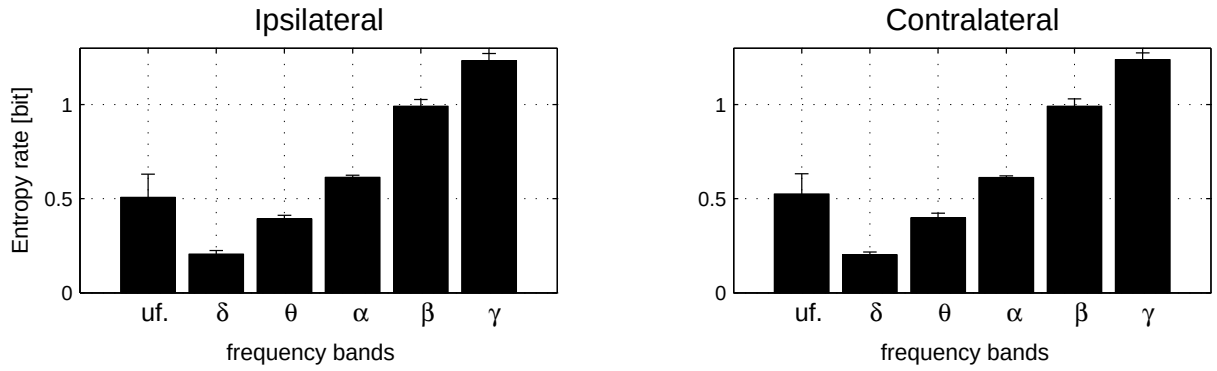


Figure 5.8.: Entropy rates in the hippocampus $\mathbf{d\bar{H}}_h$ for ipsilateral and contralateral brain hemisphere. The analysis of entropy rate was performed for raw (uf.) as well as for band-pass filtered SEEG recordings. The error bars denote standard errors over a group of 26 patients.

H_2 and $B(N_{pat}, p)$ denotes a binomial distribution with the probability $p = 1/2$). Table 5.3 shows the relative number of patients fulfilling H_2 along with the corresponding significance levels. Obtained significant values for the contralateral hemisphere indicated that our null hypothesis H_0 can be rejected for the case of unfiltered SEEG recordings as well as for cases when the SEEG data was filtered in $\theta-$, $\alpha-$, and $\beta-$ bands[5]. We also observed that for the ipsilateral hemisphere the null hypothesis cannot be rejected for the raw data as well as for all frequency bands. This finding indicates that mechanisms underlying asymmetry in the direction of interactions in the hippocampus might be impaired by epilepsy.
Similar findings were obtained by estimating hippocampal transfer entropy flows by using the corrected symbolic transfer entropy TE_c (see Fig. 5.7). However, our findings also indicated a more pronounced difference between $\mathbf{T\bar{E}}_c{}^{pa}$ and $\mathbf{T\bar{E}}_c{}^{ap}$ than between $\mathbf{T\bar{E}}^{pa}$ and $\mathbf{T\bar{E}}^{ap}$. Analysis of raw as well as band-pass filtered (in $\theta-$, $\alpha-$, and $\beta-$ bands for the contralateral and in $\theta-$, $\alpha-$ bands for the ipsilateral brain hemispheres) SEEG recordings showed that the asymmetry between hippocampal transfer entropy flows was also more pronounced when entropy transfer was estimated with TE_c. Additionally, we also observed that $\mathbf{TE}_c^{pa}$ and $\mathbf{TE}_c^{ap}$, in general, exhibited a higher variability than $\mathbf{TE}^{pa}$ and $\mathbf{TE}^{ap}$ (see error bars in Fig. 5.7). Despite a more pronounced difference between posterior-anterior and anterior-posterior transfer entropy flows the approach to estimate entropy transfer with the corrected symbolic transfer entropy did not lead to statistically more significant results (see Tables 5.3 and Tables 5.4).

Normalized transfer entropy

Frequency-selective analysis indicated relatively high mean values of transfer entropy flows $\mathbf{T\bar{E}}^{pa}$ and $\mathbf{T\bar{E}}^{ap}$ in $\alpha-$, $\beta-$, and even in $\gamma-$bands (see in Fig. 5.7). The maximal values of $\mathbf{T\bar{E}}^{pa}$ and $\mathbf{T\bar{E}}^{ap}$ could be observed in the $\alpha-$band, when entropy transfer in SEEG

[5]In this chapter, we additionally applied the Bonferroni correction for the statistical analysis of results obtained from different frequency bands (for details see [Abd06]).

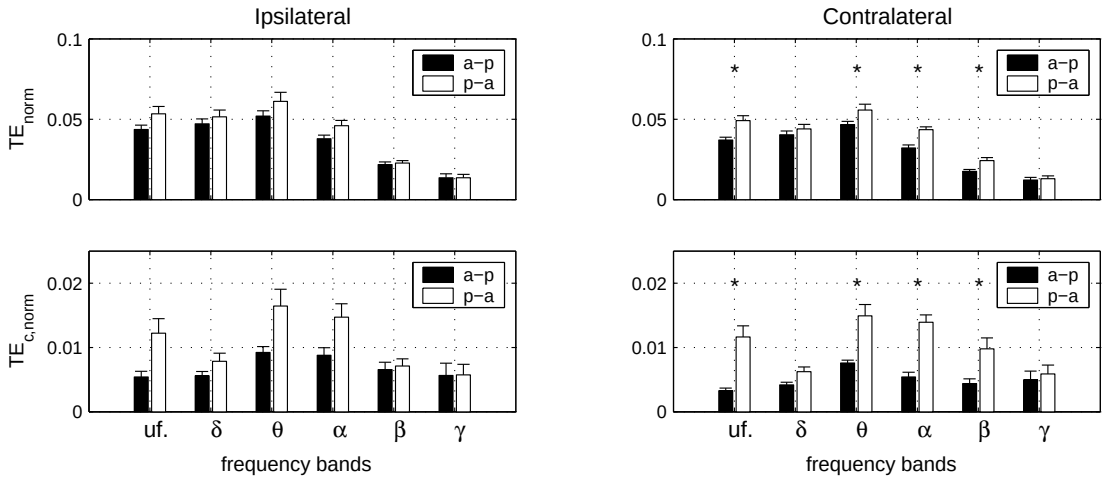


Figure 5.9.: Anterior-posterior (a-p) and posterior-anterior (p-a) transfer entropy flows in the hippocampus $\bar{\mathbf{TE}}^{pa}_{norm}$ and $\bar{\mathbf{TE}}^{ap}_{norm}$ of ipsilateral and contralateral brain hemispheres for raw (uf.) and band-pass filtered SEEG recordings. The analysis of entropy transfer was performed by using the normalized versions of either the symbolic transfer entropy TE_{norm} or corrected symbolic transfer entropy $TE_{c,norm}$ and permutation symbols approach. The error bars denote standard errors over a group of 26 epilepsy patients. The stars indicate statistically significant results.

recordings was estimated with the symbolic transfer entropy TE, and even in $\beta-$ and $\gamma-$ bands for the analysis made with the corrected symbolic transfer entropy TE_c. Since it is known that theta-oscillations represent a dominant rhythmic activity in the hippocampus [Buz02, MOA$^+$08, Lub09] our findings obtained above (see Fig. 5.7) may appear to be rather unexpected. To proceed, we need to briefly address the question as to how and to what extent the information-theoretic approach can be applied for the frequency-selective analysis of entropy production and entropy transfer in band-pass filtered data. It is known that the capacity of a continuous channel to transmit information is proportional to the frequency band of transmitted and received signals[6] [Sha48, CT91]. It is thus logical to assume that signals containing high-frequency components can provide higher entropy rate and transmit more information (per time step) than signals containing low-frequency components. In order to investigate this question we – additionally to the analysis of entropy transfer – also estimated entropy rates of SEEG recordings of epilepsy patients for the cases of raw as well as band-pass filtered data.

To proceed, we divided SEEG recordings of each patient into Λ consecutive windows of approximately 58 seconds ($N = 10^4$ data points) durations. For each window $n = 1, \ldots, \Lambda$ and each contact $i \in [\mathrm{TL1}, \ldots, \mathrm{TL10}, \mathrm{TR1}, \ldots, \mathrm{TR10}]$ (see the implantation scheme in Fig. 5.1) we computed the first order entropy rate $dH^n(i)$ (see Eq. 3.3) with the embedding parameters $m = 3$ and $\tau = 10$. We used the same embedding parameters as for the analysis

[6]For instance, the analog telephone line is bandlimited to 3300 Hz which sets the maximum capacity of this channel to transmit information to 56 kbit/s [CT91].

of the (corrected) symbolic transfer entropy made above. By averaging over all consecutive windows we then computed, for each patients and each contact i, the averaged entropy rate $\mathbf{dH}(i) = (1/\Lambda)\sum_{n=1}^{\Lambda} dH^n(i)$. Next, by further averaging over corresponding contacts we computed the mean entropy rate in the hippocampus, which was defined as

$$\mathbf{dH}_h = (1/K)\sum_{i\in\Omega_h} \mathbf{dH}(i) \tag{5.4}$$

where K denotes the number of summands in Eq. 5.4, and Ω_h denotes a set of contacts located in the hippocampus of either ipsilateral or contralateral brain hemispheres. Along with the analysis of raw data we estimated the mean entropy rate $\mathbf{dH}_h$ of SEEG recordings filtered in different frequency bands (see Table 5.1). Despite the interindividual variability between patients our findings indicated that the obtained values of $\mathbf{d\bar{H}}_h$ (bar denotes the average value over 26 patients) were significantly different for all investigated frequency bands (see Fig. 5.8). For instance, by analyzing the raw data we found that the mean entropy rate in the hippocampus was, on average, $\mathbf{d\bar{H}}_h \approx 0.5 \pm 0.1$ bit per time step. However, the filtering of SEEG data in $\delta-$ and $\theta-$ bands resulted in the reduction of this value to $\mathbf{d\bar{H}}_h \approx 0.20 \pm 0.02$ and $\mathbf{d\bar{H}}_h \approx 0.40 \pm 0.02$ bit per time step, respectively. In contrast to this, when the data was filtered in $\alpha-$, $\beta-$, and $\gamma-$ bands the values of the mean entropy rate exceeded the estimate obtained for raw SEEG data. Let us recall that according to its definition, the entropy rate approaches zero for periodic dynamics (see the analysis of entropy rate of tent map time series in chapter 3). Thus, reduced values of the mean entropy rate in the hippocampus might indicate the presence of dominant oscillations in $\delta-$ and $\theta-$ frequency bands. Whereas, e.g. in $\gamma-$band the mean entropy rate almost approached its maximal value $(\log(m!)/(m-1) \approx 1.3$; for $m = 3)$ and thus indicated the absence of oscillatory activity in this frequency band.
The comparison of mean entropy rates in the hippocampus for different physiologically relevant frequency bands revealed that obtained values of $\mathbf{d\bar{H}}_h$ were higher when the data was filtered in frequency bands which corresponded to higher frequencies (see Fig. 5.8). This finding poses a further question as to how and to what extent the absolute values of the mean transfer entropy flows $\mathbf{\bar{TE}}^{pa}$ and $\mathbf{\bar{TE}}^{ap}$ obtained for different frequency bands can be compared with each other (see Fig. 5.7). For instance, it is not clear whether the observed difference in entropy transfer between $\theta-$ and $\alpha-$ bands reflects the presence of dominant oscillations in the $\alpha-$ band or our findings just reflect the difference in entropy rates between these frequency bands (see Fig. 5.8). Following these considerations, and in order to avoid a possible influence of differences in entropy rates between different frequency bands, we defined a normalized measure for entropy transfer as

$$TE^{norm}(i,j) = TE(i,j)/dH(i) \tag{5.5}$$

where $dH(i)$ denotes the entropy rate of SEEG recording of contact i and $TE(i,j)$ is a measure of entropy transfer (in our case TE or TE_c) between SEEG recordings corresponding to a pair of contacts i and j. According to this definition, the normalized transfer entropy

is thus confined to the interval $[0, 1]$. Indeed, it is easy to see that a zero entropy transfer from i to j directly implies $TE^{norm}(i, j) = 0$ whereas the case $TE^{norm}(i, j) = 1$ implies that $TE(i, j) = dH(i)$ and therefore all entropy generated in SEEG recording of contact i is completely transfered to contact j. Thus, according to its definition, $TE^{norm}(i, j)$ is a dimensionless measure which characterizes the amount of entropy transfered from contact i to contact j relative to the amount of entropy generated in contact i. Similar approach for the normalization of transfer entropy was already used in [MK02].
To proceed, we again estimated the posterior-anterior and anterior-posterior transfer entropy flows in the hippocampus $\mathbf{TE}^{pa}_{norm}$ and $\mathbf{TE}^{ap}_{norm}$ (Eq. 5.3) but now by normalizing the corresponding values of the (corrected) symbolic transfer entropy TE (TE_c) according to Eq. 5.5. In contrast to our previous findings, we now found that the normalized mean transfer entropy flows $\bar{\mathbf{TE}}^{pa}_{norm}$ and $\bar{\mathbf{TE}}^{ap}_{norm}$ (bar denotes the average value over 26 patients) approached maximal values when SEEG data was filtered in the $\theta-$band (see Fig. 5.9). This finding appears to be physiologically motivated indicating on a presence of theta oscillations in the hippocampus. Significance levels of obtained results for different frequency bands are summarized in Tables 5.3 and 5.4.

Measuring entropy transfer in SEEG recordings using a binning symbols approach

As a final step of our analysis we address the question as to how and to what extent the findings obtained in this section might depend on the choice of embedding parameters m and τ which were used to compute the (corrected) symbolic transfer entropy TE (TE_c) between SEEG recordings. Indeed, as we could see in section 4.2, the application of the (corrected) symbolic transfer entropy to characterize the directionality of interactions between coupled dynamical model systems required a careful selection of embedding parameters. For instance, as one can see from Figs. 4.7 and 4.8 an inappropriate selection of the delay time τ can lead to an incorrect inference of the directionality of interactions between two coupled chaotic oscillators. This finding poses an important question as to whether the observed asymmetry of transfer entropy flow in the posterior-anterior direction is indeed related to the directionality of interactions in the hippocampus and to what extend this effect is a result of the incorrect selection of the symbolization parameters. The analysis of entropy transfer in model dynamical systems demonstrated that for signals exhibiting a dominant spectral component the embedding window $w_{emb} = (m - 1)\tau$ should approximately cover a basic period of an investigated dynamical system (see also [SL08]). In this section the embedding parameters were set to $m = 3$ and $\tau = 10$ such as to approximately cover a period of oscillatory activity belonging to the $\theta-$ band. By converting a real-valued SEEG data into series of permutation symbols we thus introduced an additional time parameter which is defined by the embedding window w_{emb}. This poses an important question as to what extent the findings demonstrating a maximum of the hippocampal transfer entropy flow in the $\theta-$band (see Fig. 5.9) are over-optimized and just reflect our choice of the embedding window w_{emb}.
In order to investigate the question posed above we repeated our analysis of entropy transfer in the hippocampus but now by applying a rather straightforward approach for a symbolization of real-valued data which is based on equidistant partitioning of the dynamical ranges

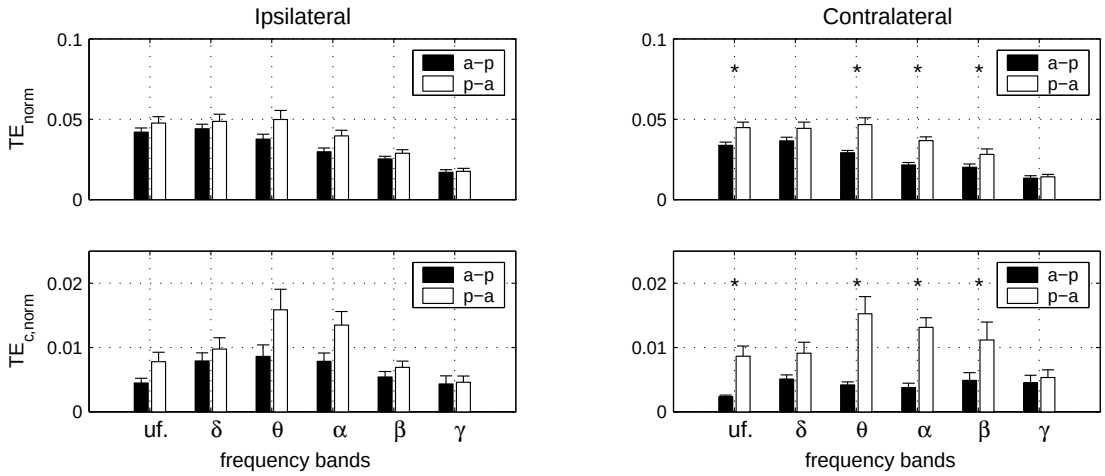


Figure 5.10.: Anterior-posterior (a-p) and posterior-anterior (p-a) transfer entropy flows in the hippocampus $\bar{\mathbf{TE}}^{pa}_{norm}$ and $\bar{\mathbf{TE}}^{ap}_{norm}$ of ipsilateral and contralateral brain hemispheres for raw (uf.) and band-pass filtered SEEG recordings. The analysis of entropy transfer was performed by using normalized versions of either the symbolic transfer entropy TE_{norm} or corrected symbolic transfer entropy $TE_{c,norm}$ and the binning symbols approach. The error bars denote standard errors over a group of 26 epilepsy patients. The stars indicate statistically significant results.

of SEEG signals (for details see section 2.3.1 and [DFT03]). In contrast to the permutation symbols approach, this symbolization procedure does not include a time parameter. To proceed, we divided the amplitude range of each SEEG time series into 6 equidistant intervals or bins. The number of these intervals defines the alphabet length of the binning symbols and it was set to be equal to the alphabet length of the permutation symbols (with embedding dimension $m = 3$). To each bin we then assigned a symbol and thus transformed SEEG recordings into series of *binning* symbols. A comparison of $\bar{\mathbf{TE}}^{pa}_{norm}$ and $\bar{\mathbf{TE}}^{ap}_{norm}$ (bar denotes the average value over 26 patients) obtained for either the permutation or the binning symbols approach revealed qualitatively similar results (see Fig. 5.10). A frequency-selective analysis of entropy transfer again demonstrated that the mean transfer entropy flows $\bar{\mathbf{TE}}^{pa}_{norm}$ and $\bar{\mathbf{TE}}^{ap}_{norm}$ again approached maximal values when SEEG data was filtered in the $\theta-$band. This indicates that the observed maximum of the hippocampal entropy transfer flow in this frequency band, obtained by a permutation symbols approach, cannot only be explained as a result of over-optimization (i.e., just reflecting our choice of the embedding parameters). Significance levels of obtained results for different frequency bands are summarized in Tables 5.3 and 5.4.

Disscussion

The application of the corrected symbolic transfer entropy TE_c indeed allowed us to better characterize the directionality of interactions in SEEG recordings. As it was shown in section 4.3, the application of the corrected symbolic transfer entropy allowed us to

Frequency bands	raw data	δ	θ	α	β	γ
Transfer entropy (permutation symbols)						
Ipsilateral						
RNP	16/26	18/26	18/26	17/26	15/26	16/26
Significance level	$8.4 \cdot 10^{-1}$	$1.4 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$8.4 \cdot 10^{-2}$
Contralateral						
RNP	22/26	17/26	22/26	23/26	21/26	18/26
Significance level	$5.2 \cdot 10^{-6}$	$1.6 \cdot 10^{-1}$	$4.3 \cdot 10^{-5}$	$4.0 \cdot 10^{-7}$	$2.6 \cdot 10^{-4}$	$1.4 \cdot 10^{-2}$
Normalized transfer entropy (permutation symbols)						
Ipsilateral						
RNP	16/26	18/26	18/26	17/26	15/26	14/26
Significance level	$8.4 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$2.7 \cdot 10^{-1}$
Contralateral						
RNP	21/26	16/26	22/26	23/26	21/26	18/26
Significance level	$2.6 \cdot 10^{-4}$	$8.4 \cdot 10^{-2}$	$4.3 \cdot 10^{-5}$	$5.2 \cdot 10^{-6}$	$2.6 \cdot 10^{-4}$	$1.4 \cdot 10^{-2}$
Normalized transfer entropy (binning symbols)						
Ipsilateral						
RNP	16/26	15/26	15/26	17/26	16/26	15/26
Significance level	$8.4 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$1.6 \cdot 10^{-1}$	$3.7 \cdot 10^{-2}$	$8.4 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$
Contralateral						
RNP	17/26	17/26	21/26	23/26	21/26	19/26
Significance level	$3.7 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$2.6 \cdot 10^{-4}$	$5.2 \cdot 10^{-6}$	$2.6 \cdot 10^{-4}$	$4.6 \cdot 10^{-3}$

Table 5.3.: The relative number of patients (RNP) showing higher values of the posterior-anterior than anterior-posterior transfer entropy flow in the hippocampus for ipsilateral and contralateral brain hemispheres. The entropy transfer between SEEG recordings was estimated by using the symbolic transfer entropy TE. Additionally, the corresponding significance levels are shown.

increase the contrast and thus to better characterize driving structures in the network of coupled chaotic oscillators. Particularly, it was demonstrated that the contrast of spurious driving structures (i.e., driving structures which only represent the difference in noise-to-signal ratios between signals) can be decreased when entropy transfer between oscillators is estimated by using the corrected symbolic transfer entropy. A qualitative comparison of posterior-anterior and anterior-posterior transfer entropy flows in the hippocampus $\mathbf{TE}^{pa}$ and $\mathbf{TE}^{ap}$ (see Eq.5.3) obtained with either symbolic transfer entropy TE or corrected symbolic transfer entropy TE_c showed that in the later case we observed, on average, a more pronounced difference between $\mathbf{TE}^{pa}$ and $\mathbf{TE}^{ap}$ (see Figs. 5.7, 5.9, and 5.10). This resulted in an increased contrast and thus a slightly better identification of the asymmetry of transfer entropy flows in the hippocampus. However, it is important to point that the question as to how and to what extent the observed asymmetry transfer entropy flow is related to a traveling wave phenomenon in the hippocampal remained open.

Frequency bands	raw data	δ	θ	α	β	γ
Transfer entropy (permutation symbols)						
Ipsilateral						
RNP	18/26	17/26	17/26	17/26	15/26	15/26
Significance level	$1.4 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$1.6 \cdot 10^{-1}$
Contralateral						
RNP	23/26	15/26	22/26	24/26	21/26	17/26
Significance level	$5.2 \cdot 10^{-6}$	$1.6 \cdot 10^{-1}$	$4.3 \cdot 10^{-5}$	$4.0 \cdot 10^{-7}$	$2.6 \cdot 10^{-4}$	$3.7 \cdot 10^{-2}$
Normalized transfer entropy (permutation symbols)						
Ipsilateral						
RNP	17/26	18/26	17/26	17/26	15/26	13/26
Significance level	$3.7 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$4.2 \cdot 10^{-1}$
Contralateral						
RNP	23/26	17/26	22/26	24/26	21/26	17/26
Significance level	$5.2 \cdot 10^{-6}$	$3.7 \cdot 10^{-2}$	$4.3 \cdot 10^{-5}$	$4.0 \cdot 10^{-7}$	$2.6 \cdot 10^{-4}$	$3.7 \cdot 10^{-2}$
Normalized transfer entropy (binning symbols)						
Ipsilateral						
RNP	17/26	15/26	15/26	16/26	17/26	13/26
Significance level	$3.7 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$1.6 \cdot 10^{-1}$	$8.4 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$4.2 \cdot 10^{-1}$
Contralateral						
RNP	21/26	15/26	21/26	23/26	21/26	17/26
Significance level	$2.6 \cdot 10^{-4}$	$1.6 \cdot 10^{-1}$	$2.6 \cdot 10^{-4}$	$5.2 \cdot 10^{-6}$	$2.6 \cdot 10^{-4}$	$3.7 \cdot 10^{-2}$

Table 5.4.: The relative number of patients (RNP) showing higher values of the posterior-anterior than anterior-posterior transfer entropy flow in the hippocampus for ipsilateral and contralateral brain hemispheres. The entropy transfer between SEEG recordings was estimated by using the corrected symbolic transfer entropy TE_c. Additionally, the corresponding significance levels are shown.

6. Estimating entropy transfer between dynamical systems exhibiting long-term memories

As we could see in chapter 3 (see section 3.1.2), estimation of KS-entropy of tent map time series exhibiting long-term memory effects (e.g. for $\rho < 1.5$ see Fig. 3.1) required high-order estimators of the entropy rate (at least up to the eighth order). It was then also demonstrated that for such time series the LZ-based estimator of the entropy rate provides a more accurate approximation of KS-entropy. In this chapter we investigate the problem of inference of the directionality of interaction from time series exhibiting long-term memory effects. In the following section we address the question how and to what extent the order k of the symbolic transfer entropy (see Eq. 4.5) can influence the correct inference of the directionality of interaction between two interacting Rössler oscillators. Next, by extending the concept of joined LZ-complexity proposed in [ZRB05] we first introduce the notion of conditional LZ-complexity and then by following [Sch00] the notion of algorithmic or LZ-based transfer entropy. Finally, by estimating entropy transfer between model dynamical systems with the symbolic transfer entropy as well as the LZ-based transfer entropy we investigate the relationship between these measures.

6.1. Directional interactions between Rössler oscillators

In contrast to the broad-band power spectrum of a Lorenz oscillator (Eq. A.3) the power spectrum of a Rössler oscillator (Eq. A.4) has a single dominating frequency component. This manifests itself in an almost oscillating autocorrelation function and thus long-term memory effects in time series (see Fig. 4.1).

In this section we investigate the problem of inference of the directionality of interaction between a pair of unidirectionally coupled Rössler oscillators with different characteristic frequencies $\omega_1 = 0.8$, $\omega_2 = 1$ (Eq. A.9; $c_{21} = 0$). With the used coupling scheme the slower first oscillator (driver) was diffusively coupled into the faster second oscillator (responder) with coupling strength c_{12} (see Fig. 6.3). The differential equations of Rössler oscillators were integrated using a fourth order Runge-Kutta algorithm with integration step $dt = 0.005$ and then downsampled to $d\tilde{t} = 0.2$. The initial conditions were normally distributed with zero mean and unit variance. In order to eliminate transients, the first 10^4 iterations were discarded. The $x-$ components of the driver and responder of length $N = 10^5$ were transformed into two series of permutation symbols S_1 and S_2. In order to reduce the influence of the finite sample effect on the estimate of the symbolic transfer entropy we

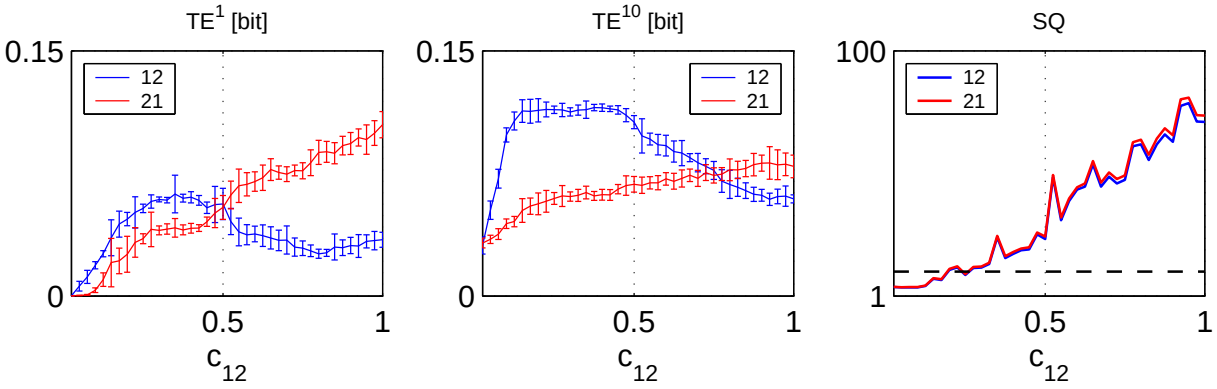


Figure 6.1.: Symbolic transfer entropies TE^1 of first (left) and tenth TE^{10} (middle) orders between $x-$ components of two unidirectionally coupled Rössler oscillators for increasing coupling strength c_{12}. Embedding parameters were set to $m = 3$ and $\tau = 16$. $x-$ components of both oscillators were transformed into series of permutation symbols. Each point on first two plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations. Statistical quotients SQ (right) of symbolic transfer entropies TE^{10} for increasing coupling strength c_{12}. Dashed line indicates $SQ = 10$.

followed the argumentation made in section 4.2.2 and set the embedding dimension to $m = 3$. To choose an appropriate delay time τ we again followed [SL08] and set the embedding window such as to approximately cover a basic period T of a Rössler oscillator, i.e., $(m-1)\tau \approx T$. Thus, for $m = 3$ we set $\tau \approx T/2$. With the used sampling rate $d\tilde{t} = 0.2$ the basic period of the Rössler oscillator corresponds to $T_{Roes} \approx 32$ integration steps (or, in absolute units as $T_{Roes}d\tilde{t} \approx 6$, see Figs 4.1 and 6.3).

Figure 6.1 shows the first order symbolic transfer entropies $TE^1(S_1, S_2)$ and $TE^1(S_2, S_1)$ between two unidirectionally coupled Rössler oscillators for the increasing coupling strength c_{12}. Similar to our results obtained in section 4.2.2 (where we investigated the entropy transfer between coupled Lorenz oscillators, see Fig. 4.7) the values of $TE^1(S_1, S_2)$ and $TE^1(S_2, S_1)$ were increasing with the coupling strength c_{12} indicating the existence of entropy transfer between the oscillators. For weak and intermediate coupling strengths ($c_{12} < 0.5$) the direction of interaction "driver-to-responder" was correctly identified, i.e., $TE^1(S_1, S_2) > TE^1(S_2, S_1)$. However, already for the case $c_{12} > 0.5$ we obtained the opposite situation,i.e., $TE^1(S_1, S_2) < TE^1(S_2, S_1)$ and thus an incorrect inference of the directionality of interaction between oscillators. To explain this counterintuitive phenomenon we need to recall that for coupling strengths $c_{12} > 0.5$ Rössler oscillators enter the regime of lagged synchronization (see Fig. 6.3). Due to the difference in the characteristic frequencies between first (driver) and second (responder) oscillators ($\omega_1 = 0.8$ and $\omega_2 = 1$) we observed that the more slower driver followed – with a time lag of several integration steps – the more faster responder[1]. This led to higher values of the first order symbolic

[1]Similar observations were made in [OMWL07] where the authors applied a phase modeling approach to infer the directionality of interaction between two unidirectionally coupled Rössler oscillators. It was

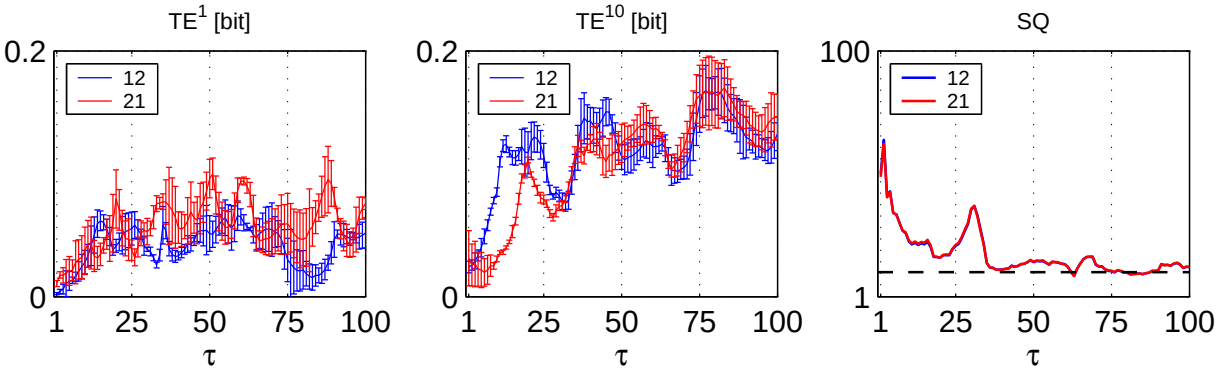


Figure 6.2.: Symbolic transfer entropies TE^1 of first (left) and tenth TE^{10} (middle) orders between $x-$ components of two unidirectionally coupled Rössler oscillators for the fixed coupling strength $c_{12} = 0.35$ and different delay time $\tau \in [1, 100]$. Embedding dimension was set to $m = 3$. Each point on first two plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations. Statistical quotients SQ (right) of symbolic transfer entropies TE^{10} for all value of delay time τ. Dashed line indicates $SQ = 10$.

transfer entropy in the direction "responder-to-driver" than in the opposite direction (i.e., $TE^1(S_1, S_2) < TE^1(S_2, S_1)$) and thus to an incorrect inference of the directionality of interaction. By further increasing the coupling strength up to $c_{12} = 1$ we observed that both oscillators got almost completely synchronized and the slow driver followed the faster responder with a small time lag of only few integration steps (see Fig. 6.3). This observation explains the significantly higher values of the first order symbolic transfer entropy in the direction "responder-to-driver" than values obtained in the opposite direction, i.e., $TE^1(S_2, S_1) \approx 3TE^1(S_1, S_2)$ when $c_{12} \approx 1$.

As we could see in chapter 3 the entropy rate estimators of high order were required for a more accurate estimation of KS-entropy of the tent map, especially in cases of tent map time series exhibiting slow decaying autocorrelation functions (e.g. for $\rho < 1.5$ see Fig. 3.1). Since the Rössler time series exhibits an oscillating autocorrelation function (see Fig. 4.1) we repeated our analysis of entropy transfer between coupled Rössler oscillators but now by using the tenth order[2] symbolic transfer entropies $TE^{10}(S_1, S_2)$ and $TE^{10}(S_2, S_1)$. Our findings revealed that, in comparison to the first order symbolic transfer entropy, estimating entropy transfer with the tenth order symbolic transfer entropies allowed us to correctly infer the direction of interaction between oscillators (i.e., $TE^{10}(S_1, S_2) > TE^{10}(S_2, S_1)$) for coupling strengths up to $c_{12} \approx 0.75$ and thus to extend the range of coupling strengths where the correct inference was still possible from $c_{12} \in [0, 0.5]$ to $c_{12} \in [0, 0.75]$ (Fig. 6.1). Analysis of entropy transfer for $c_{12} > 0.75$ revealed that for these coupling strengths $TE^{10}(S_2, S_1)$ was already slightly larger than $TE^{10}(S_1, S_2)$ indicating an incorrect inference of the di-

observed that a mismatch in characteristic frequencies of Rössler oscillators can have a strong influence on inference of the directionality of interaction between them.

[2]Analysis of entropy transfer of more higher orders ($k > 10$) were limited due to exponentially growing demands of computational resources.

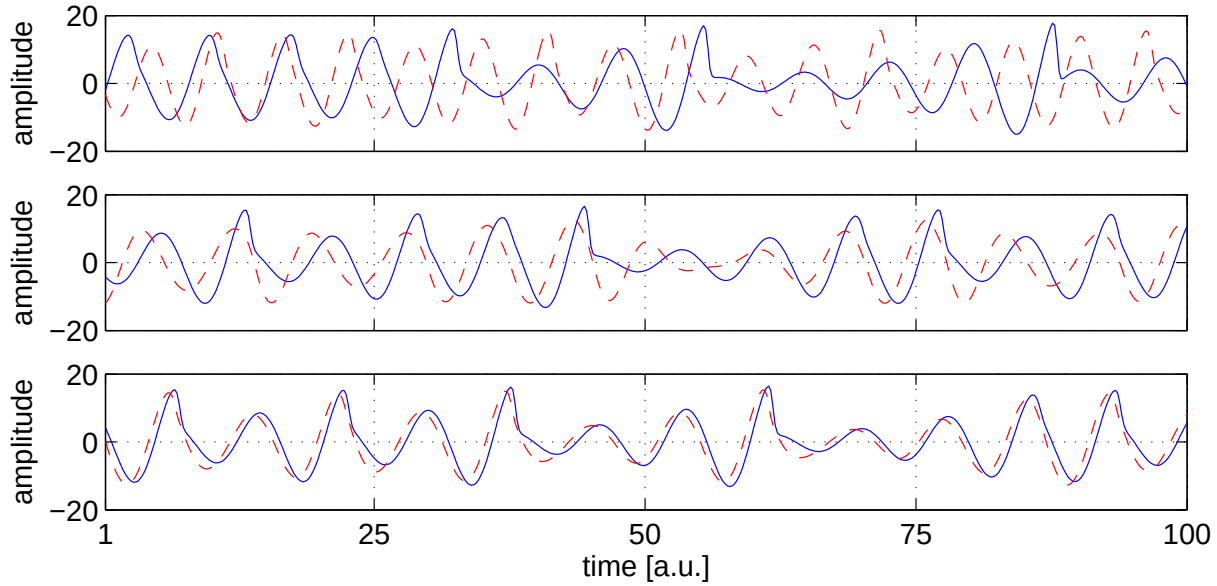


Figure 6.3.: Exemplary time series of $x-$ components of unidirectionally coupled Rössler oscillators for weak $c_{12} = 0.1$ (upper row), intermediate $c_{12} = 0.5$ (lower row), and strong $c_{12} = 1$ (third raw) coupling strengths (Eq. A.9; $\omega_1 = 1$, $\omega_2 = 0.8$ and $c_{21} = 0$). Solid and dashed lines indicate driver and responder time series.

rectionality of interaction between oscillators. Thus, the inference of the directionality of interaction between strongly coupled Rössler oscillators (i.e., when c_{12} approaches 1) may require higher order ($k > 10$) estimates of transfer entropy. However, as was already stated in [Sch00] the estimation of high order transfer entropies (Eq. 4.5) can be limited due to either computational reasons or fast growing systematic and statistical errors. As we mentioned in section 4.2.1 the influence of such errors can be neglected if the statistical quotient of an entropy-based measure fulfills the condition $SQ > 10$. According to the definition of the statistical quotient , under this condition the effect of undersampling of empirical probability distributions in many cases can be neglected (see [Kre99] and references therein). With the used embedding dimension $m = 3$ and time series length $N = 10^5$ the statistical quotient for the first order ($k = 1$) symbolic transfer entropy (see Eq. 4.6) can be estimated as $SQ = N/(m!)^{2k+1} \approx 463 >> 10$. The estimation of the statistical quotient for the tenth order ($k = 10$) symbolic transfer entropy reveals $SQ = N/W = N/(m!)^{2k+1} \approx 5^{-12}$ where W denotes a possible number of words of length $2k+1$ which can occur in symbolic strings S_1 and S_2 of length N. However, the number of different words which were found in symbolic strings S_1 and S_2, let us denote it by W^*, is considerably smaller than W and cannot exceed N. To answer the question whether the obtained dependencies of the tenth order symbolic transfer entropies $TE^{10}(S_1, S_2)$ and $TE^{10}(S_2, S_1)$ suffer from a finite sample effect due to finite length N we additionally estimated the statistical quotient $SQ^* = N/ < W^* >$, where $< W^* >$ denotes an average (over 20 realization) number of different words of length $2k+1$ which were found in symbolic strings S_1 and S_2. Our analysis of SQ^* revealed that for $c < 0.2$ obtained values of $TE^{10}(S_1, S_2)$ and $TE^{10}(S_2, S_1)$ might indeed suffer from system-

atic errors (biases) ($SQ^* < 10$, see Fig. 6.1). This explains the observed positive biases of $TE^{10}(S_1, S_2)$ and $TE^{10}(S_2, S_1)$ which were measured between uncoupled Rössler oscillators ($c_{12} = 0$). We also observed that the value of statistical quotient SQ^* was growing with increasing coupling strength c_{12}. To explain this finding we need to recall that for higher values of c_{12} the oscillators were getting more synchronized and this led to a decreasing number of different words W^*.

As we could see in section 4.2.2 the incorrect construction of the permutation partition (by selection of certain values of delay time τ) led to the incorrect inference of the directionality between a pair of interacting Lorenz (Fig. 4.7) as well as between Rössler and Lorenz (Fig. 4.8) oscillators. As it was demonstrated there, the obtained values of the symbolic transfer entropies between these oscillators, in both cases, depended on the delay time τ. For some τ the symbolic transfer entropy in the direction "responder-to-driver" exceeded symbolic transfer entropy in the direction "driver-to-responder". This led to the incorrect inference of the directionality of interaction between oscillators. Following discussions presented in [EST01] we assume here that an incorrect construction of partition (i.e., when our partition significantly deviates from the generating partition of the underlying system) may result in long-term correlations (memories) in the symbolic representation of real-valued data. To test this assumption we estimated the entropy transfer between unidirectionally coupled Rössler oscillators but now with a fixed coupling strength $c_{12} = 5$ and different values of the delay time $\tau \in [1, 100]$ using the first and tenth order symbolic transfer entropies (see Fig. 6.2). Similar to the results obtained in section 4.2.2 our findings demonstrated that both estimates of entropy transfer depend on τ. Moreover, a careful selection of delay time τ was needed in order to correctly infer the directionality of interaction between Rössler oscillators. For instance, for the first order symbolic transfer entropy we found that $TE^1(S_1, S_2) > TE^1(S_2, S_1)$ for delay time $\tau = T_{Lor}/2 = 16$. With this τ the embedding window approximately covered the basic period of Rössler oscillators T_{Roes}, i.e., $\tau(m-1) \approx T_{Roes}$. For all other values of τ (except $\tau = 26$) the first order symbolic transfer entropy in the direction "responder-to-driver" exceeded the first order symbolic transfer entropy in the opposite (i.e., "driver-to-responder") direction. Analysis of entropy transfer by means of the tenth order symbolic transfer entropy revealed that such estimates of entropy transfer between oscillators were correctly indicating the directionality of interaction (i.e., $TE^{10}(S_1, S_2) > TE^{10}(S_2, S_1)$) for all $\tau \in (5, 32)$ as well as $\tau \approx 45$, whereas for all other values of τ there were no significant differences between $TE^{10}(S_1, S_2)$ and $TE^{10}(S_2, S_1)$. To test whether obtained values of the tenth order symbolic transfer entropy suffer from a finite sample effect we computed the statistical quotient SQ^* for all values of delay time τ (see Fig. 6.2). We found that $SQ^* > 10$ almost for all τ. However, for $\tau > 32$ the statistical quotient was $SQ^* \approx 10$ indicating that a finite sample effect can already affect obtained values of $TE^{10}(S_1, S_2)$ and $TE^{10}(S_2, S_1)$.

The findings obtained in this section demonstrated that high order estimates of transfer entropy, in general, exhibited less dependency on variation of permutation partition. (see Fig. 6.2). This property of the high order symbolic transfer entropy can be useful in analyzing time series with unknown or non-stationary dynamics and especially in situations where the form (construction) of the partition has to be optimized in the course of time. Additionally, we observed that by applying high order estimates of entropy transfer between

strongly coupled Rössler oscillators allowed us to extend the range of coupling strengths where the correct inference of the directionality of interaction was still possible. However, our findings also indicated that the estimation of the high orders symbolic transfer entropy, in general, required a large amount of data (in our case we set $N = 10^5$). Moreover, obtained estimates of high orders transfer entropy may suffer (e.g. when $SQ < 10$ as for $c_{12} < 0.2$ see Fig. 6.1) from systematic and statistical errors due to the finite length of time series. In the following section we will introduce a complementary approach to estimate high order entropy transfer by extending the notion of mutual LZ-complexity proposed in [ZRB05]. Based on the concept of LZ-complexity this approach does not require the reconstruction of high dimensional empirical probability functions and can be directly applied for the analysis of high orders transfer entropy in time series of, in general, arbitrary lengths.

6.2. Conditional LZ-complexity and algorithmic transfer entropy

Since the seminal work of Shanon [Sha48] it is known that the compressibility of data is closely related to its entropy rate. Following practical and theoretical investigations of this relationship revealed a variety of approaches which were aimed to approximate the entropy rate of time series by compressing it with some compression algorithm (see [CT91, SG96] and references therein). One of the widely used approaches was proposed by Lempel and Ziv and is based on the notion of Lempel-Ziv (LZ) complexity C_{LZ} for a discrete sequence of data [LZ76, ZL77]. Later studies revealed that the notion of LZ-complexity $C_{LZ}(S)$ of some string S is closely related to the notion of entropy rate $dH(Z)$ of the underlying stochastic process Z (for more details see section 2.4.1). In a recent study [ZRB05] the authors investigated a possible extension of the notion of LZ-complexity to characterize complexity and interactions in multivariate data. In the special case of bivariate data when only two strings, let us say, S_1 and S_2 are considered the authors introduced a joined LZ-complexity $C_{LZ}(S_1, S_2)$ and it was shown that this quantity is related to the joint entropy rate $dH(Z_1, Z_2)$ of the underlying bivariate stochastic process $Z_b \equiv (Z_1, Z_2)$. Following the analogy with the notion of two-point mutual information between two random variables (Eq. A.21) the authors proposed the notion of mutual LZ-complexity $dI_{LZ}(S_1, S_2; N)$ which characterizes an extent of correlations between two finite strings S_1 and S_2 of length N. As it was shown in section 2.4.1 the mutual LZ-complexity between S_1 and S_2 can be related asymptotically ($N \to \infty$) to the sum of transfer entropies $TE(Z_1, Z_2)$ and $TE(Z_2, Z_1)$ (asymmetric parts of interaction) and mutual information rate $dI(Z_1, Z_2)$ (symmetric part of interaction) between the underlying stochastic processes Z_1 and Z_2 as

$$\lim_{N\to\infty} dI_{LZ}(S_1, S_2; N) = TE(Z_1, Z_2) + TE(Z_2, Z_1) + dI(Z_1, Z_2). \tag{6.1}$$

Let us now consider that our strings S_1 and S_2 exhibit a symbolic representation of two interacting dynamical systems, e.g. series of permutation symbols as used in the previous sections. In this case, mutual LZ-complexity between these strings can be used to characterize the degree (strength) of interaction between dynamical systems. However, from

the definition of mutual LZ-complexity (Eq. 2.27) it follows that this measure of interdependence between strings S_1 and S_2 is, by definition, symmetric and does not allow to reveal the directionality of interaction[3], i.e., $dI_{LZ}(S_1, S_2; N) = dI_{LZ}(S_2, S_1; N)$. In order to characterize the directionality of interaction between dynamical systems one has to provide an estimate of only those terms in relation 6.1 which contribute to the asymmetric part of interaction, i.e., to provide LZ-based estimates of transfer entropies in both directions.
To proceed, let us again consider a binary string $S \equiv \{s_i\}_{i=1}^{N} = 011010100\cdots$ which is a realization of a stationary and ergodic stochastic process Z. By using the LZ-compression algorithm presented in section 2.4.1 we parse this string into a set of distinct non-overlapping words: $\{s_i\}_{i=1}^{N} = (0)(1)(10)(10100)\cdots$. The obtained number of distinct words defines the LZ-complexity $C_{LZ}(S)$ of the string S. By definition of the LZ-algorithm each new word w_k has no match with the preceding words or even with any preceding substring, i.e., $w_k = \{s_i\}_{i=j}^{t} \notin \{s_i\}_{i=1}^{t-1}$ $(j \leq t)$. This guarantees that within each new word some amount of new information is generated. For instance, the 4^{th} word has no match within the preceding substring: $w_4 = \{s_i\}_{i=5}^{9} = (10100) \notin \{s_i\}_{i=1}^{8} = 01101010$. However, it is important to point out that this is no longer valid if the last bit of the word is switched, i.e., $\hat{w}_4 = (10101) \in \{s_i\}_{i=1}^{8} = 01101010$. This statement indicates that the last bit of each new word contains new information and it is true for any other word w_k which is obtained by parsing the string $\{s_i\}_{i=1}^{N}$ with the LZ-algorithm.
Now, let us consider two binary strings $S_1 \equiv \{s_i^1\}_{i=1}^{N} = 011010100\cdots$ and $S_2 \equiv \{s_i^2\}_{i=1}^{N} = 110010011\cdots$ which both are realizations of stationary and ergodic stochastic processes Z_1 and Z_2 along with the joint string $\{s_i^{12}\}_{i=1}^{N} = 231030122\cdots$ (same as in section 2.4.2) whose symbols are taken from the joined alphabet $s_i^{12} = s_i^1 \cdot 1 + s_i^2 \cdot 2$. Let us parse the joint string $\{s_i^{12}\}_{i=1}^{N}$ into a set of distinct non-overlapping words according to the LZ-algorithm: $\{s_i^{12}\}_{i=1}^{N} = \binom{0}{1}\binom{1}{1}\binom{1}{0}\binom{0}{0}\binom{10}{10}\binom{10}{01}\binom{0}{1}\cdots = (2)(3)(1)(0)(30)(12)(2)\cdots$. The obtained number of distinct words defines the joint LZ-complexity $C_{LZ}(S_1, S_2)$ of the strings S_1 and S_2. In a similar way as we already observed above (i.e., in the case of parsing the string S) each new word w_k has no match within the preceding words or even within any preceding substring, i.e., $w_k = \{s_i^{12}\}_{i=j}^{t} \notin \{s_i^{12}\}_{i=1}^{t-1}$ $(j \leq t)$. This implies that within each new word some amount of new information (entropy) is generated. For instance, the 5^{th} word $w_5 = \{s_i^{12}\}_{i=5}^{6} = (30)$ has no match with the preceding substring $\{s_i^{12}\}_{i=1}^{5} = 23103$. However, in contrast to the univariate case considered above, here, the new information (entropy) can be generated in either the S_1- (upper bit) or/and the S_2- (lower bit) component of the word. In our example we can test this by switching the last bit of either the S_1- component (upper bit) $\hat{w}_5^{S_1} = \binom{11}{10} = (31) \in \{s_i^{12}\}_{i=1}^{5} = 23103$ or the S_2- component (lower bit) $\hat{w}_5^{S_2} = \binom{10}{11} = (32) \notin \{s_i^{12}\}_{i=1}^{5} = 23103$ respectively. This allows us to identify which component of the word (upper or/and lower bit) contains new information. In the presented example the new information is generated in the S_1- (upper bit) component of the word w_5. Let us now define the *conditional LZ-complexity* $C_{LZ}(S_1|S_2)$ of string S_1 relative to string S_2 as the total number of words obtained by parsing the joint string $\{s_i^{12}\}_{i=1}^{N}$ excluding those words which are supposed to generate new information in the S_2- component, i.e., we do not count

[3]This follows from the fact that the joint LZ-complexity between two strings S_1 and S_2 is, by definition, symmetric, i.e, $C_{LZ}(S_1, S_2) = C_{LZ}(S_2, S_1)$.

those words w_k for which the following condition is fulfilled: $w_k = \{s_i^{12}\}_{i=j}^{t} \notin \{s_i^{12}\}_{i=1}^{t-1}$ and at the same time $\hat{w}_k^{S_2} = \left\{s_i^{12}, \binom{s_m^1}{\tilde{s}_m^2}\right\}_{i=j}^{t-1} \in \{s_i^{12}\}_{i=1}^{t-1}$ ($j \leq t$; $\tilde{s}$ denotes inversion of s). Defined in such a way, conditional LZ-complexity $C_{LZ}(S_1|S_2)$ characterizes the complexity of the string S_1 when the conditioning on the string S_2 is taken into account. If we now recall our assumption that strings S_1 and S_2 are realizations of two interacting stochastic processes Z_1 and Z_2 then we can expect that – by analogy to the usual LZ-complexity – the conditional LZ-complexity as is related to the conditional entropy rate of stochastic process Z_1 relative to the stochastic process Z_2 as

$$dH(Z_1|Z_2) \approx \frac{\log(N)}{N} C_{LZ}(S_1|S_2). \tag{6.2}$$

According to its definition (see section A.3 and Fig. A.2) the conditional entropy rate $dH(Z_1|Z_2)$ characterizes the reduction of the entropy rate $dH(Z_1)$ of the stochastic process Z_1 when an additional conditioning on the past of the stochastic process Z_2 is taking place. Thus, the difference between $dH(Z_1)$ and $dH(Z_1|Z_2)$ characterizes the amount of entropy that is transfered from process Z_2 to Z_1 per unit of time, i.e., $TE(Z_2, Z_1) = dH(Z_1) - dH(Z_1|Z_2)$. This gives us the definition of the transfer entropy from Z_2 to Z_1 introduced in sections 2.2.1 and A.3. In a similar way one defines the transfer entropy in the opposite direction. Following these definitions and using Eqs. 2.24 and 6.2 we define an *algorithmic- or LZ-based transfer entropy* from string S_2 to S_1 as

$$TE_{LZ}(S_2, S_1) = \frac{\log(N)}{N}(C_{LZ}(S_1) - C_{LZ}(S_1|S_2)) \tag{6.3}$$

and analogously from string S_1 to S_2 as

$$TE_{LZ}(S_1, S_2) = \frac{\log(N)}{N}(C_{LZ}(S_2) - C_{LZ}(S_2|S_1)). \tag{6.4}$$

According to the definition given in [Sch00], the transfer entropy $TE(Z_1, Z_2)$ provides a measure for entropy transfer from stochastic processes Z_1 to Z_2 (i.e., a measure for the reduction of entropy rate of process Z_2 by additional conditioning on the past of process Z_1) whereas the algorithmic transfer entropy $TE_{LZ}(S_1, S_2)$, as defined here, can be interpreted as a measure for reduction of LZ-complexity of string S_2 when the conditioning on string S_1 is taken into account. Being, by definition, asymmetric, the algorithmic transfer entropy $TE_{LZ}(S_1, S_2)$ can provide a complementary tool to reveal the asymmetry of interaction between dynamical systems when the symbolic data (i.e., symbolic representation of system observables) exhibit a high degree of correlations.

In the following section we will provide a numerical analysis for the asymptotic convergence of the algorithmic transfer entropy between two strings to the transfer entropy (proposed in [Sch00]) between the underlying stochastic processes. To do this we will use a pair of interdependent binary Markov chains which allow us to generate two symbol series with an analytically defined value of transfer entropy between them. Next, we will compute the symbolic transfer entropy of varying orders along with the algorithmic transfer entropy between two interacting tent maps and compare obtained values with each other. Finally, we

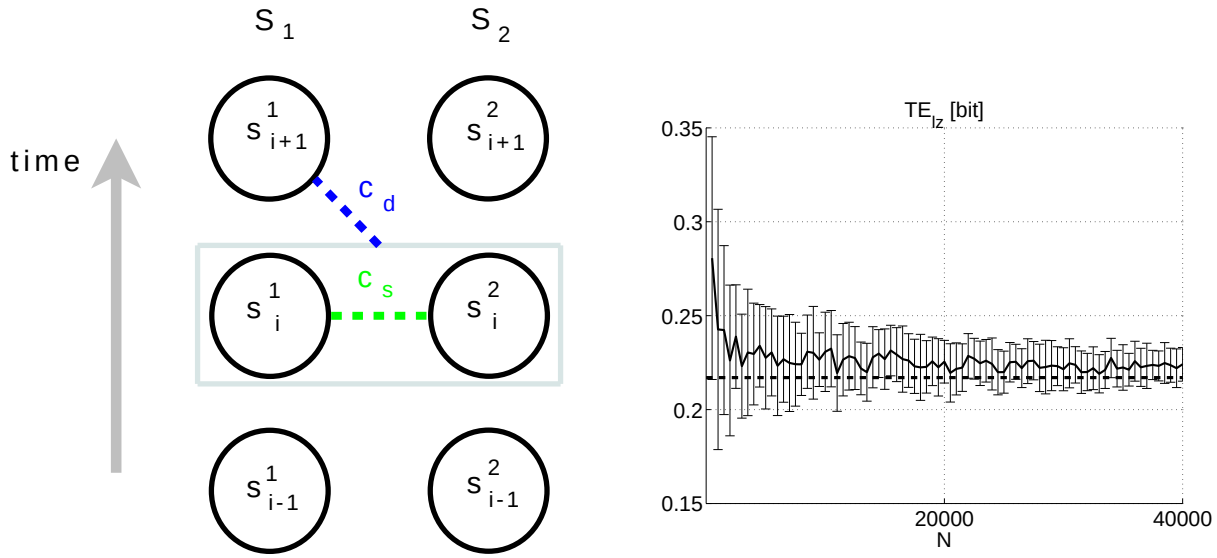


Figure 6.4.: Left: Schematic representation of two interdependent binary Markov chains S_1 and S_2. The static and dynamic couplings c_s and c_d are used to set a degree of static and dynamic correlations between Markov chains. Right: Asymptotic behavior of LZ-based transfer entropy TE_{LZ} from binary Markov chain S_2 to S_1 for increasing length N. Dashed line indicates the analytically given value of transfer entropy TE from S_2 to S_1. Each point on this plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

$P(s_i^2 \mid s_i^1)$		
	$s_i^1 = 0$	$s_i^1 = 1$
$s_i^2 = 0$	$0.5 + c_s/2$	$0.5 - c_s/2$
$s_i^2 = 1$	$0.5 - c_s/2$	$0.5 + c_s/2$

Table 6.1.: Numerical values of transition probability function $P(s_i^2|s_i^1)$ from current state s_i^1 to s_i^2 of binary Markov chains depicted in Fig. 6.4.

will consider a more complicated example of two interacting chaotic oscillators to demonstrate main similarities and differences between these two approaches to measure entropy transfer.

6.3. Estimating algorithmic transfer entropy between model dynamical systems

Two interdependent binary Markov chains

To address the question to what extent the algorithmic or LZ-based transfer entropy introduced above is related to transfer entropy proposed in [Sch00] we consider here a simple discrete time dynamical system: two interdependent binary Markov chains (BMC):

$P(s^1_{i+1}\|s^1_i, s^2_i)$				
	$(s^1_i, s^2_i) = (0,0)$	$(s^1_i, s^2_i) = (1,0)$	$(s^1_i, s^2_i) = (0,1)$	$(s^1_i, s^2_i) = (1,1)$
$s^1_{i+1} = 0$	$0.5 + (1 + c_d)/4$	0.5	0.5	$0.5 - (1 + c_d)/4$
$s^1_{i+2} = 1$	$0.5 - (1 + c_d)/4$	0.5	0.5	$0.5 + (1 + c_d)/4$

Table 6.2.: Numerical values of transition probability function $P(s^1_{i+1}|s^1_i, s^2_i)$ from current states (s^1_i,s^2_i) of two binary Markov chains to future state (s^1_{i+1}) of first binary Markov chain depicted in Fig. 6.4.

$S_1 = \{s^1_i\}^N_{i=1}$ and $S_2 = \{s^2_i\}^N_{i=1}$ with an analytically given value of first order transfer entropy $TE(S_2, S_1)$. We generate different realizations of these BMC of increasing length N and then investigate the asymptotic relationship between transfer entropy $TE(S_2, S_1)$ and its LZ-based version $TE_{LZ}(S_2, S_1)$ (Eq. 6.3).

Let us describe the dynamics of these BMC by means of the probability function $P(s^1_i)$ along with two transition probability functions $P(s^2_i|s^1_i)$ and $P(s^1_{i+1}|s^1_i, s^2_i)$ (see Fig. 6.4; tables 6.1 and 6.2). The probability function $P(s^1_i)$ (for time step i) is defined as $P(s^1_i = 1) = P(s^1_i = 0) = 0.5$. The transition probability function $P(s^2_i|s^1_i)$ (for time step i) describes the degree of static correlations between S_1 and S_2 and was parameterized with the *static coupling* $c_s \in [0, 1]$. As we can see from table 6.1, by setting $c_s = 0$ all transition probabilities between current states of the first and second BMC are equal and thus s^1_i and s^2_i are independent from each other. By setting $c_s = 1$, one has an opposite situation where $s^1_i = s^2_i$. The transition probability function $P(s^1_{i+1}|s^1_i, s^2_i)$ (for time step i) describes the temporal evolution of the process S_1 and was parametrized with the *dynamic coupling* $c_d \in [0, 1]$. It can be seen from table 6.2 that with $c_d = 0$ all transition probabilities $P(s^1_{i+1}|s^1_i, s^2_i)$ are equal to each other and the future state (s^1_{i+1}) of the process S_1 is thus independent from the current states (s^1_i and s^2_i) of processes S_1 and S_2. In the opposite situation when $c_d = 1$, there is a strong dependency between s^1_{i+1} and (s^1_i, s^2_i). Indeed – according to table 6.2 – in this case the state $(s^1_i, s^2_i) = (0, 0)$ will (with probability one) lead to the state $s^1_{i+1} = 0$. The same holds for the state $(s^1_i, s^2_i) = (1, 1)$ which leads to $s^1_{i+1} = 1$. According to the definition of transition probability function $P(s^1_{i+1}|s^1_i, s^2_i)$ the future state s^1_{i+1} of the Markov chain S_1 is independent from the past states (s^1_j,s^2_j; $j < i$) of both Markov chains S_1 and S_2. This implies that our binary symbol series $\{s^1_i\}^N_{i=1}$ and $\{s^2_i\}^N_{i=1}$ do not exhibit long-term memory effects and therefore $TE^k(S_2, S_1) = TE^1(S_2, S_1)$ for all $k > 1$.

The definitions of $P(s^1_i)$, $P(s^2_i|s^1_i)$, and $P(s^1_{i+1}|s^1_i, s^2_i)$ given above allow us to directly obtain $P(s^2_i, s^1_i) = P(s^2_i|s^1_i)P(s^1_i)$ and $P(s^1_{i+1}, s^1_i, s^2_i) = P(s^1_{i+1}|s^1_i, s^2_i)P(s^2_i, s^1_i)$ and then to analytically compute the transfer entropy of first order ($k = 1$) ([Sch00]) from Markov chain S_2 to S_1 as

$$
\begin{aligned}
TE(S_2, S_1) &= - \sum_{\left\{s_{i+1}^1, s_i^1, s_i^2\right\}} P(s_{i+1}^1, s_i^1, s_i^2) \log \frac{P(s_{i+1}^1 | s_i^1, s_i^2)}{P(s_{i+1}^1 | s_i^1)} = \\
&= - \sum_{\left\{s_{i+1}^1, s_i^1, s_i^2\right\}} P(s_{i+1}^1, s_i^1, s_i^2) \log \frac{P(s_{i+1}^1 | s_i^1, s_i^2) P(s_i^1)}{\sum_{\left\{s_i^2\right\}} P(s_{i+1}^1, s_i^1, s_i^2)}.
\end{aligned}
\tag{6.5}
$$

As a first step of our analysis, by randomly taking (with probability $P(s_i^1)$) a binary initial condition s_i^1 $(i = 1)$ and using the transition probability functions $P(s_i^2|s_i^1)$ and $P(s_{i+1}^1|s_i^1, s_i^2)$ defined above, we generated[4] 20 realizations (pairs of binary symbol series $\{s_i^1\}_{i=1}^N$ and $\{s_i^2\}_{i=1}^N$) of BMC S_1 and S_2 of length N. The dynamic and static couplings were set to $c_s = 0.5$ and $c_d = 0.9$. This allowed us to analytically define the transfer entropy from S_2 to S_1 to $TE(S_2, S_1) \approx 0.218$ (see Eq. 6.5; tables 6.1 and 6.2). In order to investigate the asymptotic behavior of the LZ-based transfer entropy $TE_{LZ}(S_2, S_1)$ and its relationship to $TE(S_2, S_1)$ we computed the LZ-based transfer entropy for increasing length of binary symbol series from $N = 500$ to $N = 40000$ with a step $\Delta N = 500$. Our findings indicated that the LZ-based transfer entropy between two interdependent binary Markov chains S_1 and S_2 indeed converges (from above) to the transfer entropy of first order introduced in [Sch00] (Fig. 6.4). In order to complement findings presented in this section, in the next section we will investigate the entropy transfer between symbol series exhibiting long-term memory effects. To to this we will consider a binary symbolic representation of two interacting chaotic maps.

Two interacting tent maps

We continue our analysis by considering a system of two unidirectionally coupled tent maps. With the used coupling scheme (Eq. A.5; for more details see [Sch00]) the first tent map (driver) is coupled into the second tent map (responder). Let $x_1(n)$ and $x_2(n)$ denote a pair of real-valued time series of length N generated by these tent maps. For both maps the control parameters were set to $\rho_j = 2$ $(j = 1, 2)$. With this choice of the control parameter the generating partition for each tent map is well known ([Str01]; see also section 3.1.2) and given by a binary partition $\Pi_{bin} = [\Pi_1; \Pi_2]$ with $\Pi_1 \equiv [0, 0.5]$ and $\Pi_1 \equiv (0.5, 1]$. It is known that the trajectory of the tent map is bounded to the range $[0, 1]$ and will, in the course of time, visit two elements of the partition Π_{bin} (Eq. A.1). To each of the two elements one can assign a symbol and the time evolution of the tent map is thus encoded as a sequence of binary symbols. Let B_n^j $(n = 1, \ldots, N)$ denote the symbolic representation of the first $(j = 1)$ and second $(j = 2)$ tent map. The series of binary symbols form two strings S_1 and S_2. In case where the tent maps are not coupled to each other the obtained symbolic representations of both maps are equivalent[5] to the original real-valued representations.

[4] In order to generate a realization of BMC we used an algorithm for the simulation of stochastic processes and fields introduced in [Hon02] (see section 5.5).

[5] Indeed, according to our finding obtained in section 3.1.2 for $\rho = 2$ the entropy rate estimator dH^k is equal to KS-entropy rate of the tent map already for $k = 1$. This indicates that symbol series S obtained

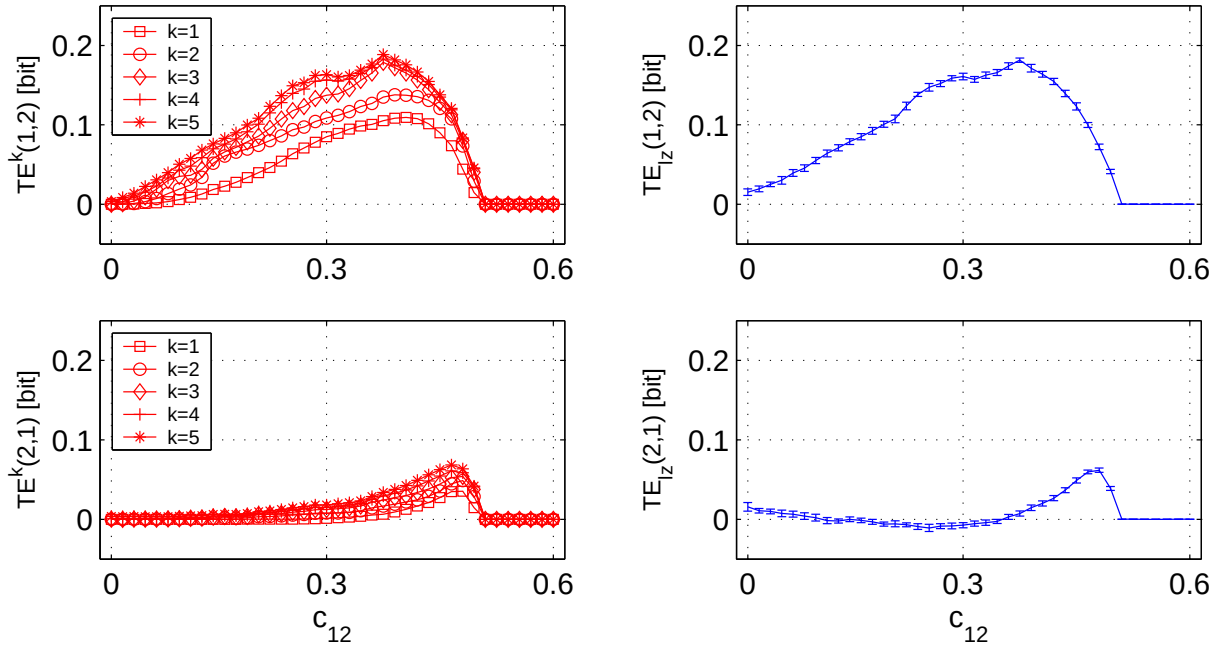


Figure 6.5.: Symbolic transfer entropies TE^k of order k and algorithmic (LZ-based) transfer entropies TE_{lz} between $x-$ components of two unidirectionally coupled tent maps for increasing coupling strength c_{12}. $x-$ components of both maps were first transformed into the symbolic representation: two series of binary symbols obtained with threshold-crossing partition. Each point on all plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

However, in case of non-zero coupling strength between these maps the used binary partition Π_{bin} may not necessarily be a generating partition. In this case we expect that our symbolic representation, i.e., our strings S_1 and S_2, may exhibit (long-term) memory effects. To test this assumption we estimated the symbolic transfer entropies of different orders $TE^k(S_1, S_2)$ and $TE^k(S_2, S_1)$ (with order $k = 1, \ldots, 5$) and compared obtained values with the estimates of algorithmic (LZ-based) transfer entropies $TE_{LZ}(S_1, S_2; N)$ and $TE_{LZ}(S_2, S_1; N)$ (see Fig. 6.5). To avoid the influence of a finite sample effect on obtained estimates of the symbolic transfer entropy we generated real-valued time series of length $N = 10^5$ [6]. In contrast to the analysis made in section 6.3, exact analytical values of entropy transfer between two interacting tent maps is not yet know (despite a considerable effort made e.g. in [LK05]). Fig. 6.5 shows a qualitatively similar behavior of symbolic transfer entropies $TE^k(S_1, S_2)$

with the binary partition Π_{bin} does not exhibit long-term memory effects and thus Π_{bin} represent a generating partition for a tent map.

[6] Indeed, the statistical quotient for the symbolic transfer entropy of the fifth order computed for a pair of binary symbol series can be estimated as $SQ = N/2^{2k+1} \approx 48$. This fulfills the condition $SQ \geq 10$ needed to reduce finite sample effects of the estimate of the symbolic transfer entropy to a satisfactory level. For further details see section 4.2.1

and $TE^k(S_2, S_1)$ as we observed for other model dynamical systems in section 4.2.1 [7]. Next, we observed that symbolic transfer entropies (in both directions) depended on the order k. The obtained values of the symbolic transfer entropies $TE^k(S_1, S_2)$ and $TE^k(S_2, S_1)$ increased for higher values of k. However, the estimates of transfer entropy obtained with the symbolic transfer entropies of fourth and fifth orders were already almost equal to each other. This converging behavior of the symbolic transfer entropy indicated the existence of memory (temporal correlations of, at least, forth order) in the binary symbol series S_1 and S_2 (which were obtained by partitioning of the original real-valued time series with Π_{bin}). The dependencies of entropy transfer on the coupling strength c_{12} between tent maps obtained with the LZ-based approach showed a qualitatively similar behavior as in the case when these dependencies were obtained with the symbolic transfer entropy of the fourth or fifth order. A similar converging behavior (with increasing order k) between Shannon and LZ-based approaches was observed in section 3.1.2 where we estimated KS-entropy of the tent map with LZ-based and Shannon entropy rate estimators of different orders. Complementing findings obtained in section 6.3 the findings obtained here indicated a close relationship between the notion of LZ-based transfer entropy and the notion of high order ($k > 1$, see Eq. 4.5) transfer entropy proposed in [Sch00].

Two interacting chaotic oscillators

We continue our analysis by considering a system of interacting chaotic oscillators. Following analysis made in sections 4.2.2 and 6.1 we considered a pair of unidirectionally coupled Lorenz (Eq. A.8; $R_{1,2} \in \mathbf{N}(28, 0.5)$) and Rössler oscillators (Eq. A.9; $\omega_1 = 1$, $\omega_2 = 0.8$ and $c_{21} = 0$). With the used coupling schemes the first oscillator (driver) was diffusively coupled into the second oscillator (responder) with the coupling strength $c_{12} > 0$. The differential equations of Lorenz and Rössler oscillators were integrated using a fourth order Runge-Kutta algorithm with integration step $dt = 0.005$ and then downsampled to $d\tilde{t} = 0.03$ (for Lorenz oscillators) and $d\tilde{t} = 0.2$ (for Rössler oscillators). The initial conditions were normally distributed with zero mean and unit variance. In order to eliminate transients, the first 10^4 iterations were discarded. The $x-$ components of the driver and responder were transformed into a pair of permutation symbols which formed two strings S_1 and S_2. Following considerations made in sections 4.2.2 and 6.1 both strings were formed by using the following embedding parameters: $m = 3$, $\tau = 10$ for Lorenz and $\tau = 16$ for Rössler oscillators correspondingly.

As a first step of the analysis we estimated entropy transfer between Lorenz (Rössler) oscillators with the symbolic transfer entropies $TE(S_1, S_2)$ and $TE(S_2, S_1)$ as well as with the LZ-based transfer entropies $TE_{LZ}(S_1, S_2)$ and $TE_{LZ}(S_2, S_1)$. The coupling strengths c_{12} between oscillators were increased from 0 to 10 ($\delta c_{12} = 0.25$) for a system of coupled Lorenz oscillators and from 0 to 1 ($\delta c_{12} = 0.025$) for a system of coupled Rössler oscilla-

[7] The value of $TE^k(S_1, S_2)$ is growing with the coupling strength c_{12} (Eq. A.5) reaching its maximum at $c_{12} \approx 0.38$ and then approaching zero for $c_{12} > 0.5$ for which the tent maps get fully synchronized. The positive values of the $TE^k(S_2, S_1)$ (despite the absence of coupling in this direction) exhibit static correlations between symbols which are increased when the tent maps get more synchronized (for further details see section 4.2.1).

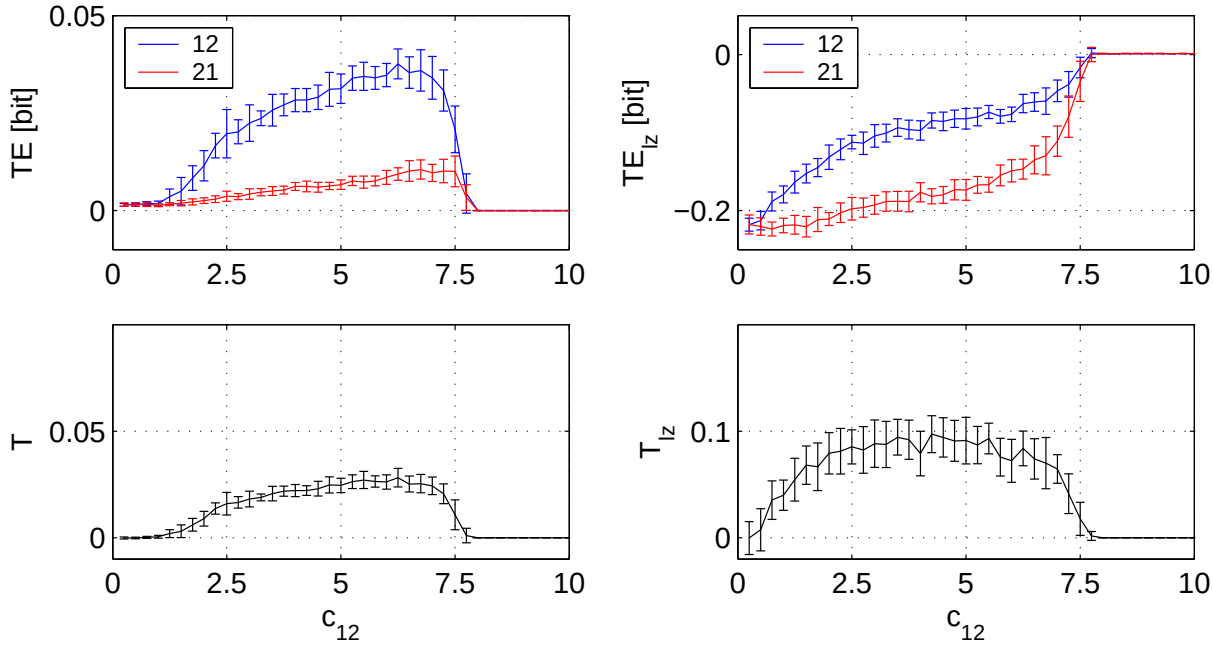


Figure 6.6.: Symbolic transfer entropies TE of the first order and corresponding asymmetry index T along with algorithmic (LZ-based) transfer entropies TE_{lz} and corresponding asymmetry index T_{lz} between $x-$ components of two unidirectionally coupled Lorenz oscillators for increasing coupling strength c_{12}. $x-$ components of both oscillators were first transformed into the symbolic representation: two series of permutation symbols (embedding dimension $m = 3$, delay time $\tau = 10$). Each point on all plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

tors. For each coupling strength c_{12} we again generated 20 realizations of real valued time series of length $N = 10^4$. As it can be seen from Figs. 6.6 and 6.7 our analysis revealed that for uncoupled Lorenz (Rössler) oscillators ($c_{12} = 0$) the obtained values of LZ-based transfer entropies were negative. To explain this phenomenon we need to recall the findings obtained in [ZRB05], where the authors demonstrated that the mutual LZ-complexity $dI_{LZ}(S_2, S_1; N)$ (Eqs. 2.26 and 2.27) can be negative, in contrast to two-point mutual information which is a positively defined function (Eqs. A.20 and A.21). The authors argued that this property of mutual LZ-complexity may limit the range of possible applications of this measure for the estimation of mutual correlation between time series and special care has to be taken for a reliable interpretation of results. The negativity of algorithmic transfer entropy $TE_{LZ}(S_1, S_2)$ indicates that – according to its definition (Eq. 6.4) – the LZ-complexity $C_{LZ}(S_1)$ of the string S_1 can be smaller then its conditional LZ-complexity $C_{LZ}(S_1|S_2)$. In contrast to this the entropy rate $dH(Z_1)$[8], by definition, can only be larger then the conditional entropy rate $dH(Z_1|Z_2)$. This implies that the Shannon estimator and

[8]Here we again assume that our strings S_1 and S_2 are realizations of two interacting stochastic processes Z_1 and Z_2.

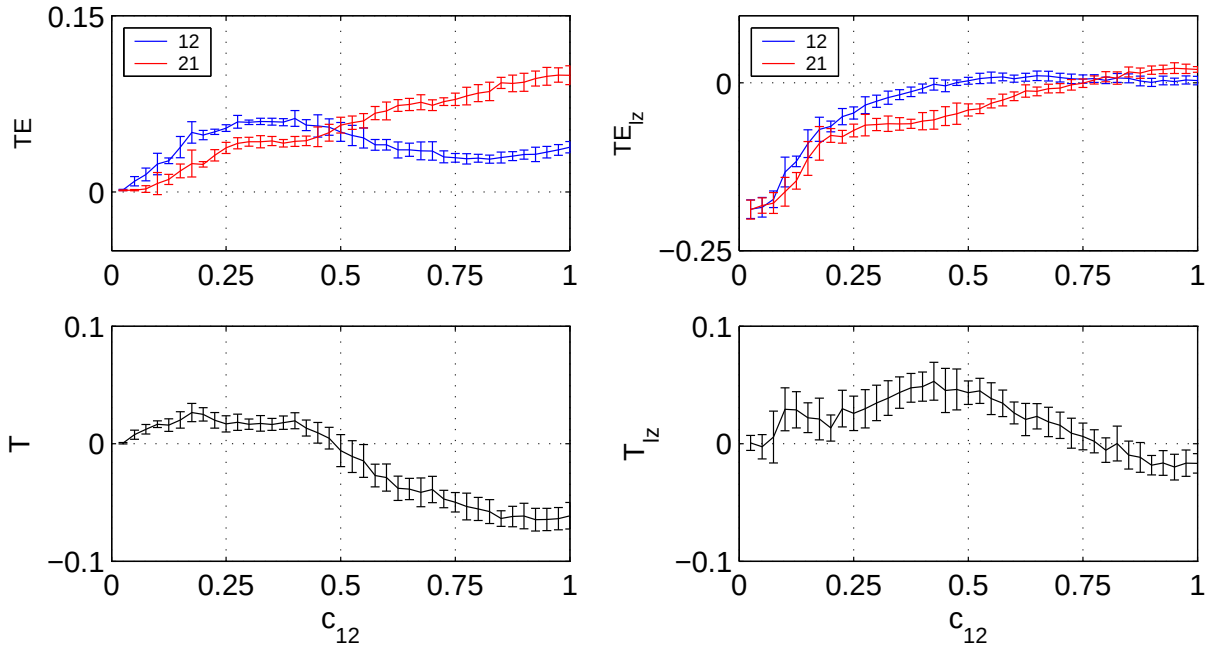


Figure 6.7.: Same as in Fig. 6.6 but for two unidirectionally coupled Rössler oscillators and embedding parameters: $m = 3$ and $\tau = 16$.

the LZ-based estimator of entropy rate are only asymptotically ($N \to \infty$) related to each other. Similarly to the problem of the negativity of mutual LZ-complexity the negativity of the absolute value of LZ-based transfer entropy may considerably limit its practical applicability. However, in contrast to mutual LZ-complexity, the LZ-based transfer entropy is an asymmetric measure of interdependence between two sequences of data. It is easy to see (Figs. 6.6 and 6.7) that LZ-based transfer entropy can, nevertheless, be used to reveal the directionality of interaction by defining the so called *LZ-based asymmetry* index as

$$T_{LZ}(S_1, S_2) = TE_{LZ}(S_1, S_2) - TE_{LZ}(S_2, S_1). \tag{6.6}$$

Thus, positive values of $T_{LZ}(S_1, S_2)$ indicate that the first system is more actively driving the second system and this results in an asymmetry of interaction in the direction "$1 \to 2$". Negative values of $T_{LZ}(S_1, S_2)$ indicate asymmetry of interaction in the opposite direction. By following the definition given in [SL08] we also defined the asymmetry index T as

$$T(S_1, S_2) = TE(S_1, S_2) - TE(S_2, S_1), \tag{6.7}$$

where $TE(S_1, S_2)$ and $TE(S_2, S_1)$ denoted the symbolic transfer entropies between S_1 and S_2. Similar to results obtained in section 4.2.2 (see Fig. 4.7) our analysis here revealed that both asymmetry indices T and T_{LZ} were positive and thus allowed to correctly identify the direction of interaction between unidirectionally coupled Lorenz oscillators (see Fig. 6.6). Moreover, the indices showed qualitatively similar dependencies on the coupling strength c_{12} approaching zero for $c_{12} > 8$ when Lorenz oscillators reached the regime of complete synchronization. The estimation of T and T_{LZ} for a system of unidirectionally coupled

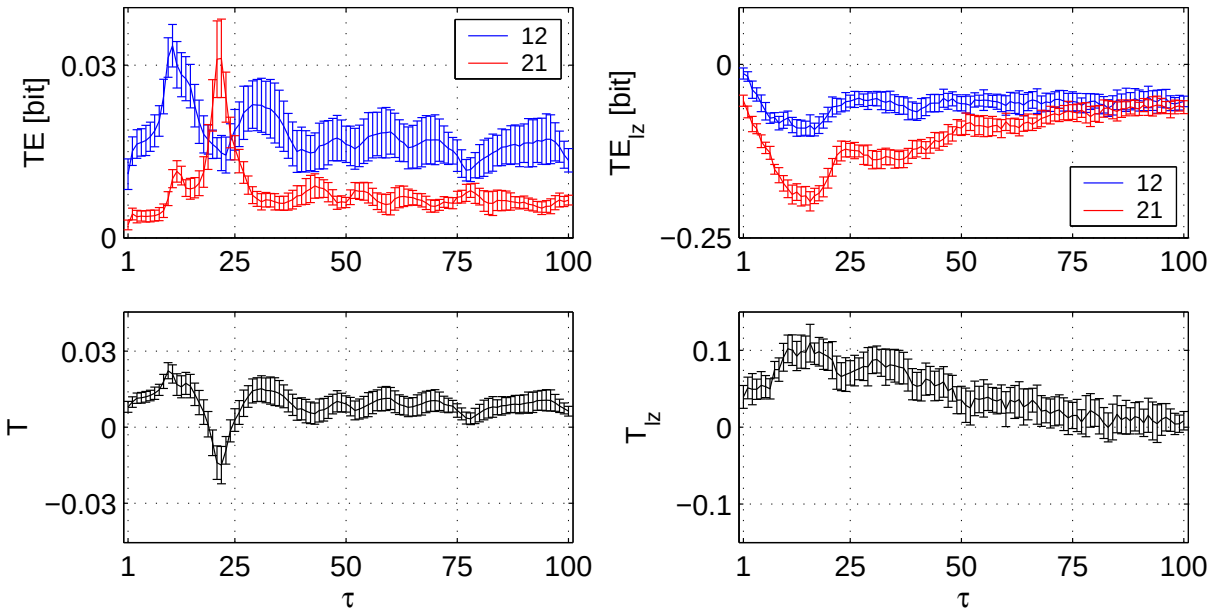


Figure 6.8.: Symbolic transfer entropies TE of first order and corresponding asymmetry index T along with algorithmic (LZ-based) transfer entropies TE_{lz} and corresponding asymmetry index T_{lz} between $x-$ components of two unidirectionally coupled Lorenz oscillators with coupling strength $c_{12} = 5$. $x-$ components of both oscillators are transformed into permutation symbols with embedding dimension $m = 3$ and varying delay times $\tau \in [1, 100]$). Each point on all plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

Rössler oscillators for increasing coupling strength c_{12} revealed a qualitative difference between dependencies of both indices on c_{12} (see Fig. 6.7). Similar to results of section 6.1 the asymmetry index T allowed us to correctly infer the direction of interaction between oscillators only for $c_{12} < 0.5$ whereas for $c_{12} \in [0.5, 1]$ the negative value of asymmetry index ($T < 0$) indicated an incorrect direction of interaction (i.e., "responder-to-driver"). Next, we repeated our analysis of entropy transfer between Rössler oscillators but now by using LZ-based transfer entropies ($TE_{LZ}(S_1, S_2)$ and $TE_{LZ}(S_2, S_1)$) and then computed the LZ-based asymmetry index T_{LZ} (Eq. 6.6). In comparison to T, the LZ-based asymmetry index remained positive and thus allowed us to correctly infer the directionality of interaction between oscillators for $c_{12} < 0.75$ extending the range of coupling strengths where the correct inference was still possible. Similar results were obtained in section 6.1 by analyzing the directionality of interaction with the tenth order symbolic transfer entropy (see Fig. 6.1). Next, we address the question how and to what extent the incorrect construction of the permutation partition (by incorrect selection of delay time τ) can effect the estimated entropy transfer between interacting Lorenz as well as Rössler oscillators. In sections 4.2.2 and 6.1 it was demonstrated that the symbolic transfer entropies between chaotic oscillators depended on the delay time τ, and for some τ, the symbolic transfer entropy in the direction "responder-to-driver" exceeded symbolic transfer entropy in the direction "driver-

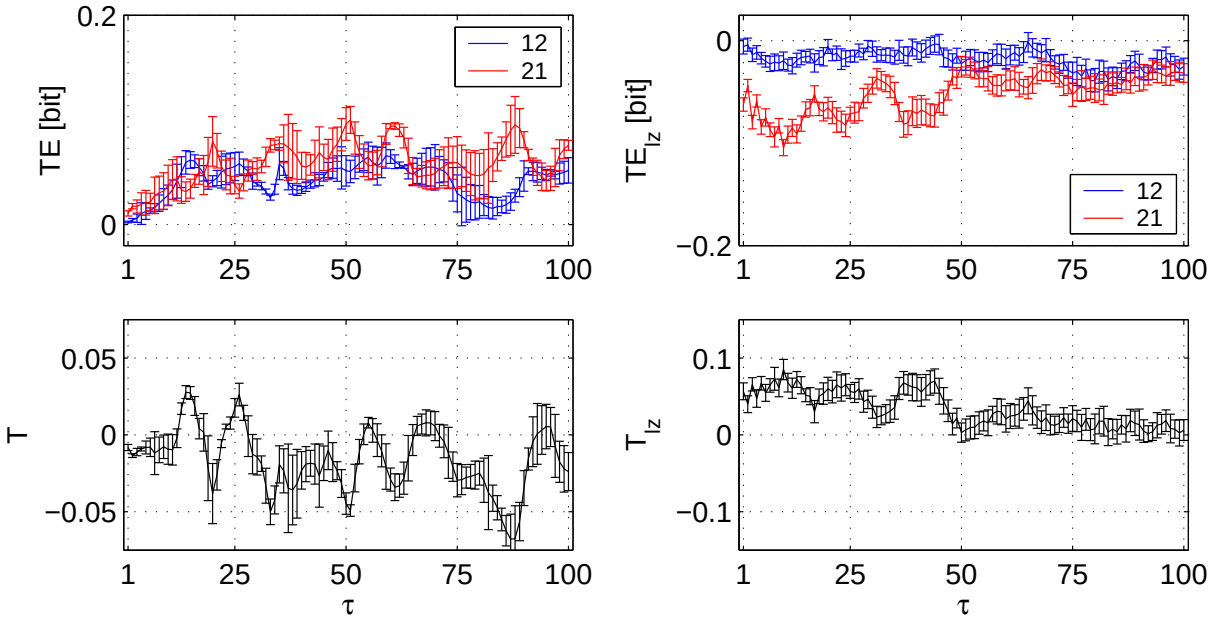


Figure 6.9.: Same as in Fig. 6.8 but for two unidirectionally coupled Rössler oscillators with fixed coupling strength $c_{12} = 0.35$.

to-responder". This led to the incorrect inference of the directionality of interaction. It was then also shown that estimating the entropy transfer between Rössler oscillators with the tenth order symbolic transfer entropy allowed us to minimize the influence of incorrect selection of the delay time τ. To test whether a similar effect can be achieved by estimating entropy transfer with LZ-based transfer entropy we again considered two unidirectionally coupled Lorenz (Rössler) oscillators with a fixed coupling strength $c_{12} = 5$ ($c_{12} = 0.35$) and different values of the delay time $\tau \in [1, 100]$. Figs. 6.8 and 6.9 show the asymmetry indices T and T_{LZ} measured with either symbolic (Eq. 6.7) or LZ-based (Eq. 6.6) transfer entropy for different τ. For a system of coupled Lorenz oscillators we observed that for $\tau \approx 20$ the values of the first order symbolic transfer entropy in the direction "responder-to-driver" significantly exceeded its values in the direction "driver-to-responder". This resulted in the negativity of the asymmetry index T and thus in the incorrect inference of the directionality of interaction between the Lorenz oscillators. Analysis of the LZ-based asymmetry index T_{LZ} revealed that its value remained positive for all τ (Fig. 6.8). For a system of coupled Rössler oscillators we observed that LZ-based asymmetry index T_{LZ} remained positive (correctly indicating the directionality of interaction "driver-to-responder") for all values of delay time τ whereas the careful selection of τ was needed ($\tau = T_{Roes}/2 = 16$) to correctly infer the directionality of interaction with asymmetry index T (Fig. 6.9).

The findings obtained in this section demonstrated a close relationship between LZ-based transfer entropy with a high order estimate of symbolic transfer entropy analyzed in section 6.1. Analogously to the results obtained with the tenth order symbolic transfer entropy we observed here that the LZ-based approach to estimate entropy transfer between chaotic oscillators exhibited less dependency on the variation in the construction of permutation

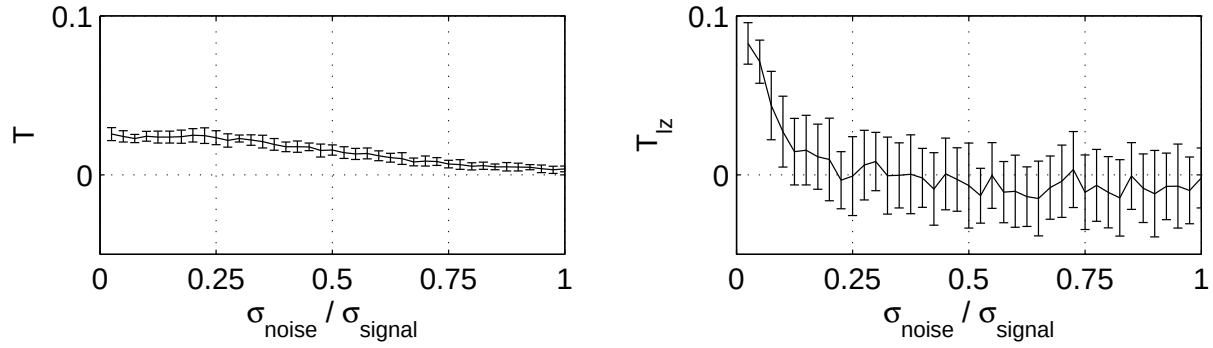


Figure 6.10.: Asymmetry indexes computed with the symbolic transfer entropies of the first order T (left) and with algorithmic (LZ-based) transfer entropies T_{lz} (right) between two unidirectionally coupled Lorenz oscillators for different noise-to-signal ratio and with fixed coupling strength $c_{12} = 5$. $x-$ components of both oscillators were first transformed into the symbolic representation: two series of permutation symbols with the embedding dimensions $m = 3$ and delay time $\tau = 10$. Each point on both plots corresponds to the mean value taken over 20 realizations. Error bars denote standard deviations.

partition than the conventional approach based on the symbolic transfer entropy of the first order. Additionally, we were able to correctly infer the directionality of interaction between strongly coupled Rössler oscillators ($c_{12} \in [0.5, 0.75]$) as we did with the tenth order symbolic transfer entropy in section 6.1. Based on the notion of the mutual LZ-complexity introduced in [ZRB05], the algorithmic or LZ-based transfer entropy represents a complementary measure to estimate high order entropy transfer in time series of, in general, arbitrary lengths.

Algorithmic transfer entropy between noise-contaminated signals

In this section we address the question of robustness of the LZ-based transfer entropy along with the first order symbolic transfer entropy against different amounts of noise in time series to infer the directionality of interaction between dynamical systems. Similar to the analysis made in section 4.2.3 we again analyzed a pair of time series generated by two unidirectionally coupled Lorenz oscillators (Eq. A.8; $R_{1,2} \in \mathbf{N}(28, 0.5)$) with a fixed coupling strength $c_{12} = 5$ and a different amount of additive noise. The first system (driver) was coupled into the second system (responder) and it was expected to find a positive value of both asymmetry indices $T > 0$ and $T_{LZ} > 0$ (Eqs. 6.6 and 6.7). In the same way as we did in section 4.2.3 the differential equations were integrated using a fourth order Runge-Kutta algorithm with integration step $dt = 0.005$ and then downsampled to $d\tilde{t} = 0.03$. The initial conditions were normally distributed with zero mean and unit variance. In order to eliminate transients, the first 10^4 iterations were discarded. Next, $x-$ components of each Lorenz oscillator were superimposed with additive white noise using different noise-to-signal ratios and transformed into a pair of series of permutation symbols S_1 and S_2 by setting the embedding dimension to $m = 3$ and delay time to $\tau = 10$. We

generated 20 realizations of Lorenz time series of length $N = 10^4$ for each value of noise-to-signal ratio. We gradually increased the standard deviation of the additive white noise $\sigma_{noise} = 0, \ldots, \sigma_{signal}$ and thus increased the noise-to-signal ratio from 0 to 1 with a step 0.025. Figure 6.10 shows the obtained values of both asymmetry indices T and T_{LZ} for different noise-to-signal ratios. For a zero noise-to-signal ratio (i.e., when $\sigma_{noise} = 0$) both indices T and T_{LZ} were positive and thus correctly indicated the asymmetry of interaction in the direction "driver-to-responder". However, both indices gradually approached zero when the amount of noise in the time series was increased. The qualitative comparison of T and T_{LZ} revealed that the latter asymmetry index was more sensible to noise than the former one and as a consequence it showed a faster decline of its value for increasing noise-to-signal ratios. Thus, our findings indicated that already a moderate amount of noise can corrupt high order correlations in data and thus limit the ability of the LZ-based approach to reliably infer the direction of interaction in noisy time series. In contrast to this, the asymmetry index T (based on the first order symbolic transfer entropies which by definition accounts for three-point correlations in symbol series (see Eq. 4.7) demonstrated less sensitivity against noise in time series.

7. Summary and outlook

In this thesis, information-theoretic measures have been evaluated with respect to their capability to characterize interactions between dynamical systems. To this end, these measures were first applied to characterize interactions between dynamical model systems with known equations of motion. Additionally, we investigated the influence of different external factors such as noise. Finally, we applied the information-theoretic approach to characterize strength and directionality of interactions in the human brain.

First (chapter 3), we addressed the question as to how and to what extent two different symbolic representations of a chaotic dynamical system can be used to estimate its Kolmogorov-Sinai (KS) entropy. KS-entropy characterizes the amount of entropy produced by a dynamical system per time unit [Kol59, Sin59, CGG89]. The estimation of KS-entropy allowed us to characterize the rate of entropy production in the system during its temporal evolution. We started our analysis by estimating KS-entropy of the tent map for different values of its control parameter. The variation of this parameter allowed us to control the KS-entropy of the tent map. In order to obtain a symbolic representation of real-valued dynamics of the tent map, we first applied the permutation symbols approach proposed in [BP02]. Similar to findings obtained in [BKP02], we found that Shannon entropy of permutation symbols converged very slowly to KS-entropy of the tent map for increasing embedding dimension. This posed a significant limitation for the numerical estimation of KS-entropy by using a Shannon entropy estimator[1]. In contrast, the estimates of KS-entropy based on first-order Shannon entropy rate demonstrated a fast converging behavior. The fact that KS-entropy can be sufficiently well approximated already with a first-order estimate of Shannon entropy rate indicated that the symbolic representation of the tent map obtained with a permutation symbols approach can be regarded as a realization of some first-order Markov process. This indicates that the permutation partition proposed in [BP02, BKP02] represents a Markov partition of the tent map. The question how the permutation symbols approach can be used to estimate KS-entropy for continuous-time dynamical systems (e.g. chaotic oscillators) remains open and demands further theoretical and numerical analyses.

In the second part of chapter 3 we again estimated KS-entropy of the tent map but now by applying a rather different symbolization approach. By following [EST01] we used a threshold-crossing partitioning scheme and transformed the real-valued time series into series of binary symbols. Our analysis revealed that a good approximation of KS-entropy of

[1] Indeed, as it was shown in section 2.3.1, the number of different permutation symbols A_{perm} increases very quickly with the embedding dimension as $A_{perm} = m!$. This may pose a serious limitation for the numerical estimation of Shannon entropy for a large embedding dimension m, and, in general, a large amount of data can be required in order to avoid significant statistical and systematic errors of obtained estimates [BKP02].

the tent map was achieved by using at least an $8^{th}-$order Shannon entropy rate estimator. This finding indicated that an obtained series of binary symbols cannot be considered as a realization of the first-order Markov processes but rather has to be considered as a realization of some high-order Markov process (see also [BSLZ01]). This may represent a serious limitation for the application of the threshold-crossing partitioning scheme to experimental data. Next, we extended our analysis by applying a complementary approach to estimate the entropy rate based on the notion of Lempel-Ziv (LZ) complexity. It is known that this approach allows to estimate the entropy rate of symbol series exhibiting long-term temporal correlations [SG96]. Obtained findings indicated that the LZ-based approach indeed provided a better approximation of KS-entropy of the tent map than low-order estimators of Shannon entropy rate. Similar results were also obtained by estimating KS-entropy of the Hénon map.

In chapter 4 we applied information-theoretic measures to characterize interactions between dynamical model systems. Real-valued time series generated with different dynamical model systems were first transformed into their symbolic representations by using a permutation symbols approach introduced in [BP02]. In order to estimate the strength of interactions between model systems we applied the symbolic mutual information which is a measure for the characterization of the amount of common information (measured in bits) shared between two time series. To verify the capability of this measure to characterize the strength of interactions between dynamical systems, we applied the symbolic mutual information to characterize interactions between two coupled chaotic oscillators. Additionally, for comparative purposes, we also estimated the degree of phase synchronization between these oscillators by using the mean phase coherence [MLDE00, Mor03, MAK$^+$03]. A qualitative comparison indicated that both measures – capturing different aspects of the dynamics – nevertheless showed a qualitatively similar dependence on the coupling strength. Analysis of time series with different noise-to-signal ratios indicated that the mean phase coherence is less sensitive to the amount of noise in the data than the symbolic mutual information.

In the second part of chapter 4, in order to characterize the directionality of interactions between two unidirectionally coupled dynamical systems, we estimated the amount of entropy transfer between them by using the symbolic transfer entropy introduced in [SL08]. As was originally pointed out in [Sch00], due to the limited amount of data in most practical applications the entropy transfer between two dynamical systems can only be estimated reliably by using a first-order estimator. As it was shown in [SL08, SL09], the application of the first-order estimator of entropy transfer allowed to correctly characterize the directionality of interactions between different dynamical model systems as well as in experimental data. Following [SL08], we characterized the directionality of interactions between dynamical model systems by using a first-order estimator of transfer entropy. Our findings showed that for all dynamical model systems considered in this chapter the directionality of interactions can be correctly identified[2]. However, it was also observed that obtained positive values of the symbolic transfer entropy in the direction "responder-to-driver" significantly deviated from zero, despite the absence of coupling in this direction. Following the analy-

[2]Similar to findings obtained in [SL08], we observed that the inference of the directionality of interactions was not possible for sufficiently strong coupling strengths when the dynamical model systems entered the regime of complete synchronization.

sis made in [MK02], we tested the assumption that an observed positive entropy transfer in the direction "responder-to-driver" just reflected misestimation of the symbolic transfer entropy due to finite sample effect [HSE94]. Our findings indicated that such a spurious entropy transfer between two dynamical systems cannot be explained only as a result of a finite sample effect but mainly reflects the influence of static correlations between two series of permutation symbols. A further investigation of this problem led us to the development of the corrected symbolic transfer entropy which allowed us to correct an observed positive entropy transfer in the direction "responder-to-driver". In order to further investigate the influence of static correlations in the data on the estimates of the symbolic transfer entropy, we considered a network of uncoupled chaotic oscillators (responders) which are driven by the external force (two drivers). Our findings indicated that already a small amount of the external force can be sufficient to induce a spurious entropy transfer between responders. Moreover, by further increasing the strength of the external force, we observed stronger synchronization of responders. This increased the degree of static correlations in the data and led to an increase of spurious entropy transfer between them. By measuring entropy transfer with the corrected transfer entropy allowed us to reduce the amount of spurious entropy transfer between responders and therefore to better identify the two drivers. Further analyses, by selectively adding a different amount of noise to the responders, demonstrated that less noisy oscillators spuriously appeared to be driven by more noisy ones. We found that the capability of the symbolic transfer entropy to correctly identify the driving structures in the network of noisy oscillators can be recovered by applying the correction scheme proposed in this chapter.

In chapter 5 we applied the information-theoretic measures presented in chapter 4 to characterize the strength and directionality of interactions in multi-channel electroencephalographic (EEG) recordings from epilepsy patients. First, we applied the symbolic mutual information to estimate the amount of common (mutual) information shared between different EEG recordings. Our findings indicated that inter-regional EEG recordings (i.e., pairs of EEG recordings corresponding to different brain structures), on average, shared less common information with each other than EEG recordings corresponding to intra-regional contact-combinations. Thus, we found that different neuroanatomical structures can be identified as clusters of higher values of mutual information. This finding is in agreement with results obtained in [Ost08, MOA$^+$08] where it was demonstrated that a higher degree of synchronization in EEG recordings corresponded to intra- rather than inter-regional channel-combinations. As it was already mentioned in [Ost08] the functional differentiation of brain structures via the analysis of EEG recordings of epilepsy patients can have a great value for neuroscientific research and also for clinical practice.

In the second part of chapter 5 we applied the symbolic transfer entropy as well as the corrected symbol transfer entropy to characterize the directionality of interactions in the human hippocampus for a group of 26 epilepsy patients. Statistical analyses indicated that for a significantly large subgroup of patients, we observed a pronounced asymmetry in the entropy transfer within this brain structure. This difference was even more pronounced when the entropy transfer between EEG recordings was estimated by using the corrected symbolic transfer entropy. The estimation of cross-correlation coefficients between EEG recordings indicated that observed asymmetry in the transfer entropy flows might be related

to a wave-traveling phenomenon[3] in the hippocampus. The frequency-selective analysis of the entropy transfer in the hippocampus showed that the mean value (averaged over the group of 26 patients) of transfer entropy flows approached maximal values when EEG recordings were filtered in the $\theta-$band. This is in agreement with a known fact that theta-oscillations represent a dominant rhythmic activity in the hippocampus [Buz02, MOA$^+$08, Lub09]. The asymmetry in the hippocampal transfer entropy flow for a significantly large subgroup of patients has only been obtained for the non-focal brain hemisphere. This indicates that underlying mechanisms of this phenomenon might be impaired by epilepsy. It is important to point out that findings obtained in this chapter can only be considered as preliminary results. Further analysis is needed to address the question as to what extent the obtained values of the transfer entropy flows in the hippocampus were biased by the recording montage of EEG signals.

In the final part of the thesis (chapter 6), we demonstrated that by measuring entropy transfer between chaotic oscillators by using a high-order transfer entropy it was possible to obtain a better characterization of the directionality of interactions between these oscillators than by using a low-order estimate of entropy transfer. Following this finding, we then introduced a complementary approach for the estimation of high-order transfer entropies. The notion of algorithmic transfer entropy, introduced in this thesis, is related to the Lempel-Ziv complexity of a symbol series and therefore does not require the reconstruction of high-dimensional empirical probability functions. The estimation of algorithmic transfer entropy between two interdependent binary Markov processes revealed that obtained estimates converged quickly (for an increasing length of symbol series) to the analytically defined value of Shannon transfer entropy between them. Further analyses of entropy transfer between different dynamical model systems showed that obtained estimates of the algorithmic transfer entropy also approached estimates of entropy transfer obtained by using the high-order symbolic transfer entropy. However, it is important to point out that our findings also indicated that a moderate amount of noise in the data reduced the capability of the algorithmic approach to characterize the directionality of interactions between dynamical model systems.

During the last decades, advances in the theory of nonlinear dynamical systems provided us a variety of tools to characterize the irregular behavior of various natural phenomena. A number of methods of time series analysis have been introduced and successfully applied for the characterization of complexity and interactions in experimental data. The symbolic time series analysis provides a solid and broadly used toolkit for the characterization of interactions between nonlinear dynamical systems. The analyses presented in this thesis allowed us to investigate several important limitations of information-theoretic measures which may appear when experimental data exhibit strong correlations. It was demonstrated that a high degree of static and/or long-term temporal correlations can, in general, lead to the incorrect inference of directionality of interactions between underlying dynamical systems. Thus, the methods of time series analysis, presented in this thesis, can find a broad range of applications in fields where the influence of correlations in data cannot be neglected.

[3]Recent findings obtained for freely behaving rats indicated that theta-oscillations in the local field potential are traveling waves which propagate along the septotemporal axis of the hippocampus [Lub09].

A. Appendix

A.1. Deterministic dynamical model systems

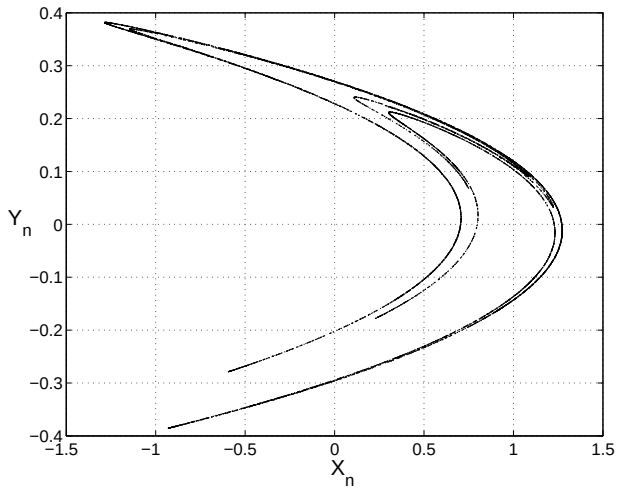
Tent map

The tent map is a widely used (see [Str01]) one-dimensional iterative map exhibiting chaotic behavior for all values of $\rho \in [1, 2]$

$$x_{n+1} = \rho(1 - 2|x_n - 0.5|). \tag{A.1}$$

Hénon map

The Hénon map is a two-dimensional invertible iterated map proposed by Michel Hénon exhibiting chaotic behavior for the canonical set of parameters $a = 1.4$ and $b = 0.3$ [H76́].



Hénon attractor

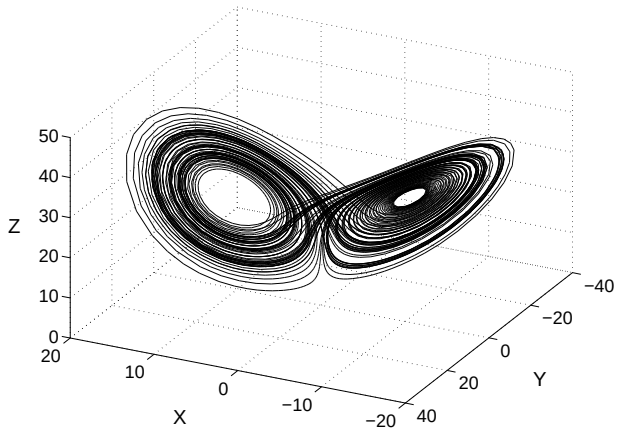
$$\begin{aligned} x_{n+1} &= 1 - ax_n^2 + by_n \\ y_{n+1} &= x_n. \end{aligned} \tag{A.2}$$

The Hénon equations (Eq. A.2) asymptotically map a set of initial conditions into a subset of points of a two dimensional plane known as Hénon attractor. This attractor is a fractal set and has a non-integer dimension.

Lorenz oscillator

The Lorenz oscillator is a system of three nonlinear ordinary differential equations exhibiting a chaotic flow. This system was introduced by Edward Lorenz to describe convection

rolls arising in the atmosphere.

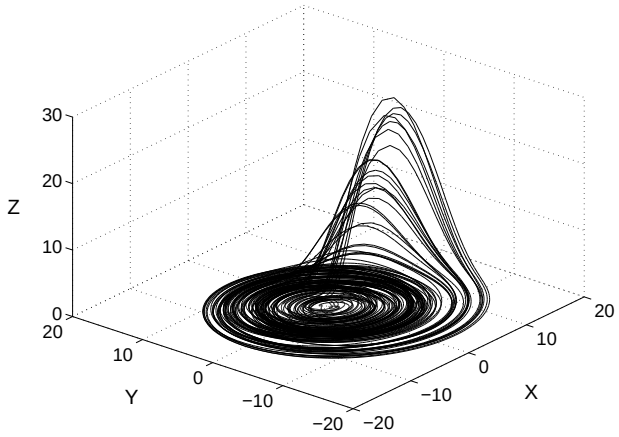


Lorenz attractor

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= x(R - z) - y \\ \dot{z} &= xy - bz \end{aligned} \tag{A.3}$$

where σ is called the *Prandtl* number and R is called the *Rayleigh* number [Lor63]. For $\sigma = 10$, $b = 8/3$ and $R = 28$ the system exhibits chaotic behavior and displays knotted periodic orbits for other values of R. The geometrical representation of the Lorenz dynamics in state space is the well known Lorenz attractor.

Rössler oscillator



Rössler attractor

$$\begin{aligned} \dot{x} &= -(y + z) \\ \dot{y} &= x + ay \\ \dot{z} &= b + z(x - c) \end{aligned} \tag{A.4}$$

The Rössler oscillator is a system of three nonlinear ordinary differential equations exhibiting a chaotic flow. This system was introduced by Otto Rössler and later has been used in modeling nonlinear processes in chemical reactions [Roe76]. For $a = 0.165$, $b = 0.2$, and $c = 10$ the system exhibits chaotic behavior. The geometrical representation of the Rössler dynamics in state space is the well known Rössler attractor.

Two coupled tent maps

Two unidirectionally coupled tent maps identical to the ones used in [Sch00], with equations

$$\begin{aligned} x_1(n+1) &= f(x_1(n)) \\ x_2(n+1) &= f(c_{12}x_1(n) + (1-c_{12})x_2(n)), \end{aligned} \tag{A.5}$$

where the function f is defined by the tent map equation (Eq. A.1; $\rho = 2$) and c_{12} denotes the coupling strength.

Two coupled Hénon maps

Two unidirectionally coupled Hénon maps identical to the ones used in [QAG00], with equations

$$\begin{aligned} x_1(n+1) &= 1.4 - x_1^2(n) + b_1 y_1(n) \\ y_1(n+1) &= x_1(n) \end{aligned} \tag{A.6}$$

for the driver, and

$$\begin{aligned} x_2(n+1) &= 1.4 - (c_{12}x_1x_2 + (1-c_{12})x_2^2) + b_2 y_2(n) \\ y_2(n+1) &= x_2(n) \end{aligned} \tag{A.7}$$

for the responder. The parameters b_1 and b_2 are both set to 0.3 for the analysis of structurally identical systems, and to 0.3 and 0.1 for the analysis of structurally nonidentical systems.

Two diffusively coupled Lorenz oscillators

Two diffusively coupled Lorenz oscillators identical to the ones used in [SL08], with equations

$$\begin{aligned} \dot{x}_1 &= 10(y_1 - x_1) + c_{21}(x_2 - x_1) \\ \dot{y}_1 &= x_1(R_1 - z_1) - y_1 \\ \dot{z}_1 &= x_1 y_1 - (8/3)z_1 \\ \dot{x}_2 &= 10(y_2 - x_2) + c_{12}(x_1 - x_2) \\ \dot{y}_2 &= x_2(R_2 - z_2) - y_2 \\ \dot{z}_2 &= x_2 y_2 - (8/3)z_2, \end{aligned} \tag{A.8}$$

where $R_{1,2} \in \mathbf{N}(28, 0.5)$ whereas c_{12} and c_{21} denote diffusive coupling strengths.

Two diffusively coupled Rössler oscillators

Two diffusively coupled Rössler oscillators identical to the ones used in [OMWL07], with equations

$$\begin{aligned}
\dot{x}_1 &= -\omega_1 y_1 - z_1 + c_{21}(x_2 - x_1) \\
\dot{y}_1 &= \omega_1 x_1 + 0.165 y_1 \\
\dot{z}_1 &= 0.2 + z_1(x_1 - 10) \\
\dot{x}_2 &= -\omega_2 y_2 - z_2 + c_{12}(x_1 - x_2) \\
\dot{y}_2 &= \omega_2 x_2 + 0.165 y_2 \\
\dot{z}_2 &= 0.2 + z_2(x_2 - 10).
\end{aligned} \tag{A.9}$$

Here, $\omega_{1,2}$ denote internal (characteristic) frequencies of Rössler oscillators , c_{12} and c_{21} are diffusive coupling strengths between oscillators.

Diffusively coupled Rössler and Lorenz oscillators

Two unidirectionally coupled Rössler (driver) and Lorenz (responder) oscillators identical to the ones used in [QAG00], with equations

$$\begin{aligned}
\dot{x}_r &= -6(y_r - z_r) \\
\dot{y}_r &= 6(x_r + 0.2 y_r) \\
\dot{z}_r &= 6(0.2 + z_r(x_r - 5.7)) \\
\dot{x}_l &= 10(y_l - x_l) \\
\dot{y}_l &= 28 x_l - y_l - x_l z_l + c_{rl} y_r^2 \\
\dot{z}_l &= x_l y_l - (8/3) z_l.
\end{aligned} \tag{A.10}$$

Here, c_{rl} denotes the coupling strength between oscillators.

Network of diffusively coupled Lorenz oscillators

A network of 20 diffusively coupled Lorenz oscillators similar to a network of coupled Rössler oscillators used in [OMWL08, SL09]. Each node j ($j = 1, \ldots, 20$) in our network obeys the differential equation:

$$\begin{aligned}
\dot{x^j} &= 10(y^j - x^j) + \zeta_d^j \\
\dot{y^j} &= x^j(R^j - z^j) - y^j \\
\dot{z^j} &= x^j y^j - (8/3) z^j
\end{aligned} \tag{A.11}$$

where $R_j \in \mathbf{N}(28, 0.5)$ and ζ_d^j denotes an unidirectional coupling term which is defined as

$$\zeta_d^j = c_d(x^4 + x^5 - 2x^j). \tag{A.12}$$

A.2. Mean phase coherence

Traditionally, phase synchronization is defined as the locking of the phases ϕ_1 and ϕ_2 of two oscillating systems [Huy73]:

$$\phi_1(t) - \phi_2(t) = const. \tag{A.13}$$

In order to quantify the degree of phase synchronization we used the *mean phase coherence* R [MLDE00, PRK01]:

$$R = \left| \frac{1}{N} \sum_{j=0}^{N} e^{i[\phi_1(j\Delta t) - \phi_2(j\Delta t)]} \right| = 1 - V, \tag{A.14}$$

where $1/\Delta t$ is the sampling rate of the discrete time series of length N. V denotes the circular variance of an angular distribution obtained by transforming the differences in phase onto the unit circle in the complex plane. By definition R is confined to the interval [0,1] where $R = 1 (V = 0)$ indicates fully synchronized systems. In order to determine the phases $\phi_{1,2}(t)$ of two signals $s_1(t)$ and $s_2(t)$, we followed [MLDE00] and used the analytic signal approach which renders an unambiguous definition of the instantaneous phase for an arbitrary signal $s(t)$:

$$\phi(t) = \arctan \frac{\tilde{s}(t)}{s(t)} \tag{A.15}$$

where

$$\tilde{s}(t) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{+\infty} \frac{s(\tau)}{t - \tau} d\tau \tag{A.16}$$

is the Hilbert transform of the signal ($\mathcal{P}$ denoting the Cauchy principal value). Application of the convolution theorem turns Eq. (A.16) into

$$\tilde{s}(t) = -i\mathcal{F}^{-1}\Big[\mathcal{F}[s(t)]\text{sign}(\omega)\Big] \tag{A.17}$$

where *sign* is a sign function, $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote Fourier and inverse Fourier transform, respectively. From this definition we can see that the Hilbert transform performs a phase shift (in frequency domain) of the signal $s(t)$ by $\pi/2$ while keeping the power spectrum of it unchanged. For the analysis of time series of finite length $N = 2^n$ (where n is integer) we applied a fast Fourier transform (for more details see [PTVF02]).

A.3. Entropy of random variables and stochastic processes

Entropy of multivariate random variables

Let $\{A_j\}_{j=1}^n$ be a set of n discrete random variables defined by a probability function $p(a_1, \ldots, a_n)$. The *block-entropy* for n random variables is defined as (see [CT91])

$$H(A_1, \ldots, A_n) = -\sum_{\{a_i\}} p(a_1, \ldots, a_n) \log(p(a_1, \ldots, a_n)). \tag{A.18}$$

A random variable A_j has a probability function $p(a_j) = \sum_{\{a_i\}:i\neq j} p(a_1, \ldots, a_n)$ and its entropy is defined as $H(A_j) = -\sum_{a_j} p(a_j) \log(p(a_j))$. If all random variables are statistically independent from each other (i.e., $p(a_1, \ldots, a_n) = p(a_1)p(a_2)\ldots p(a_n)$) then the block-entropy can be computed as $H(A_1, \ldots, A_n) = \sum_{i=1}^n H(A_i)$. In case where random variables are not independent the block-entropy is defined as a series of, in general, n terms, i.e.,

$$H(A_1, \ldots, A_n) = \sum_{i=1}^{n} H(A_i) - \sum_{i<j} I_2(A_i, A_j) + \sum_{i<j<k} I_3(A_i, A_j, A_k) - \ldots \tag{A.19}$$

where two-point correlations are taken into account by the *two-point mutual information* function

$$I_2(A_i, A_j) = \sum_{a_i, a_j} p(a_i, a_j) \log \frac{p(a_i, a_j)}{p(a_i)p(a_j)} \tag{A.20}$$

or alternatively

$$I_2(A_i, A_j) = H(A_i) + H(A_j) - H(A_i, A_j). \tag{A.21}$$

Three-point correlations are taken into account by the *three-point mutual information* function

$$I_3(A_i, A_j, A_k) = -\sum_{a_i, a_j, a_k} p(a_i, a_j, a_k) \log \frac{p(a_i, a_j, a_k)}{\hat{p}(a_i, a_j, a_k)} \tag{A.22}$$

where $\hat{p}$ is the so called *Kirkwood* approximation of p,

$$\hat{p}(a_i, a_j, a_k) = \frac{p(a_i, a_j)p(a_j, a_k)p(a_i, a_k)}{p(a_i)p(a_j)p(a_k)} \tag{A.23}$$

or alternatively

$$I_3(A_i, A_j, A_k) = \sum_{s\in\{i,j,k\}} H(A_s) - \sum_{(s_1>s_2)\in\{i,j,k\}} H(A_{s_1}, A_{s_2}) + H(A_i, A_j, A_k). \tag{A.24}$$

Following the same rules the higher order mutual information functions are obtained analogously [Mat00].

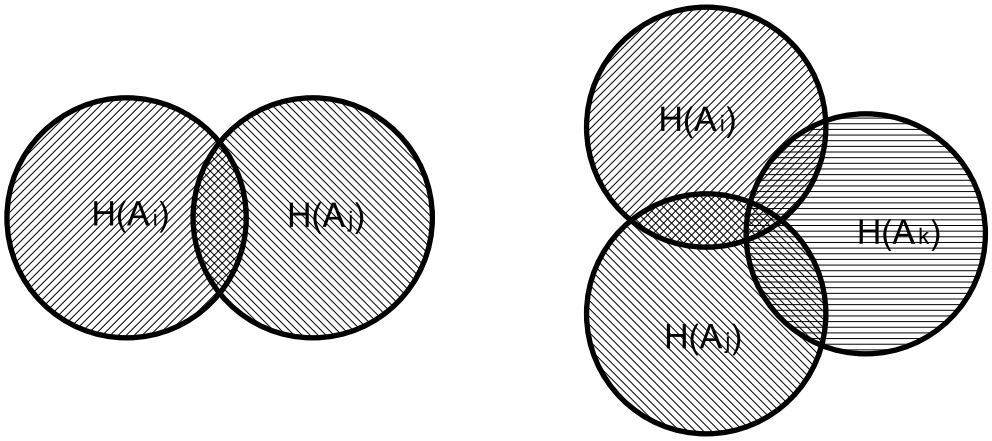


Figure A.1.: Schematic representation of entropies $H(A_i)$, $H(A_j)$, $H(A_k)$ for the cases of two (left) and three (right) random variables. The entropy of each variable is represented by a hatched circle with an area proportional to the absolute value of the entropy. If random variables are correlated then circles intersect each other. An intersection between two circles represents two-point mutual information $I_2(A_i, A_j)$ (left) whereas intersection between three circles represent three-point mutual information $I_3(A_i, A_j, A_k)$ (right). It is easy to see that both mutual information functions are symmetric under the permutation of their arguments. Additionally, $I_2(A_m, A_n) = 0$ for $\forall(m \neq n) \in \{i, j, k\}$ implies that $I_3(A_i, A_j, A_k)$=0.

Conditional entropy and mutual information of multivariate random variables

Let $\{A_i\}_{i=1}^n$ be a set of n discrete random variables defined by a probability function $p(a_1, \ldots, a_n)$. We can define a *conditional entropy* of the random variable A_1 with a condition that outcomes of the random variables $\{A_i\}_{i=2}^n$ are given by

$$H(A_1|A_2, \ldots, A_n) = -\sum_{\{a_i\}} p(a_1, \ldots, a_n) \log(p(a_1|a_2 \ldots, a_n)), \tag{A.25}$$

where $p(a_1|a_2 \ldots, a_n)$ is a conditional probability of variable a_1. According to the definition of conditional probabilities: $p(a_1|a_2 \ldots, a_n) = p(a_1, a_2 \ldots, a_n)/p(a_2 \ldots, a_n)$ it is easy to see that

$$H(A_1|A_2, \ldots, A_n) = H(A_1, A_2, \ldots, A_n) - H(A_2, \ldots, A_n). \tag{A.26}$$

By combing the equations A.21 and A.26 the *conditional mutual information* between for example random variables A_1 and A_2 with a condition that outcomes of random variables $\{A_i\}_{i=3}^n$ are given is defined as

$$I_2(A_1, A_2|A_3, \ldots, A_n) = H(A_1|A_3, \ldots, A_n) + H(A_2|A_3, \ldots, A_n) - H(A_1, A_2|A_3, \ldots, A_n). \tag{A.27}$$

The higher order conditional mutual information functions are defined analogously [Mat00].

Entropy rate of bivariate stochastic Markov process

Let us first consider an k-order ($k > 0$) Markov process (chain) Z that is defined by an $(k+1)$-dimensional density functions $g(z(t_i), z_1(t_{i-1}), \ldots, z(t_{i-k}))$. For the sake of simplicity we can rewrite the probability function as $g(z(t_i), z(t_p))$, where the variable $z(t_i)$ defines the current state of the process whereas the variable $z(t_p) = (z(t_{i-1}), \ldots, z(t_{i-k}))$ defines k preceding states of the process. The *entropy rate* of this process is then defined as (see [CT91])

$$dH(Z) = H(z(t_i), z(t_p)) - H(z(t_p)) = H(z(t_i)|z(t_p)). \tag{A.28}$$

From this definition, it is easy to see that the entropy rate provides a measure for the amount of entropy generated by the stochastic process Z per time step.
If we now consider the case of two m-order Markov processes Z_1 and Z_2 the *joint entropy rate* is given by

$$dH(Z_1, Z_2) = H(z_1(t_i), z_2(t_i)|z_1(t_p), z_2(t_p)), \tag{A.29}$$

and can – by analogy with the equation A.19 – be rewritten as (see Fig. A.2)

$$dH(Z_1, Z_2) = dH(Z_1|Z_2) + dH(Z_2|Z_1) - dI(Z_1, Z_2). \tag{A.30}$$

where $dH(Z_1|Z_2)$ and $dH(Z_2|Z_1)$ denote *conditional entropy rates* of the process Z_1 with an additional condition that the past of the process Z_2 is also given, i.e., $dH(Z_1|Z_2) = H(z_1(t_i)|z_1(t_p), z_2(t_p))$ and *vice-versa*. As one can see from the Figure A.2 the conditional entropy rate thus characterizes the amount of net entropy that is generated per time step for instance in the process Z_1 only, i.e., when the past of the process Z_2 is given. The difference between the entropy rate $dH(Z_1)$ of the process Z_1 and its conditional entropy rate $dH(Z_1|Z_2)$ provides a measure for the amount of entropy that has been transfered per time step from the process Z_2 to the process Z_1. The *transfer entropy* from the stochastic process Z_2 to Z_1 is then defined as

$$TE(Z_2, Z_1) = dH(Z_1) - dH(Z_1|Z_2) = H(z_1(t_i)|z_1(t_p)) - H(z_1(t_i)|z_1(t_p), z_2(t_p)) \tag{A.31}$$

and similarly from the process Z_1 to Z_2 as

$$TE(Z_1, Z_2) = dH(Z_2) - dH(Z_2|Z_1) = H(z_2(t_i)|z_2(t_p)) - H(z_2(t_i)|z_1(t_p), z_2(t_p)). \tag{A.32}$$

According to this definition (also see Fig. A.2) the transfer entropy is an asymmetric measure of interdependence between two stochastic processes, i.e., $TE(Z_1, Z_2) \neq TE(Z_2, Z_1)$.
The term $dI(Z_1, Z_2)$ in the equation A.30 denotes the so called *mutual information rate* between both processes. Formally, the mutual information rate is defined as a conditional mutual information between current states $z_1(t_i)$ and $z_2(t_i)$ with a condition that the complete history of both process $z_1(t_p)$ and $z_2(t_p)$ is given, i.e.,

$$dI(Z_1, Z_2) = I(z_1(t_i), z_2(t_i)|z_1(t_p), z_2(t_p)). \tag{A.33}$$

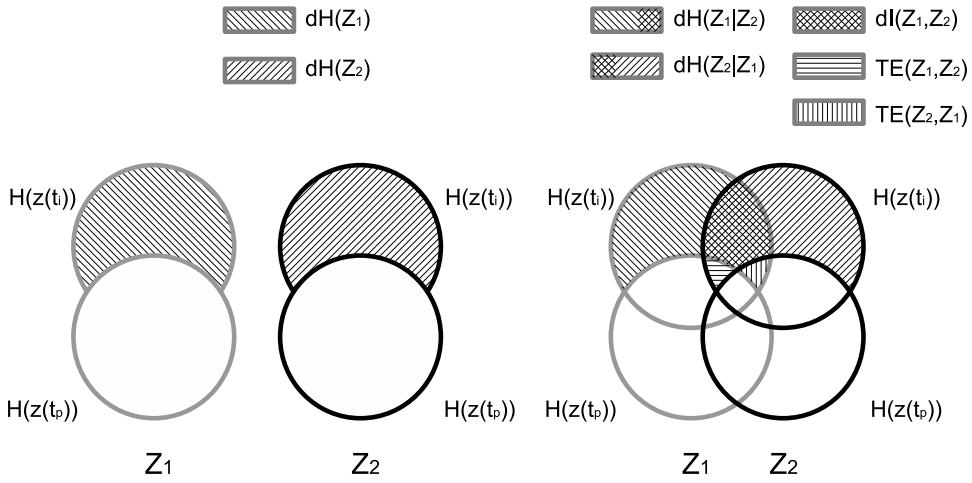


Figure A.2.: Schematic representation of two independent (left) and interacting (right) stochastic ergodic $m^{th}-$order Markov processes Z_1 and Z_2. Here, $H(z(t_i))$ denotes the entropy of the current state whereas $H(z(t_p)) = H((z(t_{i-1}), \ldots, z(t_{i-m})))$ denotes the joined entropy of m preceding states of either process Z_1 or Z_2. The hatched areas correspond to the entropy rates $dH(Z_1)$ and $dH(Z_1)$, the conditional entropy rates $dH(Z_1|Z_2)$ and $dH(Z_2|Z_1)$, transfer entropies $TE(Z_1, Z_2)$ and $TE(Z_2, Z_1)$ as well as the mutual information rate $dI(Z_1, Z_2)$ between Z_1 and Z_2. Further details are given in text.

According to this definition (also see Fig. A.2) the mutual information rate characterizes the amount of common information (entropy) that is generated per time step between processes Z_1 and Z_2. It is easy to see that – similarly to the two-point mutual information (Eq. A.20) – the mutual information rate is a symmetric measure of interdependence between stochastic processes, i.e., $dI(Z_1, Z_2) = dI(Z_2, Z_1)$.
Finally, we can represent the joint entropy rate of the bivariate stochastic processes (Z_1, Z_1) as

$$dH(Z_1, Z_2) = dH(Z_1) + dH(Z_2) - TE(Z_1, Z_2) - TE(Z_2, Z_1) - dI(Z_1, Z_2) \qquad \text{(A.34)}$$

where the transfer entropies $TE(Z_1, Z_2)$ and $TE(Z_2, Z_1)$ represent the asymmetric part and the mutual information rate $dI(Z_1, Z_2)$ the symmetric part of two-point interactions between two stochastic processes.

A.4. Measuring symbolic transfer entropy

Let us denote by x^1_n and x^2_n $(n = 1, \ldots, N)$ a pair of time series of length N and by π^1_n and π^2_n $(n = 1, \ldots, \tilde{N};\ \tilde{N} = N - (m-1)\tau)$ a pair of series of corresponding permutation symbols which form strings S_1 and S_2. According to the definitions given in [Sch00, SL08] (see also Eqs. A.31 and A.32) we can compute the symbolic transfer entropy from S_1 to S_2 of order k as

$$TE^k(S_1, S_2) = \frac{1}{m-1}[H(\pi^2_{i+1}, \ldots, \pi^2_{i-k+1}) - H(\pi^2_i, \ldots, \pi^2_{i-k+1}) \\ -H(\pi^2_{i+1}, \pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1}) + H(\pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1})], \tag{A.35}$$

where $i \in [1, \tilde{N} - k]$. The symbolic transfer entropy in the opposite direction $TE^k(S_2, S_1)$, i.e., from the string S_2 to S_1 is defined analogously. In order to compute all block entropies we used plug-in estimators for the corresponding empirical probability functions [HSPVB07]. To do this we counted the relative frequencies of corresponding blocks of symbols to occur in the string S_2 to S_1, i.e.,

$$\hat{p}(\pi^2_{i+1}, \ldots, \pi^2_{i-k+1}) = \frac{W(\pi^2_{i+1}, \ldots, \pi^2_{i-k+1})}{\tilde{N} - k}, \tag{A.36}$$

$$\hat{p}(\pi^2_i, \ldots, \pi^2_{i-k+1}) = \frac{W(\pi^2_i, \ldots, \pi^2_{i-k+1})}{\tilde{N} - k},$$

$$\hat{p}(\pi^2_{i+1}, \pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1}) = \frac{W(\pi^2_{i+1}, \pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1})}{\tilde{N} - k},$$

and

$$\hat{p}(\pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1}) = \frac{W(\pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1})}{\tilde{N} - k},$$

where $W(\pi^2_{i+1}, \ldots, \pi^2_{i-k+1})$ and $W(\pi^2_i, \ldots, \pi^2_{i-k+1})$ denote the number of occurrences of blocks of symbols $(\pi^2_{i+1}, \ldots, \pi^2_{i-k+1})$ and $(\pi^2_i, \ldots, \pi^2_{i-k+1})$ in string S_2.
$W(\pi^2_{i+1}, \pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1})$ and $W(\pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1})$ denote the number of occurrences of blocks of symbols $(\pi^2_{i+1}, \pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1})$ and $(\pi^2_i, \pi^1_i, \ldots, \pi^2_{i-k+1}, \pi^1_{i-k+1})$ in the bivariate string $S_{21} = \{\pi^2_i, \pi^1_i\}_{i=1}^{\tilde{N}-k}$.

Bibliography

[Abd06] H. Abdi. Bonferroni and sidak corrections for multiple comparisons. In N.J. Salkind, editor, *Encyclopedia of Measurement and Statistics*. Sage Publications, 2006.

[AK07] J.M. Amigó and M.B. Kennel. Topological permutation entropy. *Physica D*, 231:137 – 142, 2007.

[AKK05] J.M. Amigó, M.B. Kennel, and L. Kocarev. The permutation entropy rate equals the metric entropy rate for ergodic information sources and ergodic dynamical systems. *Physica D*, 210:77 – 95, 2005.

[BBS09] L. Barnett, A.B. Barrett, and A.K. Seth. Granger causality and transfer entropy are equivalent for gaussian variables. *Phys. Rev. Lett.*, 103:238701, 2009.

[BKK04] K. J. Blinowska, R. Kuś, and M. Kamiński. Granger causality and information flow in multivariate processes. *Phys. Rev. E*, 70:050902(R), 2004.

[BKP02] C. Bandt, G. Keller, and B. Pompe. Entropy of interval maps via permutations. *Nonlinearity*, 15:1595–1602, 2002.

[Bl89] H. Bai-lin. *Elementary Symbolic Dynamics*. World Scientific, Singapore, 1989.

[BP02] C. Bandt and B. Pompe. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.*, 88:174102, 2002.

[BSLZ01] E.M. Bollt, Th. Stanford, Y.-C. Lai, and K. Zyczkowski. What symbolic dynamics do we get with a misplaced partition? : On the validity of threshold crossings analysis of chaotic time-series. *Physica D*, 154:259 – 286, 2001.

[Buz02] G. Buzsáki. Theta oscillations in the hippocampus. *Neuron*, 33:325–340, 2002.

[Buz06] G. Buzsáki. *Rhythms of the brain*. Oxford University Press, USA, 2006.

[CGG89] T.M. Cover, P. Gacs, and R.M. Gray. Kolmogorov's contributions to information theory and algorithmic complexity. *Ann. Probabil.*, 17:840–865, 1989.

[CT91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Jon Wiley and Sons, New York, 1991.

[DCB06] M. Ding, Y. Chen, and S.L. Bressler. Granger causality: Basic theory and application to neuroscience. In B. Schelter, M. Winterhalder, and J. Timmer, editors, *Handbook of Time Series Analysis*. Wiley, Wienheim, 2006.

[DFT03] C.S. Daw, C.E.A. Finney, and E.R. Tracy. A review of symbolic analysis of experimental data. *Rev. Sci. Instrum.*, 74:915–930, 2003.

[dMvS93] W. de Melo and S. van Strien. *One-dimensional Dynamics*. Springer, New York, 1993.

[DSSW06] J. S. Duncan, J. W. Sander, S. M. Sisodiya, and M. C. Walker. Adult epilepsy. *Lancet*, 367:1087–1100, 2006.

[EFS98] W. Ebeling, J. Freund, and F. Schweitzer. *Komplexe Strukturen: Entropie und Information*. Teubner, Stuttgart, 1998.

[Eic00] H. Eichenbaum. A cortical-hippocampal system for declarative memory. *Nat. Rev. Neurosci.*, 1:41–50, 2000.

[Ein05] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys. Chem.*, 17:549–560, 1905.

[EP97] J. Engel and T. A. Pedley. *Epilepsy:* A *Comprehensive Textbook*. Lippincott-Raven, Philadelphia, 1997.

[ER85] J.P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.*, 57:617–656, 1985.

[EST01] W. Ebeling, R. Steuer, and M. R. Titchener. Partition-based entropies of deterministic and stochastic maps. *Stochastics and Dynamics*, 1:1 – 17, 2001.

[FEG+99] G. Fernández, A. Effern, T. Grunwald, N. Pezer, K. Lehnertz, M. Dümpelmann, D. Van Roost, and C. E. Elger. Real-time tracking of memory formation in the human rhinal cortex and hippocampus. *Science*, 285:1582–1585, 1999.

[FKL+01] J. Fell, P. Klaver, K. Lehnertz, T. Grunwald, C. Schaller, C. E. Elger, and G. Fernandez. Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. *Nat. Neurosci.*, 4:1259–1264, 2001.

[FP07] S. Frenzel and B. Pompe. Partial mutual information for coupling analysis of multivariate time series. *Phys. Rev. Lett.*, 99:204101, 2007.

[GK85] P. Grassberger and H. Kantz. Generating partitions for the dissipative hénon map. *Phys. Lett. A*, 113:235 – 238, 1985.

[GP83a] P. Grassberger and I. Procaccia. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. A*, 28:2591, 1983.

[GP83b] P. Grassberger and I. Procaccia. On the characterization of strange attractors. *Phys. Rev. Lett.*, 50:346, 1983.

[Gra83a] P. Grassberger. Generalized dimensions of strange attractors. *Phys. Lett. A*, 97:227, 1983.

[Gra83b] P. Grassberger. On the fractal dimension of the hénon attractor. *Phys. Lett. A*, 97:224, 1983.

[Gra88] P. Grassberger. Finite sample corrections to entropy and dimension estimates. *Phys. Lett. A*, 128:369, 1988.

[Gra01] C. W. J. Granger. Testing for causality: a personal viewpoint. In *Essays in econometrics: collected papers of Clive W. J. Granger*, pages 48–70, 2001.

[H76] M. Hénon. A two-dimensional mapping with a strange attractor. *Commun. Math. Phys.*, 50:69, 1976.

[HKS99] R. Hegger, H. Kantz, and T. Schreiber. Practical implementation of nonlinear time series methods: the TISEAN package. *Chaos*, 9:413, 1999.

[Hol90] Ph. Holmes. Poincaré, celestial mechanics, dynamical-systems theory and chaos. *Phys. Rep.*, 193:137 – 163, 1990.

[Hon02] J. Honerkamp. *Stochastical Physics*. Springer, Berlin, second edition, 2002.

[HSE94] H. Herzel, A.O. Schmitt, and W. Ebeling. Finite sample effects in sequence analysis. *Chaos, Solitons and Fractals*, 4:97 – 113, 1994.

[HSPVB07] K. Hlavácková-Schindler, M. Palus, M. Vejmelka, and J. Bhattacharya. Causality detection based on information-theoretic approaches in time series analysis. *Phys. Rep.*, 441:1–46, 2007.

[Huy73] C.H. Huygens. *Horologium Oscillatorium*. Parisiis: Apud F. Muguet, 1673.

[Izh07] E. M. Izhikevich. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press, Cambridge, MA, 2007.

[JW07] D. Ji and M. A. Wilson. Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat. Neurosci.*, 10:100–107, 2007.

[Kol59] A.N. Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. *Dokl. Akad. Nauk SSSR*, 124:754, 1959.

[Kol65] A.N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems Inform. Transmission*, 1:1–7, 1965.

[Kol68] A.N. Kolmogorov. Logical basis for information theory and probability theory. *IEEE Trans Inform Theor*, 14:662–664, 1968.

[Kre99] T. Kreuz. Symbolische Dynamik in nichtlinearen Modellsystemen und Zeitreihen hirnelektrischer Aktivität. Master's thesis, Math. Nat. Fakultät der Universität Bonn, 1999.

[KS02] A. Kaiser and T. Schreiber. Information transfer in continuous processes. *Physica D*, 166:43–62, 2002.

[KS03] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis, 2nd ed.* Cambridge Univ. Press, Cambridge, UK, 2003.

[KSJ00] E.R. Kandel, J.H. Schwartz, and T.M. Jessell. *Principles of Neural Science*. McGraw-Hill Publishing Co., 2000.

[Lan08] P. Langevin. Sur la théorie du mouvement brownien. *Comptes Rendus*, 146:530–533, 1908.

[Lan64] L. Landau. *Statistical Physics*. Nauka, Moscow, 1964.

[Lan88] L. Landau. *Mechanics*. Nauka, Moscow, 1988.

[Leh08] K. Lehnertz. Epilepsy and nonlinear dynamics. *J. Biol. Phys.*, 34:253–266, 2008.

[LK05] X. San Liang and R. Kleeman. Information transfer between dynamical system components. *Phys. Rev. Lett.*, 95:244101, 2005.

[LMO+07] K. Lehnertz, F. Mormann, H. Osterhage, A. Müller, A. Chernihovskyi, M. Staniek, J. Prusseit, D. Krug, S. Bialonski, and C. E. Elger. State-of-the-art of seizure prediction. *J. Clin. Neurophysiol.*, 24:147–153, 2007.

[Lor63] E. N. Lorenz. Deterministic non-periodic flow. *J. Atmos. Sci.*, 20:130 – 141, 1963.

[Lub09] E. Lubenov. Hippocampal theta oscillations are travelling waves. *Nature*, 459:534–539, 2009.

[LV08] M. Li and P. Vitanyi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer Verlag, third edition, 2008.

[LZ76] A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Trans. Inform. Theor.*, 22:75–81, 1976.

[MAK+03] F. Mormann, R. Andrzejak, T. Kreuz, C. Rieke, P. David, C. E. Elger, and K. Lehnertz. Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients. *Phys. Rev. E*, 67:021912, 2003.

[Mat00] H. Matsuda. Physical nature of higher-order mutual information: Intrinsic correlations and frustration. *Phys. Rev. E*, 62:3096–3102, 2000.

[MBG+05] M. Müller, G. Baier, A. Galka, U. Stephani, and H. Muhle. Detection and characterization of changes of the correlation structure in multivariate time series. *Phys. Rev. E*, 71:046116, 2005.

[MBRS08] M. Müller, G. Baier, C. Rummel, and K. Schindler. Estimating the strength of genuine and random correlations in non-stationary multivariate time series. *Europhys. Lett.*, 84:10009, 2008.

[MFA+05] F. Mormann, J. Fell, N. Axmacher, B. Weber, K. Lehnertz, C. E. Elger, and G. Fernández. Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. *Hippocampus*, 15:890–900, 2005.

[MK02] R. Marschinski and H. Kantz. Analysing the information flow between financial time series. *Eur. Phys. J. B*, 30:275 – 281, 2002.

[MLDE00] F. Mormann, K. Lehnertz, P. David, and C. Elger. Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Physica D*, 144:358–369, 2000.

[MOA+08] F. Mormann, H. Osterhage, R.G. Andrzejak, B. Weber, G. Fernández, J. Fell, C. Elger, and K. Lehnertz. Independent delta/theta rhythms in the human hippocampus and entorhinal cortex. *Front. Hum. Neurosci.*, 2:1–6, 2008.

[Mor03] F. Mormann. *Synchronization phenomena in the human epileptic brain.* Dissertation. Math. Nat. Fakultät der Universität Bonn, 2003.

[MP00] G.G. Malinetskii and A.B. Potapov. *Modern problems of nonlinear dynamics.* URSS publishers, Moscow, 2000.

[OMWL07] H. Osterhage, F. Mormann, T. Wagner, and K. Lehnertz. Measuring the directionality of coupling: phase versus state space dynamics and application to EEG time series. *Int. J. Neural. Syst.*, 17:139–148, 2007.

[OMWL08] H. Osterhage, F. Mormann, T. Wagner, and K. Lehnertz. Detecting directional coupling in the human epileptic brain: Limitations and potential pitfalls. *Phys. Rev. E*, 77:011914, 2008.

[OS09] A.V. Oppenheim and R.W. Schafer. *Discrete-Time Signal Processing.* Prentice Hall, third edition, 2009.

[Ost08] H. Osterhage. *Messungen der Stärke und Richtung von Interaktionen im epileptischen Gehirn des Menschen.* Dissertation. Math. Nat. Fakultät der Universität Bonn, 2008.

[Ott94] E. Ott. *Chaos in Dynamical Systems.* Cambridge University Press, Cambridge, USA, 1994.

[PRK01] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization – A universal concept in nonlinear sciences.* Cambridge University Press, Cambridge, UK, 2001.

[PTVF02] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, England, second edition, 2002.

[PV07] M. Paluš and M. Vejmelka. Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections. *Phys. Rev. E*, 75:056211, 2007.

[QAG00] R. Quian Quiroga, J. Arnhold, and P. Grassberger. Learning driver-response relationships from synchronization patterns. *Phys. Rev. E*, 61:5142 – 5148, 2000.

[RMS08] C. Rummel, M. Müller, and K. Schindler. Data-driven estimates of the number of clusters in multivariate time series. *Phys. Rev. E*, 78:066703, 2008.

[Roe76] O.E. Roessler. An equation for continuous chaos. *Phys. Lett. A*, 57:397–398, 1976.

[Rou99] M.S. Roulston. Estimating the errors on measured entropy and mutual information. *Physica D*, 125:285 – 294, 1999.

[Sch00] T. Schreiber. Measuring information transfer. *Phys. Rev. Lett.*, 85:461–464, 2000.

[SG96] T. Schürmann and P. Grassberger. Entropy estimation of symbol sequences. *Chaos*, 6:414–427, 1996.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technol. J*, 27:379–423, 1948.

[Sin59] Ya.G. Sinai. *Dokl. Akad. Nauk SSSR*, 124:768, 1959.

[SL08] M. Staniek and K. Lehnertz. Symbolic transfer entropy. *Phys. Rev. Lett.*, 100:158101, 2008.

[SL09] M. Staniek and K. Lehnertz. Symbolic transfer entropy: inferring directionality in biosignals. *Biomed. Tech.*, 54:323 – 328, 2009.

[SS85] M. Sano and Y. Sawada. Measurement of Lyapunov spectrum from a chaotic time series. *Phys. Rev. Lett.*, 55:1082 – 1085, 1985.

[Ste75] S. D. Stearns. *Digital Signal Analysis*. Hayden Book Company, inc., 1975.

[Str01] S.H. Strogatz. *Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering (Studies in nonlinearity)*. Perseus Books Group, Cambridge, 2001.

[SWD+06] B. Schelter, M. Winterhalder, R. Dahlhaus, J. Kurths, and J. Timmer. Partial phase synchronization for multivariate synchronizing systems. *Phys. Rev. Lett.*, 96:208103, 2006.

[Tak81] F. Takens. Detecting strange attractors in turbulence. In D. A. Rand and L.-S. Young, editors, *Dynamical Systems and Turbulence (Warwick 1980)*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer-Verlag, Berlin, 1981.

[TK10] H. Teramoto and T. Komatsuzaki. How does a choice of markov partition affect the resultant symbolic dynamics? *Chaos*, 20:037113, 2010.

[vK07] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library, Amsterdam, third edition, 2007.

[WAL+10] T. Wagner, N. Axmacher, K. Lehnertz, C. E. Elger, and J. Fell. Sleep-dependent directional coupling between human neocortex and hippocampus. *Cortex*, 46:256–263, 2010.

[WSSV85] A. Wolf, J.B. Swift, H.L. Swinney, and J.A. Vastano. Determining lyapunov exponents from a time series. *Physica D*, 16:285 – 317, 1985.

[WST07] M. Winterhalder, B. Schelter, and J. Timmer. Detecting coupling directions in multivariate oscillatory systems. *Int. J. Bifurcation Chaos Appl. Sci. Eng.*, 17:3725–3739, 2007.

[ZL77] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Trans. Inform. Theor.*, 23:337–343, 1977.

[ZRB05] S. Zozor, P. Ravier, and O. Buttelli. On lempel-ziv complexity for multidimensional data analysis. *Physica A*, 345:285 – 302, 2005.

Eigene Veröffentlichungen

Journalartikel

[CMM+05] A. Chernihovskyi, F. Mormann, M. Müller, C. E. Elger, G. Baier, and K. Lehnertz. EEG analysis with nonlinear excitable media. *J. Clin. Neurophysiol.*, 22:314–329, 2005.

[SCML05] R. Sowa, A. Chernihovskyi, F. Mormann, and K. Lehnertz. Estimating phase synchronization in dynamical systems using cellular nonlinear networks. *Phys. Rev. E*, 71:061926, 2005.

[CL07] A. Chernihovskyi and K. Lehnertz. Measuring synchronization with nonlinear excitable media. *Int. J. Bifurcation Chaos Appl. Sci. Eng.*, 17:3425–3429, 2007.

[LMO+07] K. Lehnertz, F. Mormann, H. Osterhage, A. Müller, A. Chernihovskyi, M. Staniek, J. Prusseit, D. Krug, S. Bialonski, and C. E. Elger. State-of-the-art of seizure prediction. *J. Clin. Neurophysiol.*, 24:147–153, 2007.

[CEL09] A. Chernihovskyi, C.E. Elger, and K. Lehnertz. Analysis of synchronization phenomena in human electroencephalograms with nonlinear excitable media. *EURASIP Journal on Advances in Signal Processing*, 2009.

Konferenzbeiträge

[CSF+04] A. Chernihovskyi, R. Sowa, S. Florin, C.E. Elger, and K. Lehnertz. Pattern Recognition in Noisy and Non-stationary Time Series with Cellular Neural Networks. In B Shi A. Zarandy T. Roska, M. Gili, editor, *Proc. 8th IEEE International Workshop on Cellular Neural Networks and Applications (CNNA2004)*, Budapest, Ungarn, 2004. Amulett '98 Kft.

[CSN+04] A. Chernihovskyi, R. Sowa, C. Niederhoefer, R. Tetzlaff, C.E. Elger, and K. Lehnertz. Real-time seizure detection with Cellular Neural Networks. Epilepsia 45, 2004.

[SMC+04] R. Sowa, F. Mormann, A. Chernihovskyi, S. Florin, C.E. Elger, and K. Lehnertz. Estimating synchronization in brain electrical activity from epilepsy patients with cellular neuronal networks. In B Shi A. Zarandy T. Roska, M. Gili, editor, *Proc. 8th IEEE International Workshop on Cellular Neural Networks and Applications (CNNA2004)*, Budapest, Ungarn, 2004. Amulett '98 Kft.

[CTE+04] A. Chernihovskyi, R. Tetzlaff, C.E. Elger, and K. Lehnertz. Detection of transient spectral patterns and synchronization phenomena in EEG with nonlinear excitable media. Epilepsia 47, 2006.

[CEL06] A. Chernihovskyi, C.E. Elger, and K. Lehnertz. Effect of inhibitory diffusive coupling on frequency-selectivity of excitable media simulated with cellular neural networks. In V. Tavsanoglu and S. Arik, editors, *Proc. 10th IEEE International Workshop on Cellular Neural Networks and their Applications*, pages 292–296. IEEE-Press, 2006.

[DCM+06] F. Döhler, Chernihovskyi, F. Mormann, C.E. Elger, and K. Lehnertz. Detecting structural alterations in the brain using a Cellular Neural Network based classification of magnetic resonance images. In Arik S Tavsanoglu V, editor, *Proc of the 2006 10th IEEE International Workshop on Cellular Neural Networks and their Applications.*, pages 297–300. IEEE-Press, 2006. Catalog No: 06TH8915, ISBN 1-4244-0639-0.

[KCO+06] D. Krug, A. Chernihovskyi, H. Osterhage, C. E. Elger, and K. Lehnertz. Estimating generalized synchronization in brain electrical activity from epilepsy patients with cellular nonlinear networks. In V. Tavsanoglu and S. Arik, editors, *Proc. 10th IEEE International Workshop on Cellular Neural Networks and their Applications*, pages 68–72. IEEE-Press, 2006.

[CSW+04] A. Chernihovskyi, M. Staniek, L. Wang, C.E. Elger, and K. Lehnertz. Symbolic Analysis of Multivariate Data. 3rd International Workshop on Epileptic Seizure Prediction, 2007.

[CEL08] A. Chernihovskyi, C.E. Elger, and K. Lehnertz. Analysis of synchronization phenomena in human electroencephalograms with nonlinear excitable media. In Sánchez VMB Vilariño DL, Ferrer DC, editor, *Proceedings of the 2008 11th International Workshop on Cellular Neural Networks and their Applications.*, pages 81–85. IEEE-Press, 2008. Catalog No: CFP08CNN-CDR, ISBN 978-1-4244-2090-2.

[SCEL08] M. Staniek, A. Chernihovskyi, C. E. Elger, and K. Lehnertz. Inferring directional information flows from EEG recordings during the preictal period. *Epilepsia*, 49(suppl. 7):25, 2008.

Buchartikel

[CKEL08] A. Chernihovskyi, D. Krug, C. E. Elger, and K. Lehnertz. Time series analysis with Cellular Neural Networks. In B. Schelter, J. Timmer, and A. Schulze-Bonhage, editors, *Seizure Prediction in Epilepsy. From Basic Mechanisms to Clinical Applications*, pages 131–148. Wiley-VCH, 2008.

Danksagung

An dieser Stelle möchte ich mich bei allen Personen bedanken, die zur Anfertigung dieser Arbeit beigetragen haben. Insbesondere möchte ich Herrn Prof. Dr. K. Lehnertz dafür danken, dass er mir die Möglichkeit gegeben hat, diese interdisziplinäre Arbeit an der Klinik für Epileptologie und am Helmholtz-Institut für Strahlen- und Kernphysik der Universität Bonn durchzuführen. Weiterhin danke ich ihm für die hervorragende Betreuung und Unterstützung während dieser Arbeit. Am Helmholtz-Institut für Strahlen- und Kernphysik möchte ich Herrn Prof. Dr. K. Maier für die Übernahme des Koreferats und die fortwährende Unterstützung danken. Außerdem gilt mein besonderer Dank Dr. Florian Mormann, der mir mittels zahlreicher, sehr inspirierender Gespräche zur Seite gestanden hat. Der gesamten Arbeitsgruppe Neurophysik (inklusive ehemaliger Mitglieder) möchte ich für die überaus freundliche Arbeitsatmosphäre und die große Hilfsbereitschaft danken. Nicht zuletzt danke ich meinen Eltern für die Ermöglichung meines Studiums und ihre stete Unterstützung sowie meiner Freundin Natalja, die immer für alle meine Anliegen viel Zeit und Geduld aufgebracht hat.