Department of Physics
University of Wuppertal

Measuring Synchronization in Model Systems and Electroencephalographic Time Series from Epilepsy Patients

Dissertation (PhD thesis)

by

Thomas Kreuz

September 2003

(WUB-DIS 2003-6)

Summary

The main aim of this dissertation is the comparative investigation of different measures of synchronization derived from various approaches and concepts. These include both measures for estimating the degree of dependence between two time series as well as measures which quantify the directionality of this dependence. The first group comprises the linear cross correlation, mutual information, six different indices for phase synchronization (based either on the Hilbert or on the wavelet transform) as well as symmetrized variants of two nonlinear interdependence measures and of event synchronization. The antisymmetrized variants of the last three measures form the group of measures of directionality.

In the first part of this dissertation the symmetric measures are tested in a controlled setting by means of various model systems. Using the coupling strength as a first control parameter it is investigated to which extent the different measures are able to distinguish between different degrees of dependence. Furthermore, the robustness of the measures against external noise is estimated by varying the signal-to-noise ratio as the second control parameter.

Subsequently, all measures are employed to analyze electroencephalographic recordings from epilepsy patients. This application part consists of two single studies. First a comprehensive comparison on the predictability of epileptic seizures is carried out. Object of investigation is the capability of the different measures to reliably distinguish between the intervals preceding epileptic seizures and the intervals far away from any seizure activity. Already in this study a great deal of attention is paid to the statistical validation of seizure predictions. This issue is particularly addressed in the last part of this dissertation in which the method of measure profile surrogates is introduced as an appropriate tool to distinguish between measures and algorithms unsuited for the prediction of epileptic seizures, and more promising approaches. Two of the measures of synchronization are used to illustrate this new approach.

Zusammenfassung

Hauptziel der vorliegenden Doktorarbeit ist die vergleichende Untersuchung verschiedener Ansätze zur Messung von Synchronisation zwischen zwei Zeitreihen. Diese beinhalten sowohl Maße zur Abschätzung des Grades an Abhängigkeit als auch Maße, welche die Direktionalität dieser Abhängigkeit quantifizieren. Die erste Gruppe umfasst die lineare Kreuzkorrelation, die Mutual Information, sechs verschiedene Indices für Phasensynchronisation, basierend entweder auf der Hilbert-Transformation oder auf der Wavelet-Transformation, sowie symmetrisierte Versionen zweier nichtlinearer Interdependenzmaße und der Event-Synchronisation. Aus den anti-symmetrisierten Versionen der letzten drei Maße setzt sich die Gruppe der Direktionalitätsmaße zusammen.

Im ersten Teil dieser Doktorarbeit werden die Synchronisationsmaße mit Hilfe verschiedener nichtlinearer Modellsysteme getestet. Mit der Kopplungsstärke als erstem Kontrollparameter wird untersucht, wie gut die Maße in der Lage sind, verschiedene Grade der Abhängigkeit zu unterscheiden. Die Robustheit der Maße gegenüber externen Störsignalen wird durch Variation des Signal-Rausch-Verhältnisses als zweitem Kontrollparameter abgeschätzt.

Anschließend werden die verschiedenen Maße zur Analyse elektroenzephalographischer Aufzeichnungen von Epilepsie-Patienten herangezogen. Dieser Anwendungsteil besteht aus zwei Einzelstudien. Zunächst wird eine umfassende Vergleichsstudie zur Vorhersagbarkeit epileptischer Anfälle durchgeführt. Gegenstand der Untersuchung ist die Eignung der Maße zur verläßlichen Trennung der Intervalle vor epileptischen Anfällen von den Intervallen weit weg von jeglicher Anfallsaktivität. Dabei wird bereits besonderes Augenmerk auf das häufig vernachlässigte Problem der statistischen Validierung von Anfallsvorhersagen gelegt. Ausschliesslich mit diesem wichtigen Aspekt befasst sich der letzte Teil dieser Arbeit, in dem die Methode der Maßprofil-Surrogate als geeigneter Lösungsansatz vorgestellt wird. Zwei der in dieser Arbeit untersuchten Synchronisationsmaße werden dazu verwendet, dieses neue Verfahren zu illustrieren.

First referee:	Prof. Dr. Peter Grassberger, University of Wuppertal
Second referee:	Priv. Doz. Dr. Klaus Lehnertz, University of Bonn
Day of submission:	30. September 2003
Day of oral examination:	2. Dezember 2003

Contents

1 Introduction **1**

2 Theoretical background **5**

2.1 Dynamical systems 5

2.2 State space reconstruction 6

2.3 Synchronization 7

2.3.1 Complete synchronization 8

2.3.2 Phase synchronization 8

2.3.3 Generalized synchronization 8

3 Measures of synchronization **11**

3.1 Cross correlation 11

3.2 Mutual information 12

3.3 Phase synchronization 15

3.3.1 Extracting the phase 15

3.3.2 Indices of phase synchronization 16

3.4 Nonlinear interdependencies 19

3.5 Event Synchronization 21

4 Application to coupled model systems **25**

4.1 Methods 26

4.1.1 Coupling schemes 26

4.1.2 Implementation of measures 30

4.1.3 Criterion for comparing different measures 31

4.2 Results 33
4.2.1 Dependence on coupling strength 33
4.2.2 Robustness against noise 37
4.3 Discussion 43

5 Application to the EEG of epilepsy patients 47
5.1 Epilepsy and the electroencephalogram 48
5.2 Prediction of epileptic seizures and its statistical validation 53
5.3 Statistical evaluation of the predictability of seizures 57
5.3.1 Methods 58
5.3.2 Results 67
5.3.3 Discussion 81
5.4 The method of measure profile surrogates 85
5.4.1 Methods 85
5.4.2 Results 90
5.4.3 Discussion 96

6 Summary and Outlook 99

A Nonlinear deterministic systems 103
A.1 Hénon map 103
A.2 Lorenz system 104
A.3 Rössler system 105

B Stochastic signals 106
B.1 White noise 106
B.2 Iso-spectral noise 107

Chapter 1

Introduction

The words *'synchronous'* and *'synchronicity'* originate from a combination of the Greek words $\sigma\upsilon\nu$ (syn = common) and $\chi\rho o\nu o\varsigma$ (chronos = time) and thus can be translated as 'happening at the same time'. Although the term *'synchronization'* shares this etymology, there is a subtle distinction between the notions of these two words. While a 'synchronous motion' of two or more objects is rather vaguely determined by a pure coincidence which in principle could be by chance, 'synchronization' is more rigorously defined as the (active) adjustment of rhythms of different oscillating systems due to some kind of interaction or coupling [123]. Since synchronization generally leads to synchronous motion both terms are often used synonymous.

Synchronization between dynamical systems has been an active field of research in many scientific and technical disciplines since the first description of this phenomenon in the seventeenth century. It was the Dutch scientist Christiaan Huygens who first reported on his observation of synchronization between two pendulum clocks hanging from a common support [49]. In the twentieth century systematic study of synchronization phenomena was started experimentally by Edward Appleton [12] and theoretically by Balthasar van der Pol who derived the van der Pol equation, the first and still most prominent example of a nonlinear self-oscillating system [168]. Van der Pol was also the first to apply oscillation theory to a physiological system, namely the human heart [169]. Starting in the early 1980s and only shortly succeeding the development of the theory of deterministic chaos the notion of synchronization was further extended to the case of interacting chaotic oscillators [37, 120, 2, 116]. Since the definition of chaos implies the rapid decorrelation of nearby orbits due to their high sensitivity on initial conditions, synchronization of two coupled chaotic systems is a highly counter-intuitive phenomenon. Thus it has been extensively studied and applied in many disciplines of physics. Prominent examples include electronics [121, 46, 114], laser dynamics [33, 144, 162], solid state physics [119], plasma physics [136], communication [22, 57] and control [126, 146].

According to the fact that synchronization phenomena can manifest themselves in many different ways, a unifying framework for synchronization in chaotic dynamical systems is

still missing (and might not be achievable at all). Instead various concepts for its description have been offered. The simplest case of complete synchronization can be obtained if identical systems are coupled sufficiently strong so that their states coincide [37, 120]. Phase synchronization, first described for chaotic oscillators in Refs. [139, 122, 108], is defined as the entrainment of phases, whereas the amplitudes remain chaotic and, in general, weakly correlated. Furthermore, the concept of generalized synchronization, introduced for uni-directionally coupled systems [2, 145, 1], denotes the presence of some functional relation between the states of responder and driver. Since this function does not have to be the identity, generalized synchronization is already a rather weak criterion, but it is still surpassed by the notion of interdependence [14] where the mapping of local neighborhoods in the first system onto local neighborhoods in the second system is exploited as the quantifying criterion.

Corresponding to and extending this variety of concepts, many different approaches aiming at a quantification of the degree of synchronization between two systems have been proposed. These approaches comprise linear ones like the cross correlation or the coherence function as well as essentially nonlinear measures like mutual information [43]. Furthermore, different indices of phase synchronization have been introduced [163, 98]. Here the instantaneous phases are extracted from the time series by using e.g., the Hilbert transform [139] or the wavelet transform [71]. Topological approaches to quantify generalized synchronization include the method of mutual false nearest neighbors [145] and the index based on nonlinear mutual predictions [148] as well as more recent measures like the nonlinear interdependencies [14] and synchronization likelihood [160]. Finally, the measure event synchronization [130] quantifies the over-all level of synchronicity from the number of quasi-simultaneous appearances of certain predefined events. While most of these measures are only designed to estimate the degree of synchronization between two systems, three of these measures, the nonlinear interdependencies and event synchronization, are also able to reveal possible directionalities between them and thus, at least in principle, to detect driver-responder relationships.

In the literature on the quantification of synchronization almost exclusively one single measure is applied either to model systems or to real data. Only rarely different measures are used to analyze the same system and thus a comprehensive comparison of all these different approaches in a ‘controlled setting’ is still missing. This task is addressed in the first part of this thesis in which the measures of synchronization are applied to different coupled model systems. The aim is to evaluate to which extent the analysis of these model systems can render information about the different measures of synchronization useful for a later application of these measures to field data. To address this aim, the coupling strength as well as the signal-to-noise ratio serve as control parameters. With the coupling strength as the first parameter it is tested to which extent the different measures are able to distinguish between different levels of coupling. This property is essential in most applications since rarely the absolute value of synchronization is of interest but rather it is the change of synchronization between two different states, times, or recording sites that matters. The

second parameter, the signal-to-noise ratio, is used to investigate whether the results of the different measures prove to be robust when the signals of interest are contaminated with a certain level of noise. Robustness against such contaminations is a very important prerequisite for the application of these measures to field data, since noise is an inevitable disturbance in any measurement setting.

A prominent example for the acquisition of field data and one of the most important applications for measures of synchronization is the study of human electroencephalographic (EEG) signals [90], an outstanding example for the acquisition of such data is the presurgical diagnostics of epilepsy patients. To yield sufficient information for the unequivocal localization of the seizure-generating structure (epileptic focus), often multichannel recordings are acquired using intracranial monitoring techniques, in which the brain electrical activity is recorded directly from the surface of the brain and from specific structures within the brain [32]. The excellent signal to noise ratio and the outstanding temporal and spatial resolution of these data allow their meaningful investigation by means of linear and nonlinear time series analysis techniques in order to further understand the spatiotemporal dynamics of the epileptic brain [83, 82]. Bivariate measures of synchronization seem to be particularly well suited for this purpose, since synchronization phenomena have been increasingly recognized as a key feature for establishing the communication between different regions of the brain [170, 34, 171]. Furthermore, abnormal synchronization of neuronal ensembles is regarded as the main mechanism responsible for the generation of epileptic seizures [90].

Probably the most challenging task in the analysis of EEG recordings from epilepsy patients is the prediction of epileptic seizures (cf. Refs. [87, 89, 86]). Despite the aforementioned interrelations at first mostly univariate measures have been used to address this issue, and it is only recently that the focus of attention starts to shift towards bivariate measures and in particular to the question whether measures of synchronization can render valuable information enabling the prediction of epileptic seizures. To address this question, a comprehensive comparison of the different measures with respect to their capability to discriminate the intervals preceding seizures from the intervals far away from any seizure activity is carried out. Furthermore, in order to investigate to which extent the different measures of synchronization and directionality carry independent and non-redundant information, the correlation between these measures is estimated.

Many of the studies dealing with seizure prediction suffer from a severe lack of statistical validation. Only rarely results are passed to a statistical test and are verified against some null hypothesis H_0 in order to quantify their significance. This issue has first been addressed by the method of seizure times surrogates proposed by Andrzejak and colleagues [9]. In the last study of this thesis the method of measure profile surrogates is introduced as a new and complementary approach. The method is illustrated by statistically validating the predictive performance of two of the measures of synchronization.

This thesis is organized as follows: First in Chapter 2 some theoretical background about dynamical systems and their analysis is given along with an introduction to the different

notions of synchronization. In Chapter 3 a representative selection of different approaches to quantify the degree of synchronization between two systems are introduced. These measures are applied to different coupled model systems in Chapter 4 in which they are compared with respect to their capability to reflect the strength of coupling and their robustness against noise. In Chapter 5 they are applied to electroencephalographic time series measured from the brain of epilepsy patients (Section 5.1). In particular the issue of epileptic seizure prediction and its statistical validation is addressed (Section 5.2). First in Section 5.3 the capability of the different measures to reliably detect a distinct pre-seizure state is evaluated. Furthermore, based on observed correlations between the different measures the combined use of measures is discussed. The aim of statistical validation is further pursued in Section 5.4 in which the new method of measure profile surrogates is demonstrated. The conclusions of this thesis are drawn in Chapter 6.

Chapter 2

Theoretical background

In the main part of this chapter an overview of the various concepts for the description of synchronization is given (Section 2.3). As a theoretical fundament for its understanding the theory of dynamical systems (Section 2.1) and the most important tool in their analysis, the state space reconstruction (Section 2.2), are described before.

2.1 Dynamical systems

Apart from systems with an infinite number of degrees of freedom, the state of a dynamical system can generally be described by D time dependent variables. Assigning each of these system variables to a basis vector in an abstract state space, the instantaneous state of the system is determined by a point in this state space:

$$\overrightarrow{x}(t) = (x_1(t), x_2(t), ..., x_D(t)) \tag{2.1}$$

The series of vectors consecutive in time form the *trajectory* of the system. For *deterministic* dynamics the state of the system in the next instant is unequivocally defined by the present state. If only probabilities for the following state can be given, the dynamics is said to be *stochastic*. The continuous temporal evolution of a dynamical system is a *flow* in the state space, while a discrete dynamics (this includes the equidistant sampling of a continuous trajectory) is called a *map*. In the continuous case the dynamics of a deterministic system can be described by a set of ordinary differential equations:

$$\frac{d\overrightarrow{x}(t)}{dt} = F(\overrightarrow{x}(t)) \tag{2.2}$$

For a discrete deterministic system the transition from time instant t to time instant $t + \triangle t$ is represented by a mapping of the state space onto itself:

$$\overrightarrow{x}(t + \triangle t) = F(\overrightarrow{x}(t)) \tag{2.3}$$

In both cases the temporal evolution of the state variable $\overrightarrow{x}$ is governed by a generally nonlinear function F.

In addition to this, classical mechanics distinguishes between *conservative* systems with Hamiltonian structure, in which, according to the theorem of Liouville, the state space volume occupied by the system is preserved $(divF = 0)$, and *dissipative* systems, which exchange energy with their environment and/or with microscopic degrees of freedom which are not modelled explicitly. Typically, after some initial transients this leads to a contraction of the occupied state space volume $(divF < 0)$ onto a subset with lower dimension which is termed *attractor*. For regular dynamics, this can be a fix point attractor of dimension zero, a one dimensional limit cycle or a torus of dimension equal or larger than two. But due to some stretching- and folding-mechanisms caused by the nonlinear function F, also so-called *strange attractors* can occur. These attractors are distinguished by their *self-similarity* (comparable structures on any scale) and their *fractal* (i.e., non-integer) dimension. The later property is, among others, often used to define a *chaotic system*. In this context, it is important to note that nonlinearity is a necessary but not sufficient condition for chaos.

2.2 State space reconstruction

Given a D-dimensional system with known dynamics, the equations of motions (2.2) or (2.3) allow, starting from an initial state $\overrightarrow{x}(0)$, the numerical determination of the temporal evolution of all state variables. According to the definition of the state space the trajectory $\overrightarrow{x}(t)$ gives a complete characterization of the dynamics of the system. In a typical experimental setup mostly only scalar time series are available describing the temporal evolution of a possibly high dimensional system. Nevertheless the principal possibility to reconstruct the fundamental dynamics from a single time series was shown in the embedding theorems in Refs. [161, 147]. The method of state space reconstruction via time delay embedding relies on the assumption that in the limit of an infinite number of data points and without any noise the influence of all other system variables is reflected in the temporal evolution of the measured one[1].

From a univariate discrete time series x_n, $n = 1, \ldots, N$, of a dynamical system X delay vectors can be reconstructed via

$$\overrightarrow{x}_n = (x_n, \ldots, x_{n-(m-1)\tau}) \tag{2.4}$$

[1] In case of multichannel recordings in principle a state space reconstruction using spatial embedding techniques is possible. In practice the uncertainties in the choice of a spatial delay and other related problems make this option infeasible (cf. Ref. [58]).

with m denoting the embedding dimension and τ denoting the time lag. To avoid self crossings of the reconstructed trajectory due to spurious projections, the embedding theorem of Whitney [175] demands an embedding dimension of

$$m \geq 2D + 1. \tag{2.5}$$

Furthermore, the minimization of obstructive correlations in the state space can be achieved by setting the time lag τ to the first zero crossing of the autocorrelation function [20] or the first local minimum of the mutual information function [36]. In this way it is possible using a single time series to reconstruct a state space topology which is equivalent to the state space spanned by all system variables.

2.3 Synchronization

Synchronization is rigorously defined as the active adjustment of rhythms of different oscillating systems due to some kind of interaction or coupling [123]. There are some important physical subtleties and caveats within this definition. First of all it should be possible to divide the system under investigation into different subsystems which can in principle generate independent signals. Thus all cases are excluded where the oscillating variables are just different coordinates of the same system. Instead signals could be measured from truly separated, identical or non-identical systems, but alternatively they could also stem from different parts of the same extended system. Second, synchronization takes place between autonomous systems exhibiting self-sustained oscillations. Without interaction each of these oscillators continues to generate the same steady and stable rhythm until its source of energy expires. With weak interaction the systems adjust their rhythm, a phenomenon occurring more or less persistently even in a certain range of systems' mismatch of parameters. In case of very strong coupling, however, the term synchronization might not be appropriate any more since then the originally independent subsystems now form one new unified system which can not be regarded as decomposable. A similar problem occurs when two systems with identical or almost identical dynamical properties are driven by a common hidden source. In this case they both get synchronized with the common driver and thus they will show a synchronous behavior, but according to the rigorous definition they are not synchronized with each other.

Considering these non-trivial implications in the theoretical treatment of synchronization, it is not surprising that the observation and description of synchronization becomes an even more vague venture. According to the many different ways in which synchronization phenomena can manifest themselves, various concepts for its description have been offered. These concepts which mostly focus on different distinct features will be introduced in the following.

2.3.1 Complete synchronization

The simplest case of complete or identical synchronization can be obtained if two systems X and Y are coupled sufficiently strong so that their states $\overrightarrow{x}(t)$ and $\overrightarrow{y}(t)$ coincide in the limit $t \to \infty$ [37, 120]:

$$\lim_{t\to\infty}[\overrightarrow{x}(t) - \overrightarrow{y}(t)] = \overrightarrow{0} \tag{2.6}$$

This can only be obtained for identical systems. Otherwise, if the parameters of coupled systems slightly mismatch, the states can come close to each other but still remain different. Complete synchronization is included as a special case in all other definitions of synchronization.

2.3.2 Phase synchronization

In principle, the notion of phase synchronization goes back to the observation of interactions between two pendulum clocks by Huygens [49]. It is the natural concept for the description of two coupled linear (harmonic) or nonlinear oscillators or any other systems, where the definition and determination of a phase is obvious. Only recently this concept has also been applied to chaotic oscillators [139, 122, 108] and even further extended to the analysis of almost arbitrary time series (cf. [141]). In most cases the phase is extracted from the time series, e.g., via Hilbert transform. Phase synchronization is then defined as entrainment of the phases of two oscillating systems X and Y:

$$|n\phi_x(t) - m\phi_y(t)| \leq const \tag{2.7}$$

with n and m being integers. While the phases are locked, the amplitudes remain chaotic and, in general, uncorrelated [110]. The special case, when in addition to a strictly constant phase shift the amplitudes become also completely correlated, is referred to as lag synchronization in the literature [140].

2.3.3 Generalized synchronization

Finally, the concept of generalized synchronization, originally introduced for unidirectionally coupled systems [2, 145, 1], but meanwhile also applied to bidirectional coupling schemes [180], denotes the presence of some functional relation between the state variables $\overrightarrow{x}$ and $\overrightarrow{y}$ of driver system X and responder system Y:

$$\overrightarrow{y}(t) = \psi[\overrightarrow{x}(t)]. \tag{2.8}$$

The mathematical properties demanded for the functional ψ vary throughout the literature. Usually existence and smoothness are required as minimum conditions, but sometimes also differentiability and invertibility are considered as prerequisites [117]. The relation between phase synchronization and generalized synchronization is discussed controversially in the literature. First it has been claimed that generalized synchronization implies phase synchronization [114], later for certain cases the opposite relation has been reported [179].

Unidirectionally driven systems are usually said to exhibit generalized synchronization if the largest Lyapunov exponent of the responder (called conditional Lyapunov exponent in Ref. [116]) is negative, although this represents merely a necessary and not a sufficient condition [145].

Chapter 3

Measures of synchronization

In this thesis a representative selection of different bivariate measures quantifying the degree of synchronization between two systems is applied to model systems and electroencephalographic time series measured from the brain of epilepsy patients. These measures comprise symmetric ones like the linear cross correlation, the mutual information and three different indices of phase synchronization (where the phase of the time series is extracted by using either the Hilbert transform or the wavelet transform) as well as anti-symmetric ones like two related approaches quantifying nonlinear interdependencies and the event synchronization. Similarities and differences of these approaches will be described in this Chapter.

In general one should always distinguish between a quantity defined for a mathematical model process and the analogous quantity estimated from finite data (time series) drawn randomly from this process. Since this thesis is dealing with the analysis of time series exclusively, all formulas of measures of synchronization are meant to define the respective estimators. Thus any carets (usually denoting estimators) are omitted for simplicity. Furthermore, in the following and throughout this thesis x and y with samples x_n and y_n $(n = 1, \ldots, N)$ will denote two simultaneously measured discrete univariate time series of length N from two possibly coupled dynamical systems X and Y while $\overrightarrow{x}_n$ and $\overrightarrow{y}_n$ will denote their embedded delay vectors yielded by state space reconstruction (cf. Section 2.2). Before the calculation of these measures all time series are demeaned and normalized to unit variance. Therefore also mean and variance are omitted in all equations for simplicity.

3.1 Cross correlation

The simplest and most commonly used measure of synchronization is the cross correlation defined in the time domain as a function of the time lag $\tau = -(N-1), \ldots, 0, \ldots, N-1$:

$$C_{XY}(\tau) = \begin{cases} \frac{1}{N-\tau} \sum_{n=1}^{N-\tau} x_{n+\tau} y_n & \tau \geq 0 \\ C_{YX}(-\tau) & \tau < 0. \end{cases} \tag{3.1}$$

The cross correlation is obtained by normalization of the cross covariance and thus ranges from minus one (anti-phase synchronization) to one (in-phase synchronization), while intermediate values close to zero are attained for linearly independent systems[1]. The cross correlation is a measure of linear synchronization between X and Y only.

This linear synchronization can also be quantified in the frequency domain using the cross spectrum:

$$C_{XY}(\omega) = (\mathcal{F}x)(\omega) \cdot (\mathcal{F}y)^*(\omega)x, \tag{3.2}$$

where $(\mathcal{F}x)$ is the Fourier transform of x, ω are the discrete frequencies ($-N/2 < \omega < N/2$) and the asterisk denotes complex conjugation. The cross-spectrum is a complex number whose amplitude

$$\Gamma_{XY}(\omega) = \frac{|C_{XY}(\omega)|}{\sqrt{C_{XX}(\omega) \cdot C_{YY}(\omega)}} \tag{3.3}$$

is called the coherence function. It is normalized by the autocorrelation functions of the two systems. As a function of the frequency ω this is a very useful measure when one is interested in the synchronization related to certain frequency ranges only, e.g., in the classical EEG frequency bands (cf. Ref. [90] and Section 5.1). These are two equivalent representations related via the correlation theorem and thus the cross correlation function can be calculated as the inverse Fourier transform of the cross spectrum.

The first measure used in this thesis is the maximum cross correlation which is symmetric in X and Y:

$$C_{max} = \max_{\tau}\{|C_{XY}(\tau)|\}. \tag{3.4}$$

3.2 Mutual information

In contrast to cross correlation or coherence which are measures of linear dependencies only, mutual information quantifies nonlinear dependencies as well, i.e., it is zero if and

[1] Even completely uncoupled time series can spuriously attain non-zero estimated cross-correlation values due to high values of the autocorrelation function (cf. Ref. [106]). The Bartlett estimator, a significance threshold for observed cross-correlation values, can be found in Refs. [17, 19].

only if the two time series are strictly independent. As a measure derived from information theory [43, 24] it is based on the Shannon entropy and therefore allows a straightforward interpretation (in contrast to quantities based on higher order Renyi entropies [131], which, however, are often easier to estimate).

Performing a binning of the state spaces of two systems X and Y and denoting by $p_x(i)$ [$p_y(j)$] the weight of the i-th [j-th] bin in X-space [Y-space], the mutual information is defined as

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \tag{3.5}$$

with $H(X)$ and $H(Y)$ denoting the Shannon entropies of the respective marginal distributions $p_x(i), p_y(j)$ with $i = 1, \ldots, M_x, j = 1, \ldots, M_y$, e.g.,

$$H(X) = -\sum_{i=1}^{M_x} p_x(i) \ \log \ p_x(i) \tag{3.6}$$

and $H(X,Y)$ denoting the Shannon entropy of the joint distribution $p_{xy}(i,j)$, i.e.,

$$H(X,Y) = -\sum_{i,j}^{M} p_{xy}(i,j) \ \log \ p_{xy}(i,j). \tag{3.7}$$

The Shannon entropies would diverge for vanishing bin size and continuous X, whereas this limit is finite for $I(X,Y)$ provided X and Y are not deterministically related. In this limit $I(X,Y)$ becomes an integral over densities. While the different entropies measure the information content of the marginal spaces X, Y and the joint space (X,Y), mutual information quantifies the amount of information of X obtained by knowing Y and vice versa. When taking the logarithms with base 2, all quantities are measured in bits. As mentioned above, mutual information is zero if and only if the two time series are independent, while it attains positive values with a maximum of $I(X,X) = H(X)$ for identical signals. Like cross correlation at delay $\tau = 0$, it is symmetric in X and Y, i.e., $I(X,Y) = I(Y,X)$.

Mutual information can also be regarded as a Kullback-Leibler entropy measuring the gain in information when replacing the distribution $p_x(i) \cdot p_y(j)$, yielded under the assumption of independence between X and Y, by the actual joint probability distribution $p_{xy}(i,j)$, i.e.,

$$I(X,Y) = \sum_{i,j}^{M} p_{xy}(i,j) \ \log \frac{p_{xy}(i,j)}{p_x(i)p_y(j)}. \tag{3.8}$$

The easiest and most wide spread approach for estimating mutual information from two time series x and y consists in partitioning their supports into bins of finite size and counting the numbers of points falling into the various bins. This binning can either be performed

using the time series itself or using the reconstructed state spaces rendered by applying the method of time delay embedding to the individual time series before. With $n_x(i)$ and $n_y(j)$ denoting the number of points falling into the i-th bin of X and the j-th bin of Y respectively, and $n_{xy}(i,j)$ as the number of points in their intersection, the marginal and the joint probabilities can be approximated as

$$p_x(i) \approx \frac{n_x(i)}{N}, \quad p_y(j) \approx \frac{n_y(j)}{N}, \quad p_{xy}(i,j) \approx \frac{n_{xy}(i,j)}{N}. \tag{3.9}$$

More sophisticated estimators [36, 25] use adaptive instead of fixed bin sizes to obtain approximately equal numbers $n(i,j)$ in all non-empty bins (i,j). But still these estimates suffer from systematic errors that can in principle be reduced by taking into account the first two terms of a diverging series of finite size corrections [42]. Another approach to calculate mutual information uses first order correlation integrals for the estimation of the single Shannon entropies [41, 115]. This involves the computation of probabilities within neighborhoods of a certain fixed radius around each point.

In this thesis the first of two new and improved estimates of mutual information introduced by Kraskov and colleagues in Ref. [62] (based on considerations of Ref. [129]) is used. Both estimates do not use a fixed neighborhood size, but instead are based on entropy estimates from k-nearest neighbor distances. The first variant uses adaptive (hyper-)cubes whose size is locally adapted in the joint space and then kept equal in the marginal subspaces. Its estimate reads:

$$I(X,Y) = \psi(k) - \langle \psi(n_x+1) + \psi(n_y+1) \rangle + \psi(N), \tag{3.10}$$

for any integer $k \in [1,\ldots,N]$. Here, $\psi(x)$ is the digamma function defined as $\psi(x) = \Gamma(x)^{-1} d\Gamma(x)/dx$. It can easily be calculated using the recursion $\psi(x+1) = \psi(x) + 1/x$ and $\psi(1) = -C$ with $C = 0.5772156\ldots$ the Euler-Mascheroni constant. For large x, $\psi(x) \approx \log x - 1/2x$ holds.

The second variant uses (hyper-)rectangles instead of (hyper-)cubes, i.e., their size is adjusted independently for each of the marginal subspaces. In general it gives very similar results. For both estimates systematic errors increase with k, while statistical errors decrease, with the errors of the first estimate being between those for the second variant with the same k and those for $k' = k+1$. Both estimates are very data efficient (for $k = 1$ structures down to the smallest possible scale are resolved), adaptive (the resolution is adjusted according to the local data density), and have minimal bias. Indeed, the bias of the underlying entropy estimates is mainly due to the non-uniformity of the density at the smallest resolved scale, typically giving errors of the order $O(k/N)$. Numerically, the estimator $I(X,Y)$ proved to become exactly unbiased if the densities of the distributions factorize ($\rho_{x,y} = \rho_x \rho_y$), i.e., it vanishes (up to statistical fluctuations) for independent distributions [62].

Schreiber extended the concept of mutual information and defined the so-called *transfer entropy* [153], which, as an asymmetric measure, proved to be able to in principle distinguish driver-responder relationships. Another asymmetric measure which is based on the concept of mutual information has recently been proposed by Palus [111].

3.3 Phase synchronization

The first step in quantifying phase synchronization between two time series x and y is the determination of their phases $\phi_x(t)$ and $\phi_y(t)$. Here this is achieved either via Hilbert transform or via wavelet transform[2]. From the rendered phase distributions three different indices are calculated. Combining these two steps this adds up to a total of six measures of phase synchronization used in this thesis and described in the following.

3.3.1 Extracting the phase

As a first way to determine the phases of the time series a method is used which is based on the *analytic signal* approach [38, 113]. From the continuous time series $x(t)$ first the analytic signal is defined

$$Z_x(t) = x(t) + i\,\tilde{x}(t) = A_x^H(t)e^{i\phi_x^H(t)}, \tag{3.11}$$

where $\tilde{x}(t)$ is the *Hilbert transform* of $x(t)$:

$$\tilde{x}(t) \equiv (\mathcal{H}x)(t) = \frac{1}{\pi} p.v. \int_{-\infty}^{+\infty} \frac{x(t')}{t-t'} dt' \tag{3.12}$$

(here p.v. denotes the Cauchy principal value). This renders the unambiguous definition of the so-called *instantaneous phase*:

$$\phi_x^H(t) = \arctan \frac{\tilde{x}(t)}{x(t)}. \tag{3.13}$$

Analogously, A_y^H and ϕ_y^H are defined from $y(t)$.

In the frequency domain the Hilbert transform performs a phase shift of the original signal by $\frac{\pi}{2}$ leaving the power spectrum unchanged, thus creating an artificial imaginary part for the real time series. This can be seen by application of the convolution theorem

$$\tilde{x}(t) = -i \cdot FT^{-1}[FT[x(t)]sign(\omega)], \tag{3.14}$$

[2]Besides that, many other methods of extracting a phase have been proposed (cf. e.g., Refs. [21, 142]).

where FT denotes the Fourier transform and FT^{-1} its inverse.

The second method used to extract the phases from the time series is based on the *wavelet transform* and has recently been introduced by Lachaux et al. [71, 70]. In this approach the phase is determined by the convolution of the respective signal with a complex Morlet wavelet (here slightly modified according to Ref. [128])

$$\Psi(t) = (e^{i\omega_0 t} - e^{-\omega_0^2\sigma^2/2}) \cdot e^{-t^2/2\sigma^2}, \tag{3.15}$$

where ω_0 is the center frequency of the wavelet and σ denotes its rate of decay. This is proportional to the number of cycles and related to the frequency span by the uncertainty principle.

The convolution of $x(t)$ with $\Psi(t)$ yields a complex time series of wavelet coefficients

$$W_x(t) = (\Psi \circ x)_{(t)} = \int \Psi(t')\, x(t-t')\, dt' = A_x^W(t) \cdot e^{i\phi_x^W(t)}, \tag{3.16}$$

from which the phases can be defined as

$$\phi_x^W(t) = \arctan \frac{ImW(t)}{ReW(t)}. \tag{3.17}$$

In the same way $W_y(t)$ and $\phi_y^W(t)$ are defined from $y(t)$.

Although based on very different approaches these two different definitions of phase are indeed closely related, as demonstrated practically in Ref. [77] and explained theoretically by Quian Quiroga and colleagues [128]. In short, the phase based on the wavelet transform $\phi_x^W(t)$ corresponds approximately to the phase based on the Hilbert transform $\phi_x^H(t)$ which would be rendered after band pass filtering the time series. This correspondence would even get exact if the wavelet approach would be performed by a convolution with an analytic wavelet and if this wavelet would be used for the band pass filtering in the Hilbert approach. Therefore in the wavelet approach the center frequency ω_0 and the frequency width σ of the wavelet can serve as parameters to adjust the frequency range of interest. In contrast to this, the actual phase extraction based on the Hilbert transform is free of parameters. This phase retains information from the entire power spectrum. Thus a comparison of narrow band and broad band synchronization is achievable just by using both approaches without applying any extra filtering.

3.3.2 Indices of phase synchronization

Based on the phase distributions $\phi_x(t)$ and $\phi_y(t)$ extracted in the first step, three different indices of phase synchronization are calculated.

Two of these indices are derived from information theory and have been introduced by Tass and coworkers in 1998 [163]. As a prerequisite for the calculation of both indices an equidistant binning of the interval $[0, 2\pi[$ is performed. An appropriate value for the number of bins L is dependent on the length of the two time series [141]:

$$L = e^{0.626+0.4\ln(N-1)}. \tag{3.18}$$

The first index is calculated from the binned distribution of phase differences. The probability for a phase difference to belong to a certain bin l is roughly estimated by the relative number of phase differences of the given time series in this bin:

$$p_l = \frac{\#(\ \{\ \phi_x(t_j) - \phi_y(t_j) \in [\frac{l}{L}2\pi, \frac{l+1}{L}2\pi[\ \}\ )}{N}, \quad l = 1 \ldots L, \tag{3.19}$$

where $\#(\{\cdot\})$ indicates the number of elements in the set $\{\cdot\}$. With

$$S = -\sum_{l=1}^{L} p_l \cdot \ln p_l \tag{3.20}$$

denoting the Shannon entropy [157] of this distribution and with

$$S_{max} = \ln L \tag{3.21}$$

its maximum possible value (attained for homogenous distributions), the normalized *index based on Shannon entropy* is yielded as[3]:

$$\gamma_{se} = \frac{S_{\max} - S}{S_{\max}}. \tag{3.22}$$

The second index quantifies the conditional probability for ϕ_y to fall into a certain bin given the bin containing ϕ_x. This *index based on conditional probability* is defined as

$$\gamma_{cp} = \frac{1}{L}\sum_{l=1}^{L} |r_l| \tag{3.23}$$

It is the average over

[3]This definition is used following Ref. [163], although it is not quite correct. Instead of the uniform distribution, S_{max} should be estimated from independent pairs of phases. Since in general the phase distribution is not uniform, the distribution of phase differences is not either. This issue has first been addressed in Ref. [92].

$$r_l = \frac{1}{M_l} \sum_{\substack{j \\ \phi_x(t_j) \in [\frac{l}{L}2\pi, \frac{l+1}{L}2\pi[}} e^{i\phi_y(t_j)} \tag{3.24}$$

with

$$M_l = \#\ \{\ \phi_x(t_j) \in [\frac{l}{L}2\pi, \frac{l+1}{L}2\pi[\ \} \tag{3.25}$$

denoting the number of phase values $\phi_x(t)$ belonging to bin l.

The third index of phase synchronization applied in this thesis is the *index based on circular variance*:

$$\gamma_{cv} = \left| \frac{1}{N} \sum_{j=1}^{N} e^{i[\phi_x(t_j) - \phi_y(t_j)]} \right| = 1 - CV. \tag{3.26}$$

Here CV denotes the circular variance [92] of an angular distribution obtained by transforming the phase differences onto the unit circle in the complex plane. In the literature this index has been introduced by Mormann and colleagues using the term *mean phase coherence* [104], later is has also been referred to as the intensity of the first Fourier mode of the phase distribution [141].

The last index is based on the circular variance, the second statistical moment of a phase distribution, and thus bears the advantage over the other two indices that it can easily be adapted to statistical moments of higher order[4]. Furthermore, it is the only index with a straightforward way of determining a significance threshold[5] to avoid spurious detections of synchronization in uncoupled time series (equivalent to the Bartlett estimator for C_{max}). On the other hand, γ_{cv} can underestimate the actual phase synchronization in case of a multi-modal distribution of ϕ_{xy}. This can occur when the phase difference remains fairly stable but occasionally jumps between different values [178]. Although the signals are synchronized then (except at the times of the jumps), the phase differences ϕ_{xy} may cancel in the time average, thus rendering a low γ_{cv}. A multi-modal distribution of the phases can also appear if the signals are investigated for a $1:1$ synchronization but the real relationship is $1:2$. In this context, it is noteworthy that all three indices could easily be adapted to the more general case of $n:m$ synchronization [163, 104]. Yet throughout this thesis the analysis remains restricted to the standard case of $1:1$ phase synchronization (since in our case there is no reason to expect any different $n:m$ synchronization).

Combining two different ways to extract the phases with three different indices yields a total of six measures of phase synchronization, namely γ_{se}^H, γ_{cp}^H and γ_{cv}^H as well as γ_{se}^W,

[4]Notice, however, that the same caveats apply to it as mentioned in the previous footnote.

[5]According to the Rayleigh test of uniformity [92] the significance threshold at a confidence level of $p = 0.05$ is given by $\gamma_{cv} = \sqrt{5.991/2N}$.

γ_{cp}^W and γ_{cv}^W. All these indices are confined to the interval [0,1]. Values close to zero are attained for phase differences forming a rather uniform distribution (no phase synchronization) while the maximum value corresponds to Dirac-like distributions (perfect phase synchronization). The most important feature of these indices is that they are only sensitive to phases, irrespective of the amplitude of the two signals. This feature has been illustrated first in Ref. [139] using bidirectionally coupled model systems.

All measures described so far are symmetric by definition and therefore are not suited to exploit the directionality of interaction. Just recently two asymmetric extensions of the concept of phase synchronization have been proposed, the first one by Rosenblum et al. [138, 137] and the second one, an information-theoretic approach, by Palus and Stefanovska [112]. The former approach has been adapted for the application to short and noisy time series in Ref. [159]. The property of asymmetry is shared by the two measures of nonlinear interdependence and the event synchronization. The underlying concepts of these measures are introduced in the following two Sections.

3.4 Nonlinear interdependencies

As measures for generalized synchronization between two time series x and y, the nonlinear interdependencies S and H have been introduced by Arnhold and coworkers in 1999 [14]. They are related to earlier attempts to detect generalized synchronization like the method of mutual false nearest neighbors [145] and the mutual cross predictability introduced in Ref. [148] and applied to biological data in Refs. [75, 78]. In contrast to these measures the nonlinear interdependencies S and H do not assume a strict functional relationship between the dynamics of the underlying systems X and Y.

A prerequisite for the calculation of the nonlinear interdependencies S and H is the state space reconstruction of the individual time series x and y. Using the method of embedding described in Section 2.2, time delay vectors $\overrightarrow{x}_n = (x_n, \ldots, x_{n-(m-1)\tau})$ and $\overrightarrow{y}_n = (y_n, \ldots, y_{n-(m-1)\tau})$ with embedding dimension m and time lag τ are reconstructed. Subsequently, the criterion is investigated whether neighborhood in the state space of Y implies neighborhood in the state space of X for equal time partners, and the same is done for the other direction just by exchanging X and Y.

Denoting the time indices of the k nearest neighbors of $\overrightarrow{x}_n$ and $\overrightarrow{y}_n$ with $r_{n,j}$ and $s_{n,j}$, $j = 1, \ldots, k$, respectively, for each $\overrightarrow{x}_n$, the squared mean Euclidean distance to its k neighbors is defined as

$$R_n^{(k)}(X) = \frac{1}{k} \sum_{j=1}^{k} \left(\overrightarrow{x}_n - \overrightarrow{x}_{r_{n,j}} \right)^2. \tag{3.27}$$

Then, by replacing the nearest neighbors by the equal time partners of the closest neighbors of $\overrightarrow{y}_n$, the *y-conditioned* squared mean Euclidean distance is given as:

$$R_n^{(k)}(X|Y) = \frac{1}{k}\sum_{j=1}^{k}\left(\overrightarrow{x}_n - \overrightarrow{x}_{s_{n,j}}\right)^2. \tag{3.28}$$

If the systems are strongly correlated, then $R_n^{(k)}(X|Y) \approx R_n^{(k)}(X)$ is very small when compared with the average squared distance $R(X) = \frac{1}{N}\sum_{n=1}^{N} R_n^{(N-1)}(X)$ of $\{\overrightarrow{x}_n\}$, while $R_n^{(k)}(X|Y) \approx R_n^{(N-1)}(X) \gg R_n^{(k)}(X)$ holds for independent systems. Accordingly, the nonlinear interdependence measure $S^{(k)}(X|Y)$ is defined as

$$S(X|Y) = \frac{1}{N}\sum_{n=1}^{N}\frac{R_n^{(k)}(X)}{R_n^{(k)}(X|Y)}. \tag{3.29}$$

Since $R_n^{(k)}(X|Y) \geq R_n^{(k)}(X)$ by construction, a proper normalization

$$0 < S(X|Y) \leq 1 \tag{3.30}$$

is achieved with low values suggesting independence between X and Y and high values indicating interdependence.

The second nonlinear interdependence measure $H(X|Y)$ is defined as

$$H(X|Y) = \frac{1}{N}\sum_{n=1}^{N}\log\frac{R_n(X)}{R_n^{(k)}(X|Y)}. \tag{3.31}$$

In the definition of $S(X|Y)$ the conditional distance $R_n^{(k)}(X|Y)$ is compared to the local distance $R_n^{(k)}(X)$ in the time series x itself, in $H(X|Y)$ it is related to the mean squared distance between two state vectors:

$$R_n(X) = \frac{1}{N-1}\sum_{\substack{j=1\\ j\neq n}}^{N}(\overrightarrow{x}_n - \overrightarrow{x}_j)^2. \tag{3.32}$$

Also H is close to zero if X and Y are completely independent, while it is positive if neighborhood in Y implies neighborhood in X for equal time partners. But in contrast to S it is not bounded and it even can attain slightly negative values for the unlikely but not impossible case that close pairs in Y correspond mainly to distant pairs in X.

Exchanging systems X and Y yields the opposite interdependencies $S(Y|X)$ and $H(Y|X)$ defined in complete analogy. They are generally not equal to $S(X|Y)$ and $H(Y|X)$ and this asymmetry of S and H is the main advantage over other nonlinear measures such as mutual information or indices of phase synchronization. It can correctly reflect driver-responder relationships [14, 127, 151], but can also just be due to different dynamical properties of the individual time series [14, 127, 118, 160].

This aspect is isolated by quantifying the over-all degree of synchronization with one measure, i.e.,

$$S_s = \frac{S(X|Y) + S(Y|X)}{2} \quad \text{and} \quad H_s = \frac{H(X|Y) + H(Y|X)}{2} \tag{3.33}$$

and the asymmetry with another measure, i.e.,

$$S_a = \frac{S(X|Y) - S(Y|X)}{2} \quad \text{and} \quad H_a = \frac{H(X|Y) - H(Y|X)}{2}. \tag{3.34}$$

For the latter positive (negative) values can indicate driver-responder relationships with X being the driver (responder) and Y being the responder (driver).

3.5 Event Synchronization

The last measure applied in this thesis is the so-called event synchronization proposed by Quian Quiroga and colleagues in 2002 [130]. This measure, which is based on the relative timings of certain events in the time series, is different from the others in many respects. First of all there exist two variants, the first one quantifying the over-all level of synchronicity from the number of quasi-simultaneous appearances of these events, the second one dealing with time delay patterns extracted from the precedence of events in one signal with respect to the other. Secondly it allows the tracking of changes in synchronization and delay with a much higher time resolution than all the other measures. And finally it is very adaptive since for each application a different choice of events is possible thus enabling to focus the attention on the pattern of interest without any disturbance from structures irrelevant to the application in mind.

Before the actual calculation of event synchronization, suitable events have to be defined. These can be characteristic patterns in the time series such as spikes but also rather common features like local maxima and/or minima. Then in the first step the respective time series x and y are scanned for these events and the times of their occurrence are marked as t_i^x and t_j^y $(i = 1, \ldots, m_x; j = 1, \ldots, m_y)$ with m_x and m_y denoting the respective number of events. Thus the original time series are actually replaced by new series of event times (a procedure similar to the well known method of symbolic dynamics) subject to the quantification of their synchronicity in the last step. Actually this reduction of information is the most crucial part since event synchronization does not really measure the synchronization between the two time series as a whole but rather the synchronicity between the defined events only. Thus the event synchronization of two time series can attain totally different values for different choices of events. Positively speaking the method of event synchronization is very adaptive, negatively speaking it might not be robust at all.

A rather useful type of event is given by the following definition of a local maximum with K denoting the width and h the height of the event (a local minimum is defined accordingly):

$$x(t_i) > x(t_{i\pm k}) \;\; for \;\; k = 1, \ldots, K-1 \;\; and \;\; x(t_i) > x(t_{i\pm K}) + h \tag{3.35}$$

In order to achieve an appropriate representation of the original time series in the sequence of events and to obtain a suitable statistics, it is crucial in the definition of events to ensure a sufficient number of events. Therefore, in this thesis event times are defined as the simplest local maxima and local minima with parameters set to $K = 1$ and $h = 0$. Thus the only conditions to be fulfilled are $x(t_i) > x(t_{i\pm1})$ and $x(t_i) < x(t_{i\pm1})$ and the same in y. In the last step the synchronicity of events in X and Y is quantified. This can be done in two different ways.

The first variant is designed for time series with a rather constant event rate. Depending on this event rate a maximum time lag τ is chosen until which two events are still considered to be synchronous. To avoid double counting, τ should at least be smaller than half the minimum inter-event distance. Using

$$J_{ij}^{\tau} = \begin{cases} 1 & \text{if} \;\; 0 < t_i^x - t_j^y \leq \tau \\ 1/2 & \text{if} \;\; t_i^x = t_j^y \\ 0 & \text{else} \end{cases} \tag{3.36}$$

to quantify the relative timing of all events, the number of times an event appears in x shortly after it appears in y is counted by

$$c(x|y) = \sum_{i=1}^{m_x} \sum_{j=1}^{m_y} J_{ij}^{\tau}. \tag{3.37}$$

With the opposite value $c(y|x)$ defined accordingly, the symmetrical and anti-symmetrical combinations are given by

$$Q = \frac{c(y|x) + c(x|y)}{\sqrt{m_x}\sqrt{m_y}} \tag{3.38}$$

and

$$q = \frac{c(y|x) - c(x|y)}{\sqrt{m_x}\sqrt{m_y}}. \tag{3.39}$$

These measures are designed to quantify the synchronization of the time events and their delay behavior, respectively. They are properly normalized to $0 \leq Q \leq 1$ and $-1 \leq q \leq 1$

with $Q = 1$ if and only if all events of the signals are synchronous and $q = 1$ if and only if the events in x always precede those in y and $q = -1$ if vice versa.

In cases where the event rate changes during the recording and a global time scale τ does not seem appropriate, the second variant uses a local definition of τ_{ij} for each event pair (ij):

$$\tau_{ij} = \min\{t^x_{i+1} - t^x_i, t^x_i - t^x_{i-1}, t^y_{j+1} - t^y_j, t^y_j - t^y_{j-1}\}/2. \tag{3.40}$$

The factor $1/2$ is introduced to avoid double counting in case that, e.g., two events in x are close to the same event in y. As in the definition of events, the optimal choice of τ_{ij} depends on the problem, other choices like $\tau'_{ij} = \min\{\tau, \tau_{ij}\}$ are also possible. With τ replaced by τ_{ij} the quantity J_{ij} is then defined as in Eq. (3.36).

Like all the other measures of synchronization introduced so far, Q and q defined in Eqs. (3.38) and (3.39) can be calculated for two simultaneously measured time series of length N resulting in two values quantifying the level of event synchronization Q and delay asymmetry q in this time frame, respectively. But in contrast to the other approaches this method also allows to follow and visualize changes in synchronization and delay with a much higher temporal resolution. Time resolved variants of Q and q are simply obtained by modifying Eq. (3.37) to

$$c_n(x|y) = \sum_i \sum_j J_{ij}\, \Theta(n - t^x_i) \tag{3.41}$$

with $n = 1, \ldots, N$ and Θ denoting the Heaviside step function (i.e., $\Theta(a) = 0 \; for \; a \leq 0 \; and \; \Theta(a) = 1 \; for \; a > 0$). The same way $c_n(y|x)$ is obtained by exchanging x and y.

From this, the symmetrical and anti-symmetrical combinations are given by

$$Q(n) = c_n(y|x) + c_n(x|y) \qquad \text{and} \qquad q(n) = c_n(y|x) - c_n(x|y) \tag{3.42}$$

respectively, allowing the time resolved visualization of event synchronization and delay asymmetry as random walks. First, if an event is found both in x and y within the window τ (respectively τ_{ij}), the event synchronization $Q(n)$ increases one step, otherwise it does not change. Of course, $Q(n)$ also does not change as long as there are no new events at all. The increase of the monotonic function $Q(n)$ is proportional to the number of synchronous events in the two time series.

On the other hand, the delay asymmetry $q(n)$ performs a random walk taking one step up every time an event in x precedes one in y and one step down in the opposite case. Whenever an event occurs simultaneously in both signals or appears only in one of them, the random walker does not move. In case of non-synchronized signals a random walk with the typical diffusion behavior is obtained, whereas for signals with a delayed synchronization $q(n)$ shows a bias going up if x precedes y and going down if vice versa. Such a

bias clearly shows the presence of a time delay of the one signal with respect to the other, but does not necessarily prove the existence of a driver-responder relationship, although it might suggest it. In fact, as already mentioned in Section 2.3, the two signals might be driven by a common hidden source and the bias could just indicate different delays. Also, internal delay loops in one of the two systems could fool the interpretation. In this respect, it is not different from other indicators of directionality like those of Refs. [138, 137]

Normalized local variants of event synchronization and delay asymmetry can be obtained by taking the local derivative of the accumulative quantities $Q(n)$ and $q(n)$. The degree of synchronization at time n, averaged over the last Δn time steps, is given by

$$Q'(n) = \frac{Q(n) - Q(n - \Delta n)}{\sqrt{\Delta m_x \Delta m_y}} \tag{3.43}$$

with Δm_y and Δm_x denoting the respective number of events in the interval $[n - \Delta n, n]$. Similarly, the local delay asymmetry is defined as

$$q'(n) = \frac{q(n) - q(n - \Delta n)}{\sqrt{\Delta m_x \Delta m_y}}. \tag{3.44}$$

Chapter 4

Application to coupled model systems

The study of synchronization phenomena in chaotic systems has been a topic of increasing interest since the early 1990s (for an overview cf. Ref. [123]). A great deal of attention is paid to identifying certain regimes of synchronization in coupled identical or non-identical systems with varying parameters (e.g., different frequency mismatches or an increasing coupling strength). Various methods exist for the detection of the different types of synchronization (cf. Section 2.3). Complete synchronization can easily be recognized by plotting a component of the driver versus the respective component of the responder. Phase synchronization can be established by a vanishing mean frequency difference [139]. As a practical criterion for the existence of generalized synchronization in unidirectionally driven systems, the negativeness of the largest Lyapunov exponent of the responder is usually employed [116].

Investigations are typically carried out either analytically or by numerically analyzing long and noise free time series (typically at least in the order of 10^5 data points) that are generated using model equations. Only rarely the dependence between coupled model systems is evaluated by applying bivariate measures to short time series (in the order of 10^3 data points) with or without additive noise. If at all, then almost exclusively a single approach to measure synchronization is used. Examples of investigations on phase synchronization include Refs. [163, 104, 112, 159], while approaches to detect generalized synchronization have been studied in Refs. [145, 148, 127, 150, 160]. In Ref. [109] the linear cross correlation, the cross correlation sum [54] as well as different measures for generalized synchronization (including the nonlinear interdependencies S and H) have been employed on several coupled model systems. Different approaches like the cross correlation, measures of phase synchronization as well as the nonlinear interdependencies as measures of generalized synchronization have been compared qualitatively in Ref. [99] by analyzing two bidirectionally coupled Rössler systems.

Nevertheless, a comprehensive comparison of different approaches analyzing different coupled model systems in a 'controlled' setting is still missing and thus declared as the aim

of the present study. For this aim, the measures of synchronization[1] described in Chapter 3 are applied to three coupled model systems with different individual properties (e.g., power spectra, dimension). These comprise coupled Hénon maps as well as coupled Rössler and coupled Lorenz systems. This study is divided into two parts corresponding to two different control parameters. With the coupling strength as the first parameter it is tested to which extent the different measures are able to distinguish between different degrees of coupling. This property is essential in many field applications since rarely the absolute value of synchronization is of interest, but rather it is the change of synchronization between two different states, times, or recording sites that matters. The second parameter, the signal-to-noise ratio, is used to investigate whether the results of the different measures prove to be robust when the signals of interest are contaminated with a certain amount of noise. Robustness against noise is an important prerequisite for the application of these measures to field data, since noise is an inevitable disturbance in any measurement setting.

The remainder of this study is organized as follows: The coupling schemes are introduced in Section 4.1.1, whereas the underlying model systems are described in Appendix A. In Section 4.1.2 the parameters used in the practical implementation of the different measures of synchronization are given followed by the statistical evaluation designed to compare these measures in Section 4.1.3. Results on the measures' capability to reflect the strength of coupling and their robustness against noise are presented in Sections 4.2.1 and 4.2.2, respectively. Finally, conclusions are drawn in Section 4.3.

4.1 Methods

4.1.1 Coupling schemes

To investigate and compare the performance of the different bivariate measures, for each measure the synchronization between the first components of the following coupled model systems is calculated.

4.1.1.1 Hénon - Hénon

The first example consists of two unidirectionally coupled Hénon maps (cf. Section A.1). This coupling scheme was proposed in Ref. [148] and later analyzed in Refs. [127, 150, 111, 160, 109]. The equations of motion read for the driver

$$\begin{aligned} x_1' &= 1.4 - x_1^2 + b_x\, x_2 \\ x_2' &= x_1, \end{aligned} \tag{4.1}$$

[1] The directionality of the different approaches has already been addressed in Refs. [127, 150, 128], a comparison of different approaches is subject to current research [158].

and for the responder

$$\begin{aligned} y_1' &= 1.4 - (C\, x_1 y_1 + (1-C)\, y_1^2) + b_y\, y_2 \\ y_2' &= y_1. \end{aligned} \tag{4.2}$$

The parameters are set to $b_x = b_y = 0.3$ to yield identical systems. In Fig. 4.1 the attractor of the responder as well as plots of the first component of the driver versus the first component of the responder are shown for increasing values of the coupling strength C. The attractor of the responder looks the same for $C = 0$ and $C = 0.8$ (left), but while for $C = 0$ driver and responder are completely independent (the apparent structure in the right plot is due to the non-uniform densities of the individual systems), identical synchronization between driver and responder can be observed for $C = 0.8$ (right). In between a rather sharp transition slightly below $C = 0.7$ takes place.

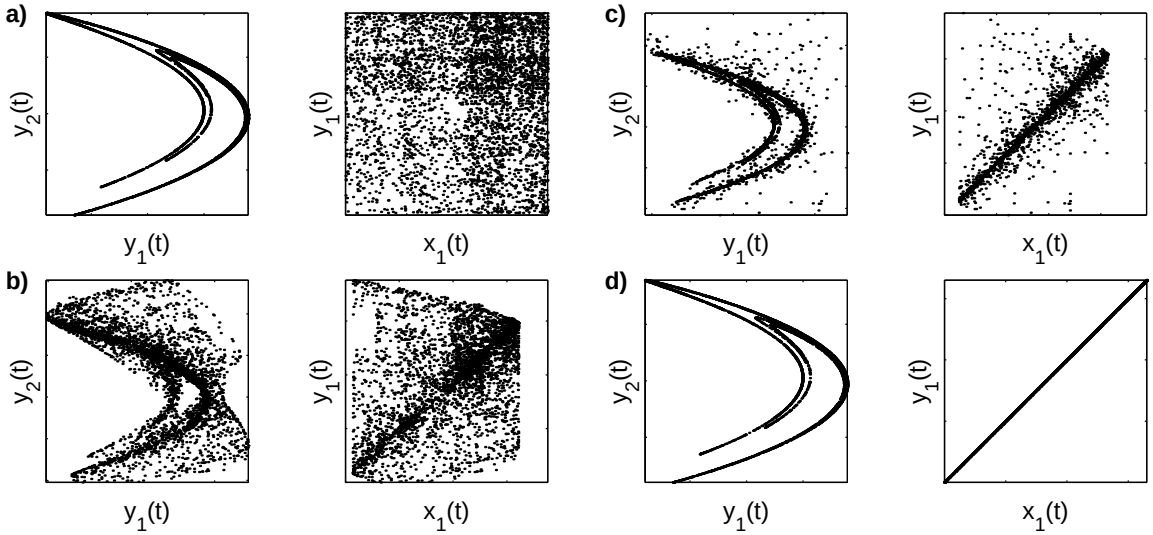


Figure 4.1: Coupled identical Hénon maps. The coupling strength C is varied non-equidistantly: a) 0.0, b) 0.6, c) 0.7, d) 0.8. For each coupling strength the attractor of the responder is depicted on the left, whereas on the right the first component of the driver is plotted versus the first component of the responder.

4.1.1.2 Rössler - Rössler

For the two Rössler systems (cf. Section A.3) a unidirectional coupling is employed using an additional diffusive coupling term. This coupling scheme has been studied in Refs. [104, 179, 112, 109]. The equations of motion read

$$\begin{aligned} \frac{dx_1}{dt} &= -\omega_x\, x_2 - x_3 \\ \frac{dx_2}{dt} &= \omega_x\, x_1 + 0.15\, x_2 \\ \frac{dx_3}{dt} &= 0.2 + x_3\,(x_1 - 10) \end{aligned} \tag{4.3}$$

for the first Rössler system and

$$\begin{aligned} \frac{dy_1}{dt} &= -\omega_y\, y_2 - y_3 + \;\; C\,(x_1 - y_1) \\ \frac{dy_2}{dt} &= \omega_y\, y_1 + 0.15\, y_2 \\ \frac{dy_3}{dt} &= 0.2 + y_3\,(y_1 - 10). \end{aligned} \tag{4.4}$$

for the second. A small parameter mismatch between the two systems is introduced by setting $\omega_x = 0.95$ and $\omega_y = 1.05$. The coupling strength C is varied from 0 to 2 in steps of 0.025. In Fig. 4.2 the attractor of the responder as well as the first component of the driver X versus the first component of the responder Y are plotted for increasing values of the coupling strength C. A clear tendency towards the identity can be observed (although complete synchronization will never be reached due to the parameter mismatch).

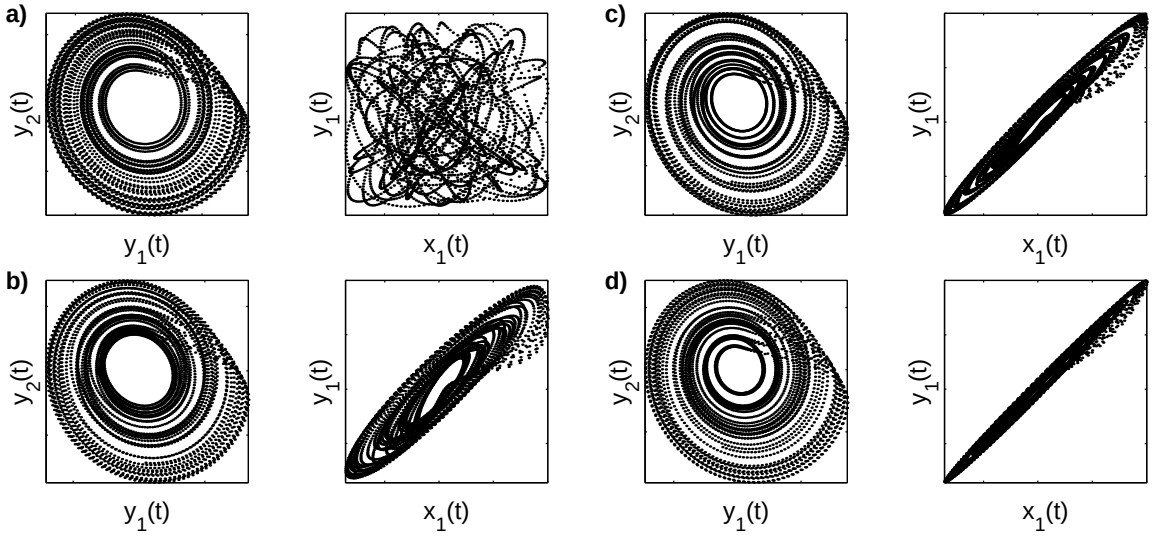


Figure 4.2: Same as Fig. 4.1, but for the coupled Rössler systems. The coupling strength C is increased non-equidistantly according to a) 0.0, b) 0.5, c) 1.0, d) 2.0.

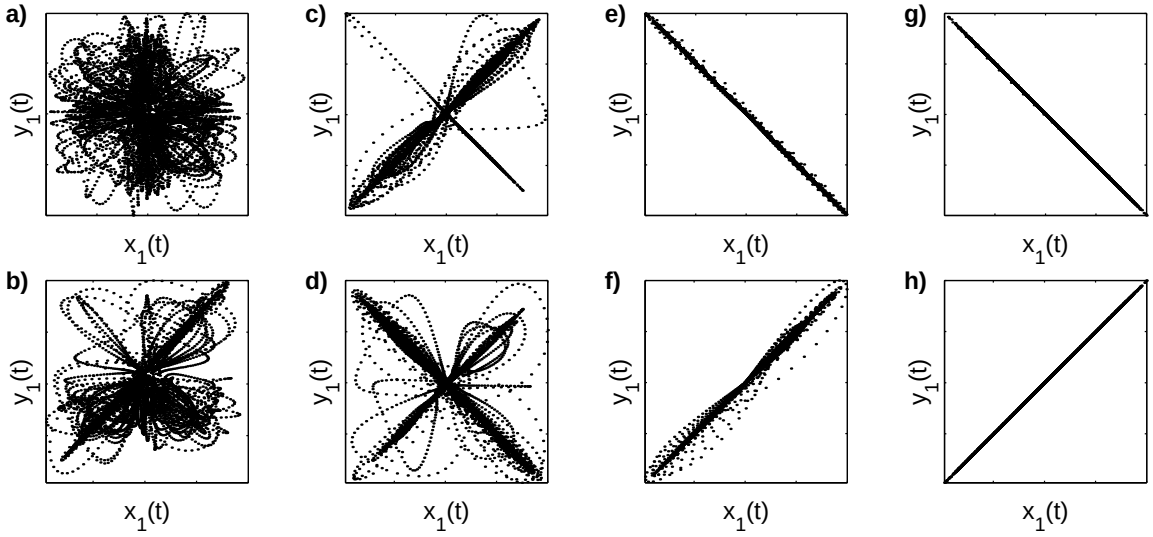
4.1.1.3 Lorenz - Lorenz

For the two Lorenz systems (cf. Section A.2) the same diffusive coupling scheme already employed in Section 4.1.1.2 is used. The equations of motion of this coupling scheme, which has already been investigated in Refs. [180, 109], read for the first Lorenz system

$$\begin{aligned} \frac{dx_1}{dt} &= 10(x_2 - x_1) \\ \frac{dx_2}{dt} &= x_1\,(28 - x_3) - x_2 \\ \frac{dx_3}{dt} &= x_1\,x_2 - \frac{8}{3}x_3 \end{aligned} \tag{4.5}$$

and for the second

$$\begin{aligned} \frac{dy_1}{dt} &= 10(y_2 - y_1) \\ \frac{dy_2}{dt} &= y_1\,(28.001 - y_3) - y_2 \\ \frac{dy_3}{dt} &= y_1\,y_2 - \frac{8}{3}y_3 + \;C\,(x_1 - y_1). \end{aligned} \tag{4.6}$$



[ht]

Figure 4.3: Coupled Lorenz systems. First component of the first system versus first component of the second system. The coupling strength C is increased according to a) 0.0, b) 1.2, c) 1.3, d) 1.325, e) 1.375, f) 1.4, g) 1.475, h) 1.5.

Please note the small parameter mismatch introduced in the second component. The coupling strength C is varied from 0 to 2 in steps of 0.025. In Fig. 4.3 plots of the first component of the first system X versus the first component of the second system Y are shown for increasing values of the coupling strength C. Here the transition towards a synchronized state (but again no complete synchronization) is much less smooth and there are even some distinct regimes of anti-correlation in between.

4.1.2 Implementation of measures

For each of the coupled model systems and every parameter combination (coupling, noise level) time series of length 8192 are created, and the first and the last quarter are discarded leaving $N = 4096$ data points in the middle for the analysis. The analyzed segments are taken as the middle part of a larger time series [59] in order to avoid edge effects in the calculation of the measures for phase synchronization. This needs to be explained in more detail: The extraction of the instantaneous phases in principle requires integration over infinite time. Since in practice only the values of neighboring data points give significant contributions, the only problem left is the extraction of phases close to the edges of the time series. This problem is solved by the described procedure, since then there are no edges inside the analyzed time series left. Thus the whole 8192 data points are used for the determination of the phases, but only the 4096 phases of the middle part contribute to the phase distributions needed for the calculation of the measures of phase synchronization. The value $N = 4096$ is selected according to the number of data points used in the analysis of EEG data in Chapter 5. It is chosen as a power of 2 to enable the use of Fast Fourier Transform (FFT) algorithms [124] for the calculation of the Fourier-based measures, i.e., cross correlation and all indices for phase synchronization (cf. Sections 3.1 and 3.3). This reduces the computation time from order N^2 to order $N \log N$. The second step to enable the extraction of adequate phases is demeaning (i.e., setting the DC Fourier coefficient ($\omega = 0$) to zero) which is performed in the preprocessing in order to yield a sufficient number of zero crossings and thus a reasonable progression of the phase (cf. Ref. [60]). Furthermore, all time series are normalized to unit variance.

In addition to these steps of preprocessing certain parameters have to be selected for a meaningful application of the different measures. Only cross correlation and the three indices of phase synchronization based on the Hilbert transform are free of parameters. From the remaining measures only mutual information (cf. Section 3.2) and event synchronization (cf. Section 3.5) are calculated using the same parameters for all four coupled model systems. Mutual information is estimated without embedding (embedding dimension $m = 1$, no time delay τ needed) and with the number of nearest neighbors set to $k = 1$ in order to resolve the smallest possible scales (at the expense of less robustness). For the calculation of event synchronization events are defined as local maxima and minima and the method of fixed time lag was chosen. The remaining parameters and the respective values chosen for the different systems are listed in Tab. 4.1. These comprise

Parameter	Hénon-Hénon	Rössler-Rössler	Lorenz-Lorenz
SR	-	20	100
ω_c	0.42	0.35	0.17
nc	5	2	1
m	3	5	5
τ	1	5	5
T	50	50	50
k	10	10	10

Table 4.1: Parameters for the application of the measures to the different coupled model systems: Sampling rate SR, center frequency ω_c (in units of the Nyquist frequency), number of cycles nc, embedding dimension m, time delay τ, Theiler correction T and number of neighbors k.

the center frequency ω_c and the number of cycles nc of the mother wavelet for the three indices of phase synchronization based on the wavelet transform. For the calculation of the nonlinear interdependencies (cf. Section 3.4) a reasonable state space reconstruction (cf. Section 2.2) is a prerequisite. Thus embedding dimension m and time delay τ have to be chosen, furthermore a Theiler-correction T (necessary to exclude temporally related neighbors [165]) and the number of nearest neighbors k. All parameters are adapted to the respective system, e.g., the center frequency ω_c of the mother wavelet is chosen as the maximum frequency component of the system (cf. Appendix A). If possible, parameter values are chosen according to earlier studies (e.g., [127, 150, 128, 109, 99]).

4.1.3 Criterion for comparing different measures

In the first part of this study the different measures are compared in their capability to distinguish between different degrees of coupling between two systems. To evaluate this dependence on the coupling strength a measure of order is introduced as follows:

$$M = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign(s_j - s_i). \tag{4.7}$$

Given a sequence of values $s_i, i = 1, \ldots, n$, this measure evaluates for every possible pair of values in this sequence, whether the correct order is attained, i.e., whether the second value (the value with the higher index) is larger than the first value (the value with the lower index). It is properly normalized as a measure of ascending order, i.e., it attains the value 1 for a sequence with a strict monotonic increase, the value -1 for a sequence with a strict monotonic decrease, and the value 0 for a constant sequence. Its way of measuring order is illustrated in Fig. 4.4 using a linear and a square function, each with an increasing level of noise. Differences in steepness or the polynomial order of the increase do not matter, since

they could be easily eliminated by a suitable monotonic transformation. A linear monotony is as good as a monotony of any higher or lower order. As a means to compare different measures M is designed to yield its maximum value for a measure with a strict monotonic increase. For such a measure higher values of coupling strength necessarily lead to higher values of synchronization.

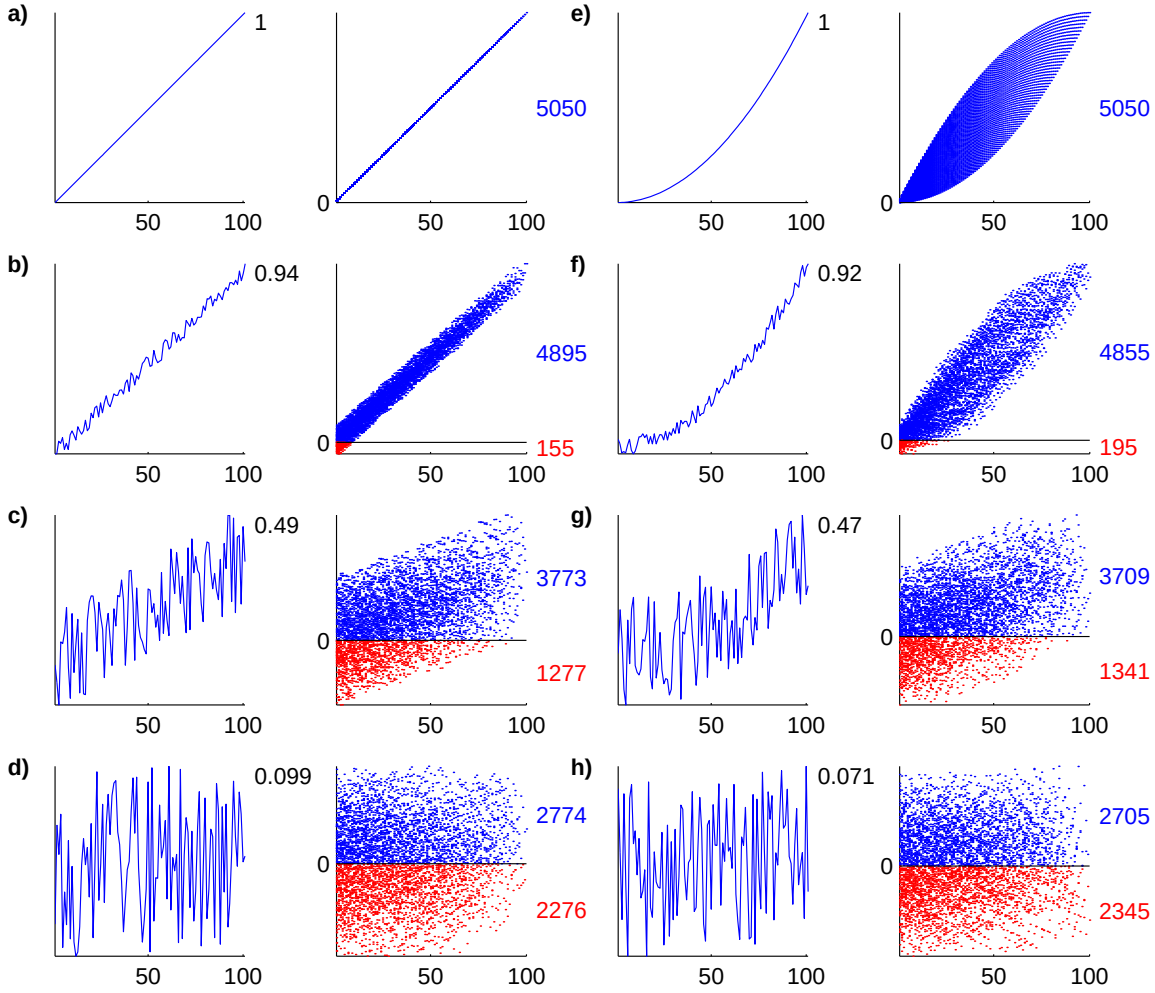


Figure 4.4: Simple examples to illustrate the measure of order M. a-d) Linear function with increasing amount of noise, e-h) Square function with increasing amount of noise. Left: Examples with 101 data point each (yielding 5050 pairs). Beside the plots the respective value of M is given. Right: y-distances of all single pairs versus the respective x-distances. Blue (Red): Pairs of value in correct (false) order. Zero line marks separation. Beside the plot the counted numbers of pairs with correct and false order are given. M is obtained as their difference divided by the total number of pairs.

In the second part of this study it is investigated to which extent the different measures of synchronization yield robust results when the signals of interest are contaminated with noise. For this aim the noise-to-signal ratio, defined as $NSR = \sigma_{noise}/\sigma_{signal}$, is used as the second parameter. For each coupled model system and every coupling strength this noise-to-signal ratio is increased according to

$$NSR = 10^{-2+n*0.1} \quad with\, n = 0, \ldots, 30, \tag{4.8}$$

thereby covering the range from -0.01 to 10 equidistantly on a logarithmic scale. For each NSR value, 10 different realizations of Gaussian white noise as well as iso-spectral noise (cf. Appendix B.1 and B.2) are generated and additively superimposed on the coupled model system. Subsequently, for every measure of synchronization the mean value over the different realizations is calculated.

To evaluate the robustness against noise, again the measure of order M is used. Its dependence on the noise-to-signal ratio serves as a means to track to which extent the initial order (i.e., the dependence on the coupling strength in the absence of noise) is destroyed by increasing levels of noise. To render a criterion for the comparison of the different measures, first for each measure M is normalized to the value of the respective systems without noise. By this means the robustness against noise can be regarded independently from the initial order. Subsequently, for each measure and every system the critical noise-to-signal ratio NSR_C is defined as the noise-to-signal ratio for which the normalized order M_n for the first time falls below the threshold $M_n^* = 1/\sqrt{2}$. To mark the special case that M_n does not cross M_n^*, NSR_C is set to a value beyond the maximum noise level analyzed. The higher this critical noise-to-signal ratio for a certain measure of synchronization, the more robust against noise is this measure.

4.2 Results

4.2.1 Dependence on coupling strength

In Fig. 4.5 the dependence of six representative measures of synchronization on the coupling between two identical Hénon-systems is shown. Along with these profiles the maximum Lyapunov exponent of the responder system is depicted (cf. Ref. [127]). It becomes negative for couplings larger than 0.7, when identical synchronization between the systems takes place (cf. Fig. 4.1). In the regime $0.47 < C < 0.52$ it is also slightly negative indicating weak generalized synchronization. This behavior is reflected by the bivariate measures showing higher degrees of synchronization with increasing coupling strength. While this is quite consistent among measures, considerable differences are found regarding their monotony and their sensitivity for weak couplings. The dependence on the coupling strength is rather monotonic for most measures, only for the index based on Shannon

entropy using the wavelet Transform large fluctuations can be observed (reflecting that γ_{se}^{W} is not adapted to the broad frequency spectrum of the Hénon-systems). For this measure quite high values are obtained already for uncoupled or only weakly coupled systems. This holds true even more for event synchronization Q whose increase, however, is much more steady and starts already at very low coupling strengths. The remaining measures also rise very slowly with the nonlinear interdependence S_s and mutual information I showing the most monotonic behavior. Cross correlation C_{max} and the index based on circular variance using the Hilbert Transform γ_{cv}^{H} are slightly less monotonic.

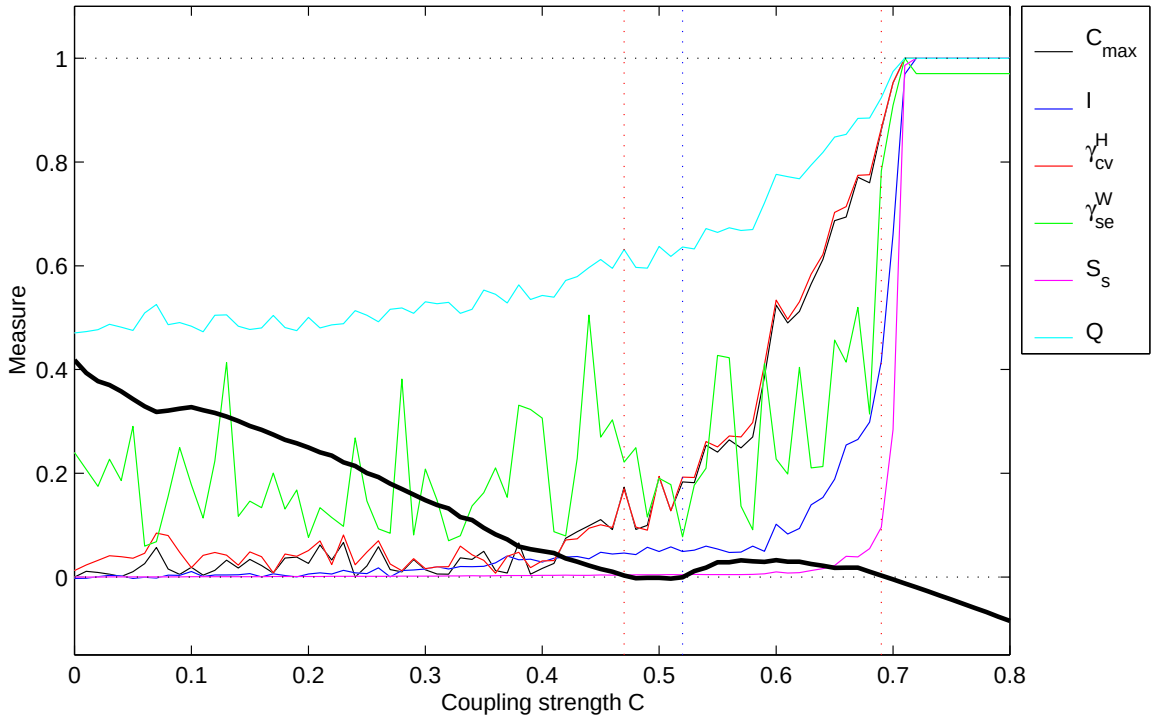


Figure 4.5: Dependence on the coupling strength for six measures of synchronization applied to coupled Hénon systems without noise. The maximum Lyapunov exponent of the responder system is depicted by the thick black line. Vertical lines mark its zero crossings.

The same dependencies are depicted in Fig. 4.6 for the coupled Rössler systems. For these quasi-periodic systems with a slight frequency mismatch the maximum Lyapunov exponent of the responder system crosses zero already at a very weak coupling and further decreases rather monotonic. The measures of synchronization consistently show an increase of synchronization for higher coupling strengths. Here γ_{se}^{W} attains an even higher value than Q for uncoupled systems (due to the quasi-periodicity of the two Rössler systems), but its increase towards the maximum value is quite slow as compared to event synchronization. The increases of Q and γ_{cv}^{H} are the steepest. These measures even show a distinct ceiling effect already for low values of the coupling strength. The increase of the remaining measures is rather gradual with C_{max} showing less fluctuations than S_s and I.

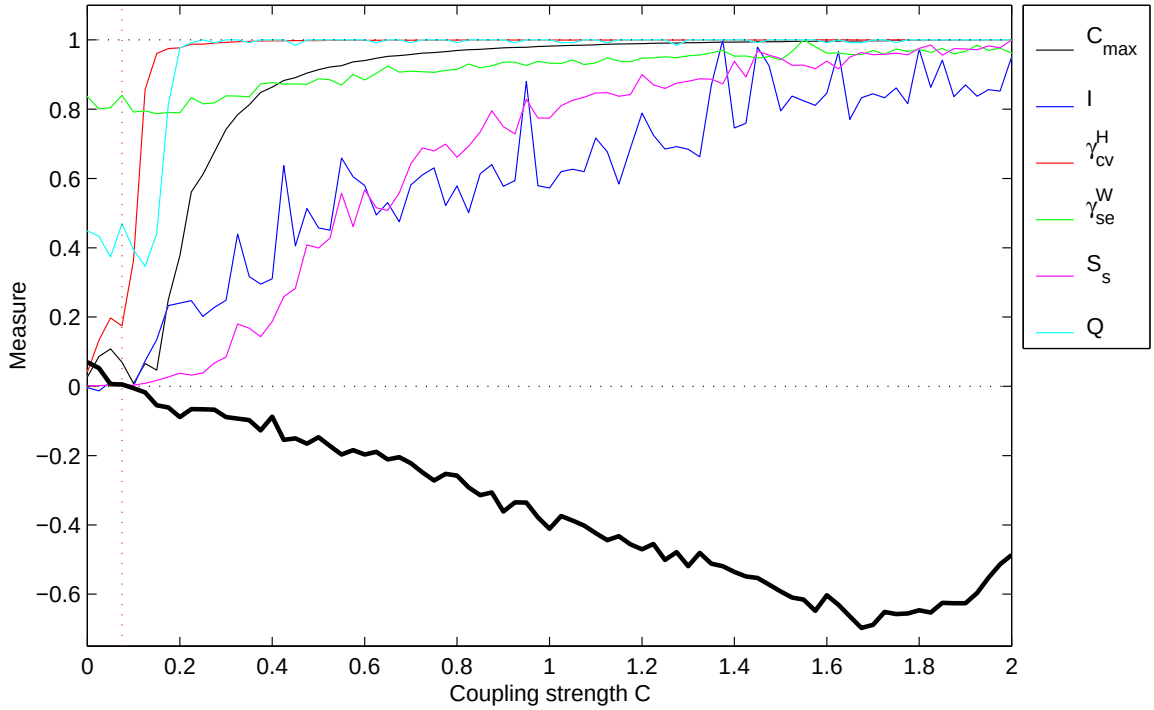


Figure 4.6: Same as Fig. 4.5, but this time for the coupled Rössler systems.

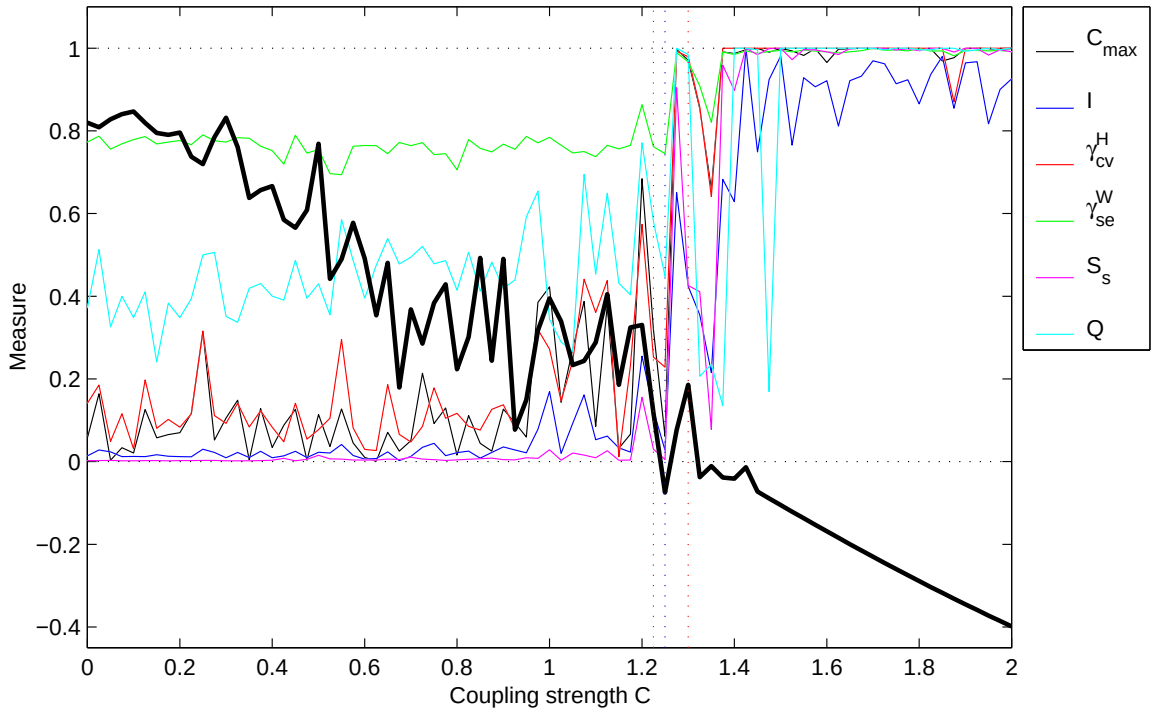


Figure 4.7: Same as Fig. 4.5, but this time for the coupled Lorenz systems.

Results for the coupled Lorenz systems are shown in Fig. 4.7. Here the maximum Lyapunov exponent of the responder system shows much more fluctuations than for the other systems. The change of sign takes place at an intermediate coupling strength $C = 1.3$. For the measures of synchronization considerable fluctuations can be observed as well, but for coupling strengths below the transition to generalized synchronization these stay on a rather constant level. This transition is consistently reflected by an increase of synchronization. Beyond the transition most measures reach their upper limit. Some measures exhibit rare deviations from this limit, and only mutual information fluctuates steadily below.

In the next step these qualitative results are quantified by means of the measure of order M. The results for each measure and every model system as well as the average values over measures and over systems are given in Fig. 4.8. The order in the dependence on the coupling strength differs considerably, both among measures as well as among systems. When comparing the different measures highest values of order are obtained for the nonlinear interdependencies followed by cross correlation and mutual information. The index based on conditional probability using the wavelet Transform γ^W_{cp} exhibits the least order by far. Regarding the different model systems it turns out that highest values are obtained either for the coupled Hénon systems or the coupled Rössler systems but never for the coupled Lorenz systems. This is due to the fluctuations seen already in Fig. 4.7.

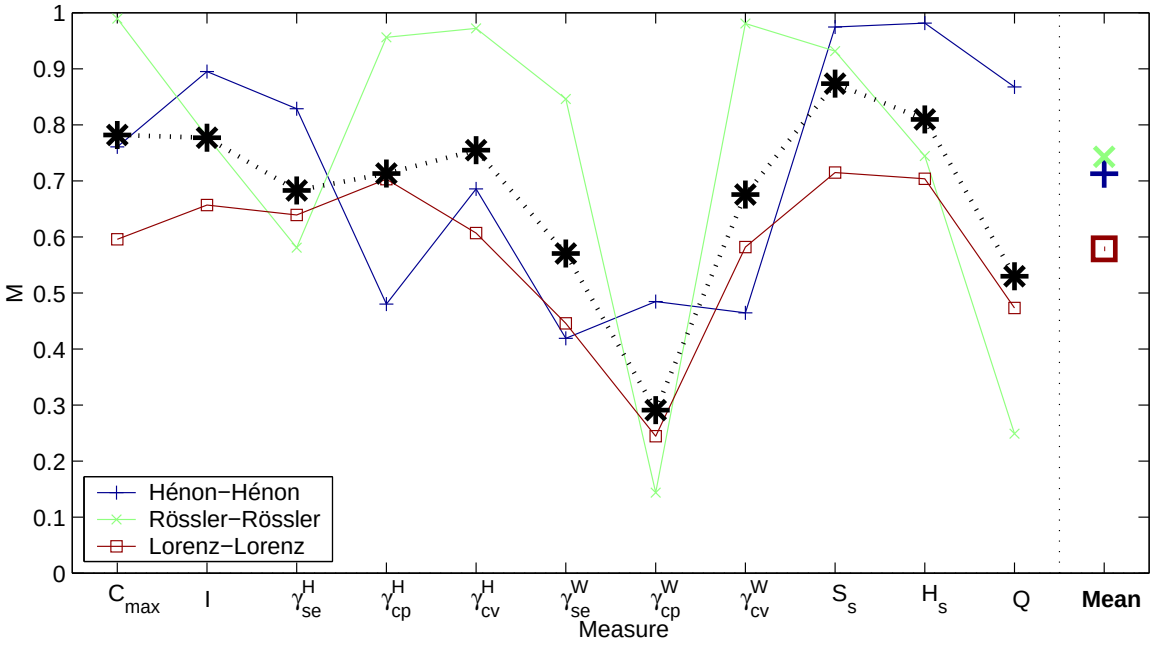


Figure 4.8: Over-all comparison for the case of noise-free model systems: Measure of order M for each measure and every system. The dotted black line depicts the mean over the three model systems. In the last column for each model system the mean value over measures is displayed.

4.2.2 Robustness against noise

In this Section the dependence of the different measures of synchronization on the coupling strength is evaluated for an increasing noise-to-signal ratio.

The dependence on the noise-to-signal ratio is displayed in Fig. 4.9 for the coupled Hénon systems and the same measures as in the previous Section. In order to assess to which extent the dependence on the coupling strength in the absence of noise is destroyed by increasing amounts of noise, the noise-free case $NSR = 0$ is depicted in the planar cross-section beyond the smallest noise-to-signal ratio $NSR = 0.01$. Thus this layer contains the values depicted in Fig. 4.5.

For all measures the increase of the noise-to-signal ratio leads to a gradual masking of the dependence on the coupling strength. For the maximum noise-to-signal ratio $NSR = 10$

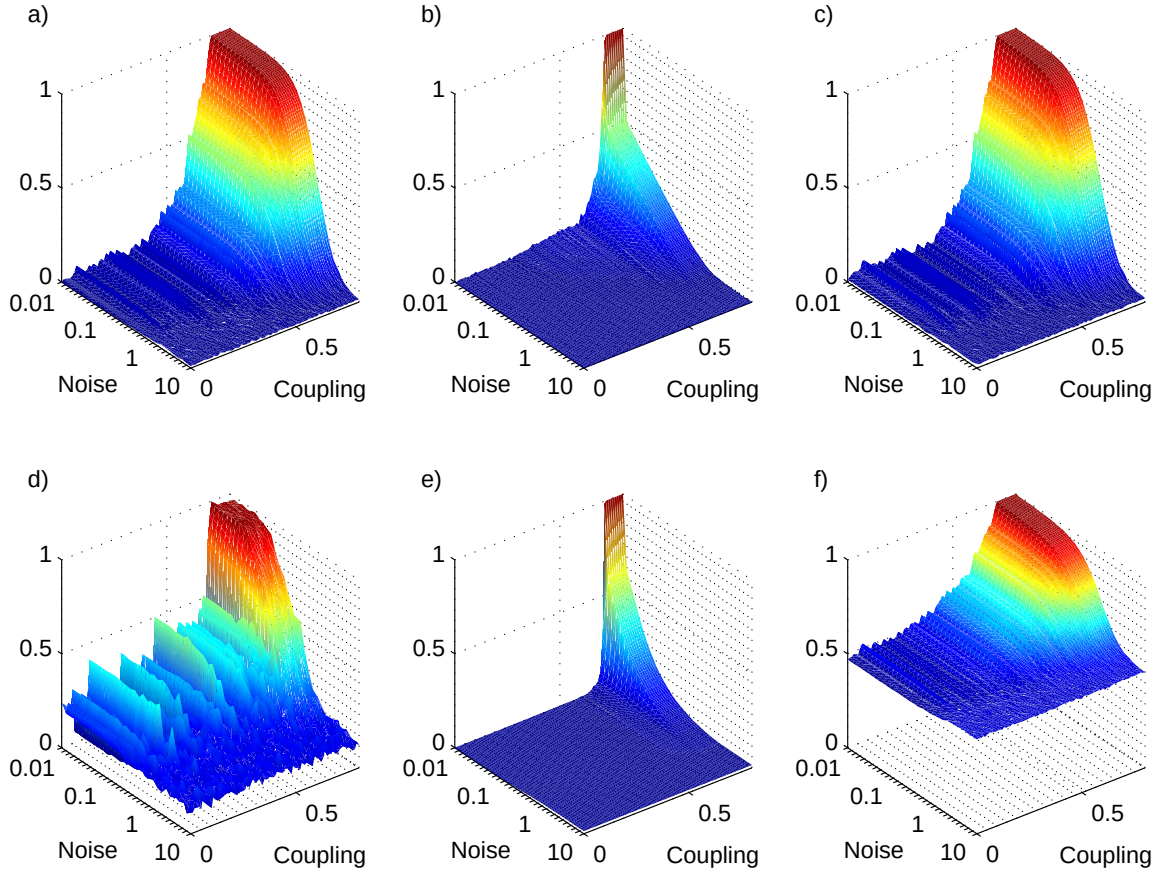


Figure 4.9: Coupled Hénon systems: Dependence on the coupling strength and the noise-to-signal ratio: a) Cross correlation C_{max}, b) Mutual information I, c) Index based on circular variance using the Hilbert Transform γ_{cv}^{H}, d) Index based on Shannon entropy using the wavelet Transform γ_{se}^{W}, e) Nonlinear interdependence S_s, f) Event synchronization Q. For the sake of completeness, for each measure the case of noise-free systems ($NSR = 0$) is plotted beyond the smallest noise-to-signal ratio $NSR = 0.01$.

all measures attain rather constant values close to the respective level obtained for uncoupled systems without noise. When looking at the complementary cross-section at maximum coupling $C = 0.8$ distinct differences between different measures regarding their decrease of values can be observed. While cross correlation C_{max} and the index based on circular variance using the Hilbert Transform γ_{cv}^{H} are able to maintain their high degree of synchronization up to intermediate ranges of the noise-to-signal ratio, the synchronization obtained for mutual information I and the nonlinear interdependence S_s is concealed rather quickly. The index based on Shannon entropy using the wavelet Transform γ_{se}^{W} and event synchronization Q rank between these extremes.

Results obtained for the coupled Rössler systems are shown in Fig. 4.10. For all measures, except for event synchronization, a decrease of values with increasing noise-to-signal ratio can be observed, again gradually for some measures and more abrupt for others. For the highest amount of noise a constant level not dependent on the coupling strength is obtained for each measure. Mostly this level is of the same order as the values obtained for uncoupled systems without noise, but for γ_{se}^{W} it is even below this level. This effect can be explained by the quasi-periodicity of the two Rössler systems which is overshadowed by the noise. For event synchronization the initial synchronization collapses with a little

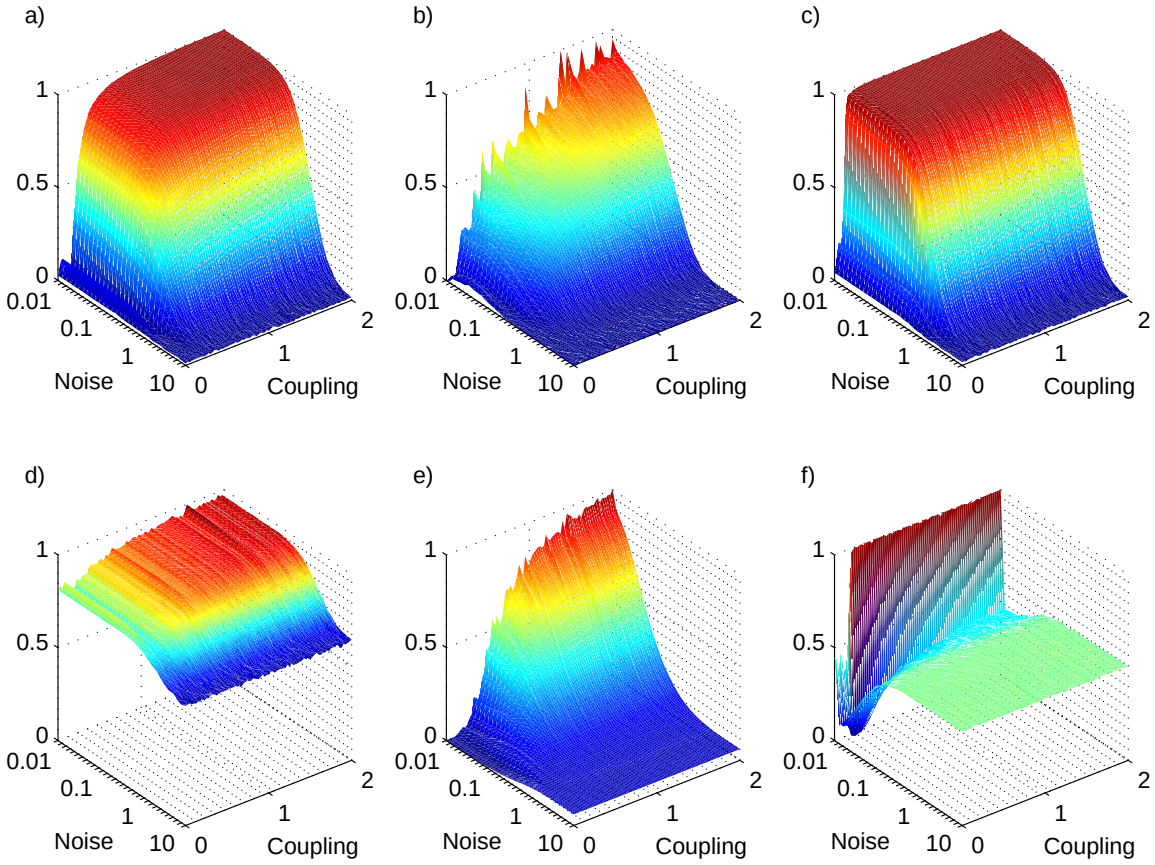


Figure 4.10: Same as Fig. 4.9, but this time for the coupled Rössler systems.

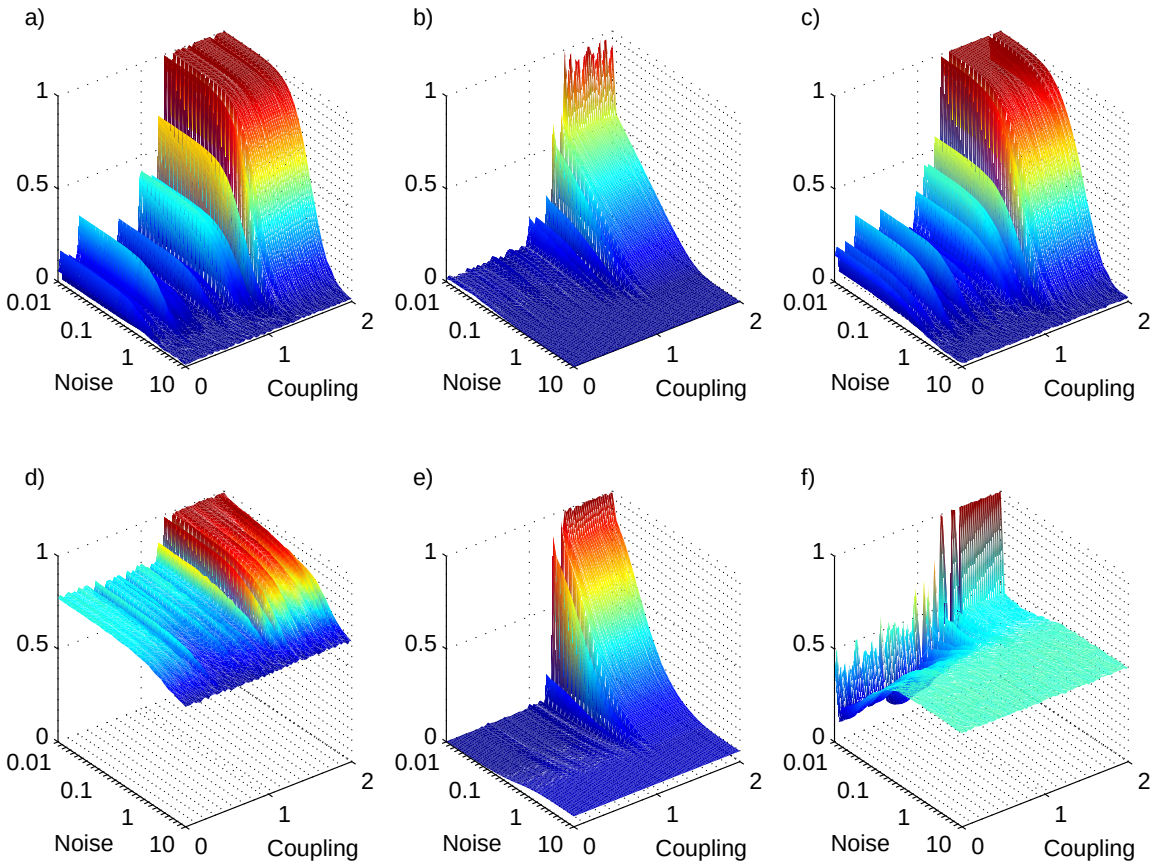


Figure 4.11: Same as Fig. 4.9, but this time for the coupled Lorenz systems.

noise. Here especially in the uncoupled case very low values are attained. With a further increase of the noise-to-signal ratio the level of the systems without coupling and without noise is reached. As displayed in Fig. 4.11 the behavior obtained for the coupled Lorenz systems is very similar to the case of the coupled Rössler systems. The same effects as described above take place with the only difference that more fluctuations can be observed which, however, disappear for higher noise-to-signal ratios.

So far the robustness against noise has been described qualitatively, in order to compare the different measures quantitatively again the measure of order M is used. It allows to track how the order obtained for noise-free systems is destroyed with increasing noise-to-signal ratio. This dependence on the noise-to-signal ratio is plotted exemplarily for the coupled Hénon systems in Fig. 4.12. The values of order for the noise-free systems are depicted on the left side of the smallest noise-to-signal ratio. For most measures these degrees of order are still maintained for lower amounts of noise. Mutual information I and the index based on Shannon entropy using the wavelet Transform γ_{se}^{W} even exhibit a slight rise in order and maximum values are obtained for intermediate noise-to-signal ratios. But from a certain level of noise for each measures a decrease of order can be observed.

However, the measures with the highest initial order do not necessarily prove to be most

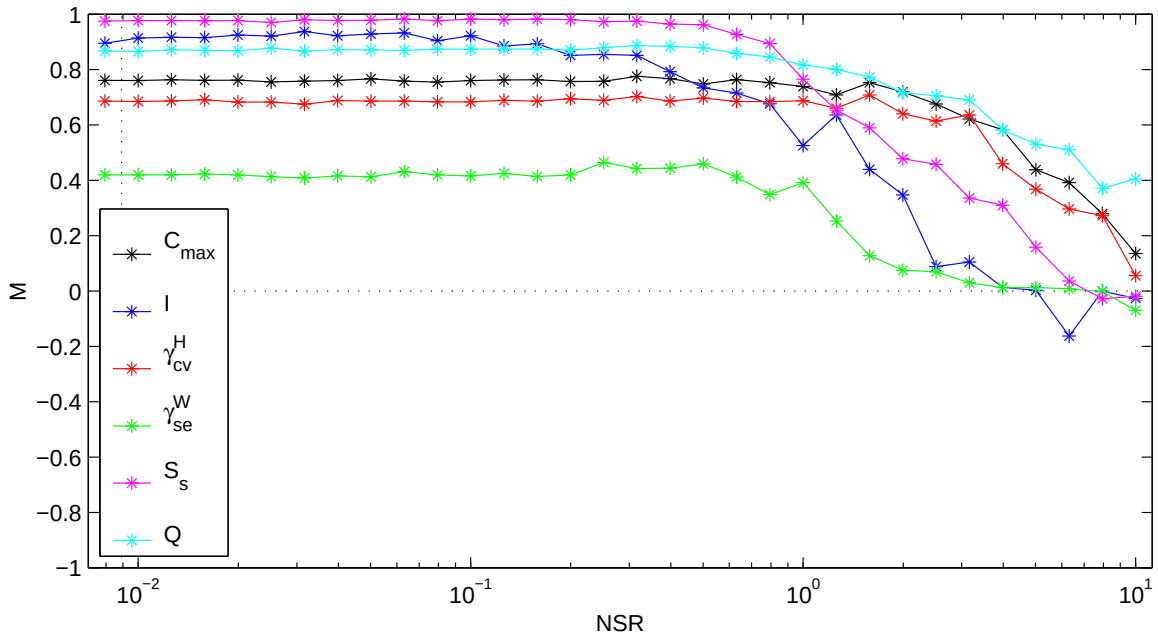


Figure 4.12: Coupled Hénon systems: Dependence of the measure of order M on the noise-to-signal ratio for the same measures as before. Also in this plot the case of noise-free systems ($NSR = 0$, already displayed in Fig. 4.8) is depicted at the left side of the smallest noise-to-signal ratio $NSR = 0.01$.

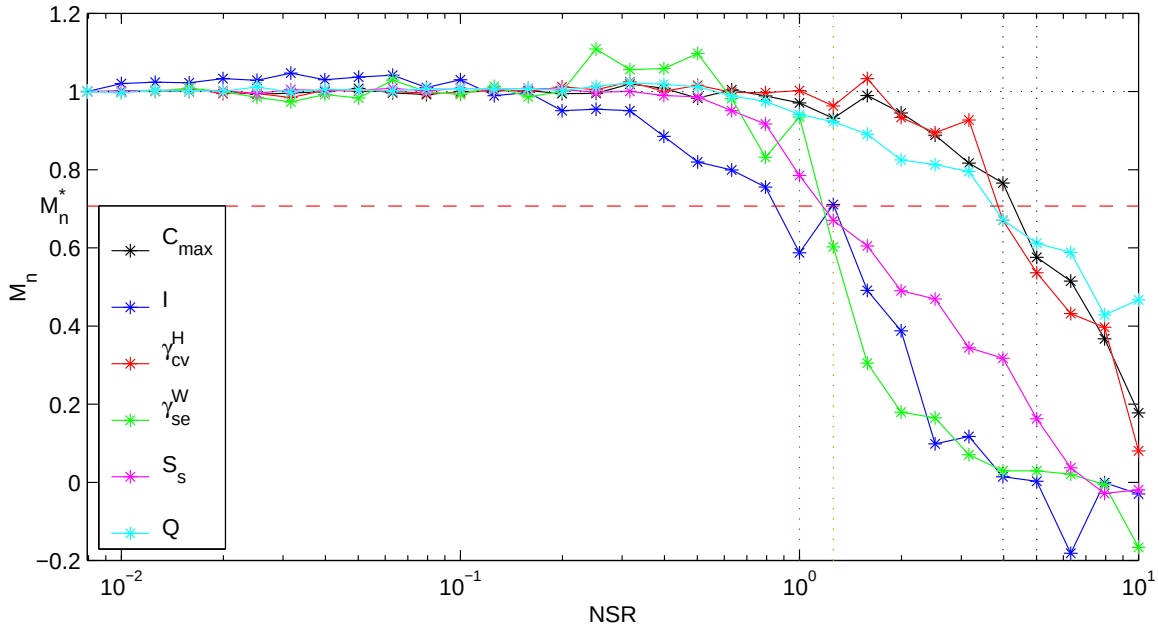


Figure 4.13: Same as Fig. 4.12, but for the normalized measure of order M_n. The dashed red line marks the threshold value $M_n^* = 1/\sqrt{2}$. Vertical lines indicate the critical noise-to-signal ratios NSR_C of the respective measure for the coupled Hénon systems.

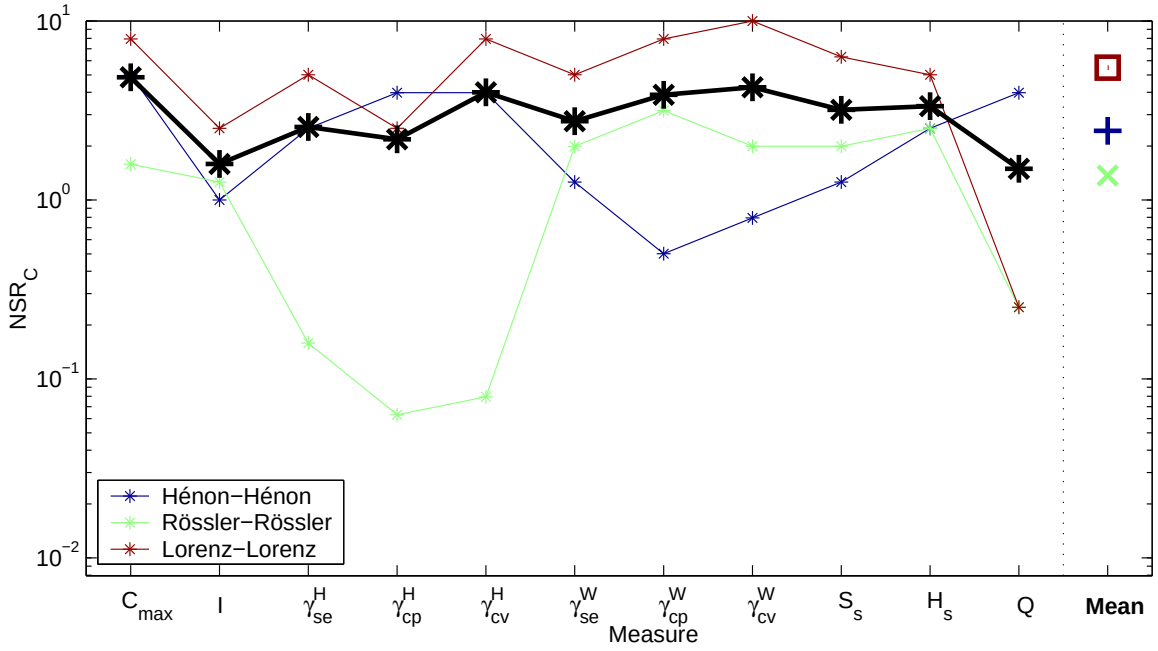


Figure 4.14: Over-all comparison for model systems contaminated with white noise: Critical noise level NSR_C for each measure and every system. The thick black line depicts the mean over the three model systems. In the last column for each model system the mean value over measures is displayed.

robust as can be seen for the nonlinear interdependence S_s and the mutual information I. These measures are most ordered in the absence of noise but this order starts to vanish already for intermediate noise levels. Finally, for high amounts of noise lowest values of order are rendered. To compare the robustness of the different measures, the critical noise-to-signal ratio NSR_C is determined as the first noise-to-signal ratio after the normalized profiles of order have crossed the threshold $M_n^* = 1/\sqrt{2}$ (cf. Fig. 4.13).

In Fig. 4.14 for each measure and every model system the critical noise-to-signal ratio NSR_C is shown. Furthermore, again the average values over measures as well as over systems are given. Results for different measures are more consistent than results for different systems. In the comparison of the measures highest robustness is obtained for cross correlation followed by the two indices based on circular variance. The least robust are mutual information and event synchronization. For mutual information this is due to the fact that parameters have been chosen to resolve the smallest possible scale (i.e., only one nearest neighbor has been regarded). Of course this scale is affected considerably by even small amounts of noise. On the other hand, events have been defined as local maxima and minima in order to yield a sufficient statistics and by this the robustness of the single event is not guaranteed when the signals become contaminated by high levels of noise. As for the different model systems, the order obtained for the coupled Lorenz systems without noise proves to be most robust followed by the Hénon and the Rössler systems. Regarding

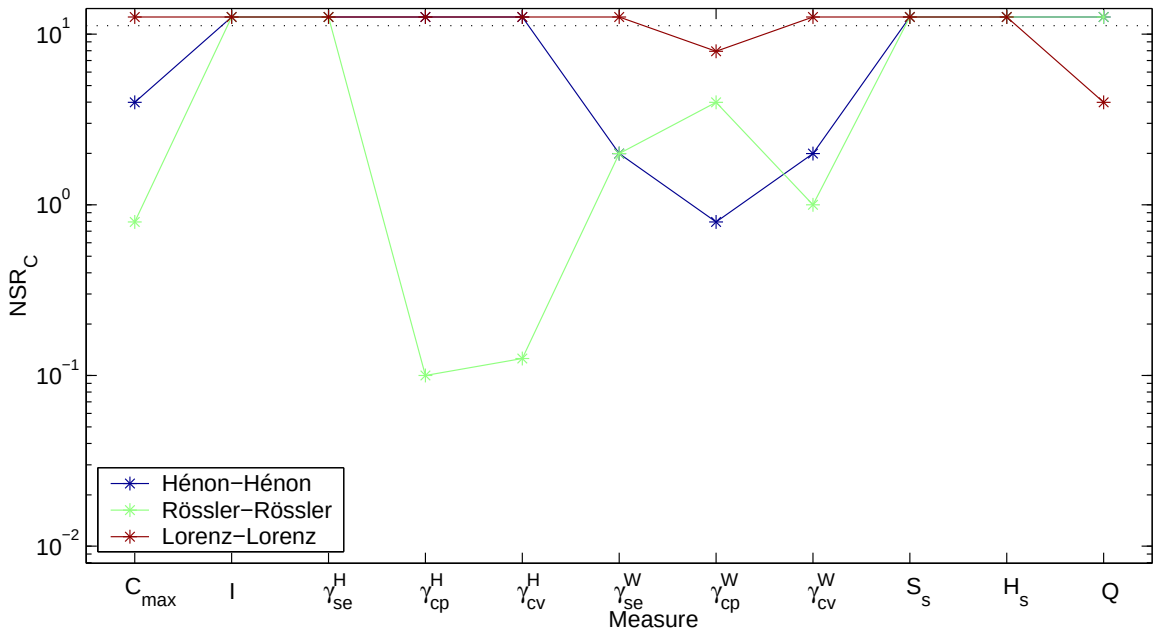


Figure 4.15: Over-all comparison for model systems contaminated with iso-spectral noise: Critical noise level NSR_C for each measure and every system. If the threshold M_n^* is not crossed, this is marked by an NSR_C-value beyond the maximum noise-to-signal ratio. Due to this effect an average value of NSR_C over different measures or systems is not useful and thus omitted.

the opposite ranking in the noise-free case (cf. Fig. 4.8), it seems that the more ordered a system has been without noise, the less noise is necessary to destroy this order.

As for the case of iso-spectral instead of white noise, results for the equivalent analysis are given in Fig. 4.15. For this type of noise for the different systems the same ranking is obtained as in the case of white noise. But here for some measures and some coupling schemes M_n does not cross M_n^* at all, i.e., the order of the noise-free case is not destroyed when the coupled systems become contaminated with noise.

For the Hénon systems in these cases the dependence on the noise-to-signal ratio is maintained or only slightly weakened, but for the coupled Rössler and Lorenz systems a new effect can be observed. This effect is illustrated exemplarily for the coupled Rössler systems in Fig. 4.16. For some measures (e.g., mutual information I and the nonlinear interdependence S_s) the rise of synchronization with increasing coupling strength first starts to fade away with increasing NSR but then gets again more pronounced for high values of NSR. In these cases for high coupling strengths a (spurious) synchronization between the contaminating noise-signals can be observed.

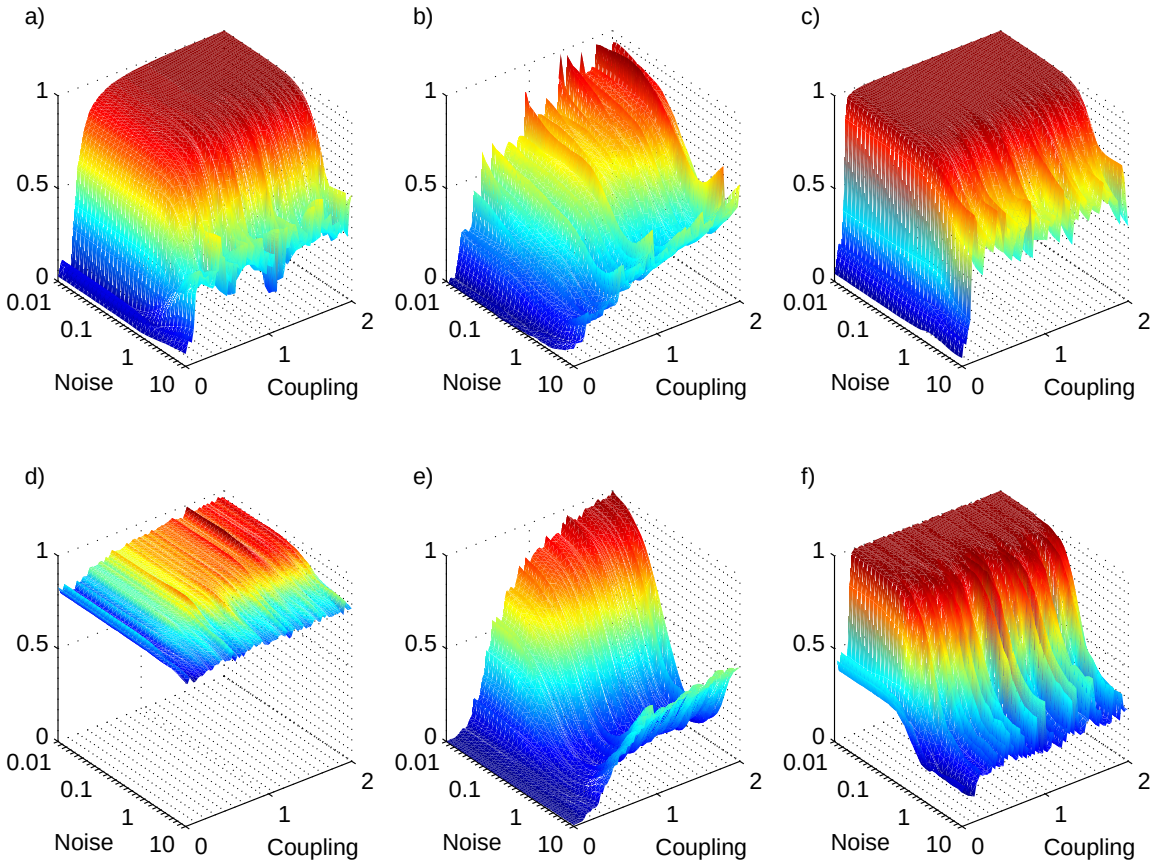


Figure 4.16: Same as Fig. 4.9, but this time for the coupled Rössler systems and isospectral noise.

4.3 Discussion

The aim of this study was to evaluate whether the analysis of model systems can contribute useful information in order to decide which measure of synchronization is most suitable for an application to field data. To address this aim, in the first part of this study a comparison of the different measures of synchronization was carried out regarding their capability to reflect different degrees of coupling between two model systems. The dependence on the coupling strength was evaluated using a measure of order M. There are some caveats in this evaluation since the approach to use M is built on some very simplifying assumptions. The most fundamental assumption is that an increase of coupling necessarily leads to an increase of synchronization. Furthermore, this should hold true for all types of synchronization no matter which model systems are investigated and which coupling schemes are

employed.

It is quite simple to construct counter-examples in which only one type of synchronization takes place while all other types are not at all or much less affected (confer also the ongoing discussion on generalized synchronization versus phase synchronization [114, 179]). In such a case only the respective measures designed to detect the type of synchronization of interest should increase while all other measures should remain unchanged. Thus the second assumption does not hold in general. But also the first assumption is not always fulfilled and it might happen that certain properties of the individual systems or some special features in the interaction of the two systems diminish or even reverse the effect of a higher coupling.

As could be seen when comparing the plots of the first driver component versus the first responder component in Figs. 4.1-4.3 and the courses of the maximum Lyapunov exponent of the responder system in Figs. 4.5-4.7 also the different systems investigated in this study exhibit very different behavior in their transition to synchronization with increasing coupling strength. In particular for the coupled Lorenz system amidst an over-all increase of synchronization distinct regimes of anti-synchronization and, correspondingly, high fluctuations of the Lyapunov exponent could be observed. On average this behavior was reflected by the different measures and, furthermore, correctly tracked by the measure M showing the lowest order for this coupling scheme.

Due to these caveats in the analysis of model systems it seems that the notion of a 'controlled' setting is valid only up to a certain extent. Each and every time a careful examination of the underlying model systems and coupling schemes is unavoidable. Thus the question which measure is best suited can not be answered in general. Rather it seemed that the result is clearly dependent on the model system and the coupling scheme. This is in line with the rather inconsistent results that were obtained for different systems and measures.

But even if the underlying assumptions do not hold rigorously and the order can not be used as a strict criterion to compare measures in the first part of this study, M still proves useful as a means to reduce information since it compresses the dependence on the coupling strength into a single value. The crucial point is the interpretation of a low value of M indicating a non-monotonic course. Is such value due to a failure of the measure to distinguish between different levels of coupling or does the measure correctly reflect certain peculiarities in the systems' transition to synchronization? Most of all, such a value should encourage a more careful investigation of the respective dependence on the coupling strength.

Whereas there were some caveats in the first part of this study, the application of the measure of order in the second part remained unaffected. Given the order without noise, M served as an appropriate means to track the changing of this order due to the additive noise. Using a threshold criterion designed to quantify the robustness against noise, the influence of Gaussian white noise was evaluated. Regarding the comparison of the different systems

again robustness was dependent on the system under investigation. For systems with high order in the absence of noise the least robust results were obtained. As for the comparison of the different measures, on average all measures proved to be rather robust against white noise. Slight advantages were observed for cross correlation and the phase synchronization indices based on circular variance. This is consistent with results obtained in Ref. [99].

Concerning contaminations with iso-spectral noise, for some model systems and some measures a (spurious) synchronization between the contaminating noise signals could be observed. This effect was also reported for bidirectionally coupled Rössler systems in Ref. [99]. It can be explained by the narrowness of the power spectrum of the respective systems [99], which, by construction of iso-spectral noise (cf. Appendix B.2), leads to the same narrow power spectrum of the two noise signals added. In case of strong coupling, the peaks in the power spectrum of the coupled systems (and thus those of the noise signals as well) align at the same frequency [114]. For low levels of the noise this causes synchronization of the coupled signals, for high levels of noise (spurious) synchronization between the noise signals can be detected. In the coupling regimes investigated here this was most prominently observed for the nonlinear interdependencies.

Summing up, the question which measure is best suited for the application to field data (e.g., of biological or medical origin) can not be answered a priori. Regarding the capability to distinguish different coupling strengths, due to the individual properties and peculiarities of every system, except for very special cases an obvious and objective criterion to compare different measures is not at hand. However, as for the robustness against noise, judgmental statements can be made. Depending on the conditions of data acquaintance and the expected noise level, the measure should be chosen accordingly. On the other hand, in most cases the measure to be applied to a certain task can be chosen rather pragmatically as the measure which most reliably yields valuable information (e.g., information useful for diagnostic purposes) in test applications. Such an application is presented in the next Chapter.

Chapter 5

Application to the EEG of epilepsy patients

In this Chapter the bivariate measures of synchronization introduced in Chapter 3 and tested on model systems in Chapter 4 are applied to real data obtained from a system with poorly understood dynamics. For this purpose the electrical activity of the brain recorded from epilepsy patients by electroencephalography (EEG) is analyzed. Due to the physiological and pathophysiological variations in the brain these electroencephalographic time series represent a prominent example of biological data showing a rich and diverse appearance and therefore constitute a great challenge for the application of methods derived from the theory of dynamical systems. Since the EEG is typically measured simultaneously in different regions of the brain, it is highly suited to be investigated by measures of synchronization. First in Section 5.1 a short introduction to the disease epilepsy and the most important tool of diagnosis, the EEG, is given. The EEG is then analyzed to investigate to which extent the different measures of synchronization are capable to reflect the temporal variability of the epileptic process. In particular, the measures are tested for their ability to successfully address the most challenging clinical task, namely the prediction of epileptic seizures. In these investigations a great deal of attention is paid to the statistical validation of seizure predictions, a very important aspect which is frequently neglected. In Section 5.2 a new approach to address this issue, the method of measure profile surrogates, is introduced and compared against the existing method of seizure times surrogates. Due to feasibility the latter method is applied in the comparison of the bivariate measures by means of a comprehensive statistical evaluation of the predictability of seizures in Section 5.3. Furthermore, the combined use of measures is discussed based on observed correlations between the different measures. Finally, in Section 5.4 the new method of measure profile surrogates is illustrated by exemplarily evaluating the predictive performance of two measures of synchronization.

5.1 Epilepsy and the electroencephalogram

The word *'epilepsy'* is derived from the Greek verb $\varepsilon\pi\iota\lambda\alpha\mu\beta\alpha\nu\varepsilon\iota\nu$ (*epilamvanein* = 'to be seized' or 'to be attacked'). In ancient history epilepsy was considered to be the sacred disease ("morbus sacer") representing attacks by the gods or evil spirits. Nowadays it is understood that the hallmark of epilepsy, the epileptic seizure, is the (sudden) occurrence of an intermittent malfunction of the brain. In more detail, it represents the clinical manifestation of an excessive, synchronous, abnormal high-frequency firing pattern of neurons in the brain [32, 107]. Epileptic seizures are fundamentally divided into two main classes - generalized and partial. While generalized seizures are bilateral and involve almost the entire brain, partial (focal) seizures have clinical or electroencephalographic evidence of a localized onset and usually stay confined to one hemisphere [73].

The onset of a focal seizure is assumed to be initiated by abnormally discharging neurons in a circumscribed region of the brain, the ictal onset zone also called the *epileptic focus*. These so-called burster neurons are altered in their fundamental excitability and start to recruit and entrain neighboring neurons into some 'critical mass' of hypersynchronous activity [167, 177, 23]. This is accompanied by a lack of inhibition which allows the spreading of these pathological discharges both in local areas and also, via preferred synaptic pathways, into distant brain regions. During the course of seizure propagation the involved neuronal populations are no longer able to maintain their usual coordinated physiological information processing and clinical symptoms become evident. These are mostly specific for the affected brain region and depending on seizure type and severity comprise loss or impairment of consciousness, absence, tonic contractions or clonic convulsions (or both combined), localized paralysis as well as sensory, autonomic or psychic symptoms [176].

But epilepsy is much more than seizures. Also temporally distant from these rare, so-called *ictal* states (lat. ictus = seizure), i.e., during the *inter-ictal* state, many neurons exhibit different forms of epilepsy-specific pathophysiological behavior. The most pronounced manifestation is the so-called *paroxysmal depolarization shift (PDS)*, a distinct shift of the resting membrane potential accompanied by bursts of action potentials (up to 800 per second) and followed by periods of inhibition [39, 94, 95]. Most often such events occur isolated, and relatively little perturbation of function can be detected. The area of the cortex that generates these *inter-ictal epileptiform discharges* is called the irritative zone. It is clearly related but often not identical to the ictal onset zone.

Approximately 5% of the world's population have an epileptic seizure at least once in their life, but this does not mean that they suffer from epilepsy, which is only the case when such seizures occur chronically. The estimated lifetime cumulative incidence (rate at which new cases of a disease occur) of epilepsy is 3% and the prevalence (frequency of all current cases of a disease) is 0.5%. Approximately 60% of patients undergo partial seizures. Although most patients have few seizures or seizures that are well controlled by antiepileptic or anticonvulsive drugs, an estimated $5 - 10\%$ are medically intractable and

may become candidates for surgical treatment, i.e., the tailored resection of the epileptic focus [32, 11]. This highly invasive therapy is feasible only if an exact localization of the epileptic focus and its delineation from functionally important areas can be achieved by an extensive presurgical evaluation making use of various clinical and diagnostic tools. First of all different neuroimaging techniques are applied to identify and delineate potential epileptogenic brain lesions. These are supplemented by a variety of basic neurological and neuropsychological examinations mainly aiming at an estimation of the extent and pattern of cognitive deficits possibly caused by an underlying brain disruption.

Despite these modern techniques the most widely used diagnostic tool is still the acquisition of the EEG, a technique dating back to the late 20s of the past century [18]. Nowadays the EEG is clinically useful in many neurologic disorders but its use in epilepsy is unique. It allows a physiologically meaningful visualization of the complex activity of the brain during and between seizures. Thus the easiest and most reliable method to exactly localize the epileptic focus is the electroencephalographic recording of several epileptic seizures and the identification of the ictal onset zone by visual inspection of the EEG. If non-invasive techniques are not sufficient to yield the desired information for an unequivocal localization, to date usually multichannel recordings are performed using intracranial monitoring techniques like the electrocorticogram (ECoG) and the stereo-EEG (SEEG). In these cases the brain electrical activity is recorded directly from the surface of the brain and from specific structures within the brain. In Fig. 5.1 a typical intracranial implantation scheme as it is used at the Department of Epileptology, University of Bonn, Germany, is depicted. The excellent signal to noise ratio and the outstanding temporal and spatial resolution yielded by implanted electrodes allow a substantially increased precision in the design of the surgical intervention justifying the high degree of invasiveness[1]. Recent technical progress like the development of combined digital video-EEG monitoring systems as well as the use of modern computers together with high capacity storage facilities in principle enable the processing and analysis of continuous long-term multichannel recordings in real time.

The human brain consists of approximately 10^{11} neurons with a total of 10^{14} to 10^{15} synaptic connections. The electrical activity recorded at each electroencephalographic channel is assumed to be generated by postsynaptic sum potentials of a very large number of individual neurons, each of them showing a highly nonlinear discharging behavior (and sometimes even bursting). The superposition of all these elementary processes of the central nervous system typically leads to EEG voltage amplitudes of the order of some mV. In Fig. 5.2 the SEEG of the left and right hippocampal formation recorded from a patient with the epileptic focus located in the left temporal lobe is depicted. The upper and lower traces show

[1]However, noninvasive multichannel magnetoencephalography (MEG) and scalp EEG analysis, together with magnetic resonance imaging (MRI) and single photon emission computed tomography (SPECT), are expected to supersede invasive EEG recording techniques in many candidates for epilepsy surgery. In the near future, when these combine with increased local cerebral metabolic information from more sensitive positron emission tomography (PET) and expected developments in metabolic applications of MRI, the need for depth and subdural recording should continue to diminish [32].

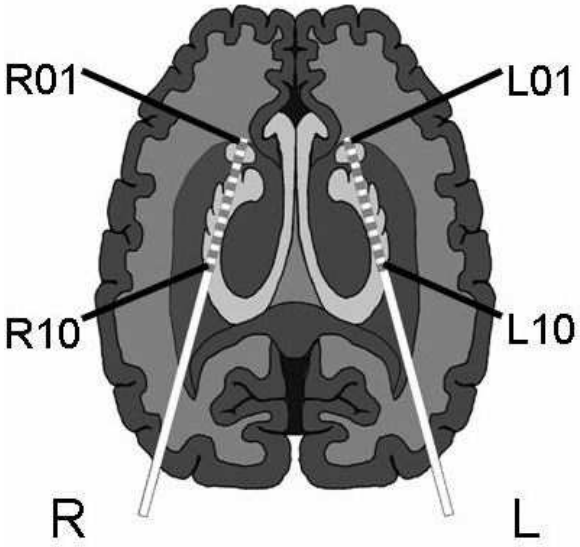


Figure 5.1: Schematic view of implanted depth electrodes TL and TR. Each depth electrode is equipped with 10 contacts of a nickel-chromium-alloy (diameter: 1 mm, length: 2.5 mm, inter-contact distance: 4 mm).

an inter-ictal and an ictal interval, respectively. The seizure onset in the left hippocampal formation is easily recognizable.

In Figs. 5.3 and 5.4 two time series representative of the inter-ictal and the ictal interval are shown together with the respective power spectrum and state space portrait.

The first measures calculated from EEG data have been linear ones, namely the relative power contained in certain characteristic frequency bands of the EEG. These comprise: δ [0-4 Hz], ϑ [4-8 Hz], α [8-13 Hz], β [13-30 Hz] and γ [30-48 Hz]. These intervals have been defined according to predominant activities in the EEG related to certain states of vigilance and/or pathology. Other traditional linear measures used in EEG analysis include the Hjorth parameters [48], time domain based measures derived from the autocorrelation function and statistical moments like the variance. Recently a variant of the latter resurfaced as accumulative energy [88].

In most studies the EEG is analyzed retrospectively by means of a moving window technique [16] dividing the recordings into short segments typically of the order of some tens of seconds. Univariate nonlinear measures that have been applied using this technique comprise estimates of an effective correlation dimension [81, 84, 8, 15], the correlation density [93], the measure ξ quantifying the fraction of nonlinear determinism [4, 5, 10], the largest Lyapunov exponent [52, 174, 72], measures derived from the theory of symbolic dynamics [63, 68] and the loss of recurrence as a measure for non-stationarity [134, 133, 132]. Frequently the method of surrogate data [166] is used to focus on specific nonlinear properties of the EEG [35, 7]. Furthermore, bivariate measures have been used to perform a univariate analysis on a single channel: One time window is kept as a fixed reference and the

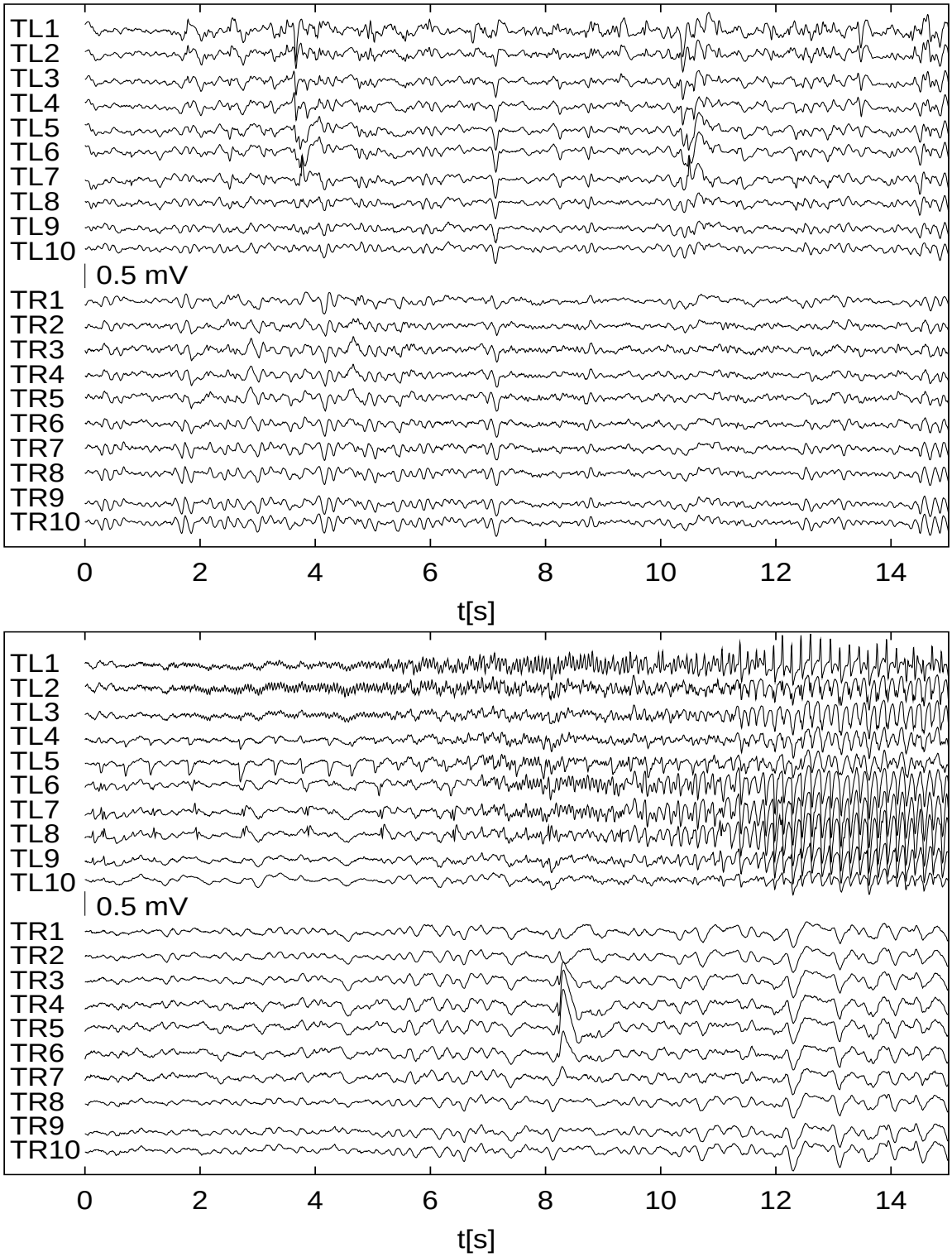


Figure 5.2: Two exemplary multichannel SEEG recordings of the same patient with mesial temporal lobe epilepsy originating in the left hippocampal formation. Top: Inter-ictal activity. Bottom: Onset of a seizure and beginning ictal activity.

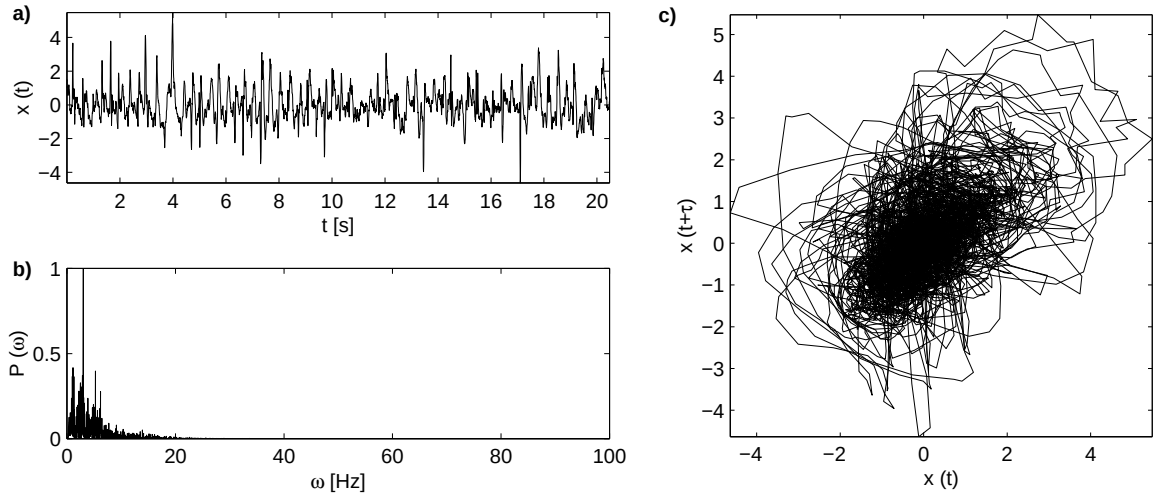


Figure 5.3: Inter-ictal EEG: a) Exemplary time series with $N = 4096 \ \ (\approx 20 \ sec)$. b) Normalized power spectrum. c) Two-dimensional state space portrait (time delay $\tau = 6$).

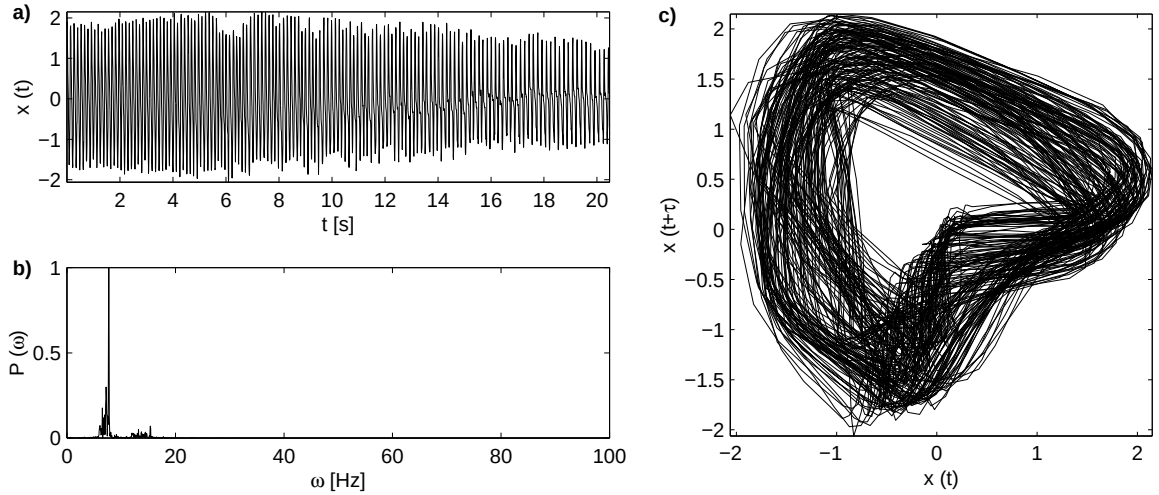


Figure 5.4: Ictal EEG: a) Exemplary time series with $N = 4096 \ \ (\approx 20 \ sec)$. b) Normalized power spectrum. c) Two-dimensional state space portrait (time delay $\tau = 6$).

dynamical changes are tracked by comparing the succeeding windows of the same channel against this reference [79].

It is only recently that bivariate measures have been more widely applied to human EEG data. In Refs. [14, 13, 69, 67, 66] the nonlinear interdependencies S and H have been used to investigate spatio-temporal patterns in the brain of epilepsy patients. The index based on circular variance using the Hilbert Transform γ_{cv}^{H} (also termed mean phase coherence) as a measure of phase synchronization has been applied to EEG data from epilepsy patients by Mormann and colleagues [98, 104, 101, 100]. In the latter two studies also the linear cross correlation has been used. Phase synchronization with the phase based on the

wavelet transform has been introduced in Refs. [71, 70] to quantify synchronization between different areas of the brain during cognitive tasks and also applied in Ref. [61]. In Ref. [77] phase synchronization measures based on the Hilbert transform and those based on the wavelet transform have been compared also using EEG data. Quian Quiroga and colleagues performed a comparison of the performance of different bivariate approaches in a case study analyzing EEG data from an animal model of epilepsy [128]. Further studies dealing with synchronization in the EEG include Refs. [53, 111].

5.2 Prediction of epileptic seizures and its statistical validation

Most epileptic seizures occur ”like a bolt from the blue”, although there are several environmental and physiological factors known as common seizure precipitants, e.g., sleep deprivation, hyperventilation, drugs and alcohol or even more frequently withdrawal of alcohol. Also the level of consciousness, the state of vigilance, and nonspecific emotional situations in some susceptible individuals are regarded as potential modulating factors [3]. In reflex epilepsies seizures are triggered by highly specific stimuli, either visual, auditory or somatosensory in nature [164]. Finally, some patients even can induce seizures by themselves. But still, depending on the patient, seizures may occur frequently or infrequently, only at night or after awakening, in a cyclic pattern, only with highly specific triggers, in many other permutations, and, most commonly, without any apparent predictability.

To date it remains an open question whether this unpredictability of epileptic seizures can be overcome by the analysis of the electroencephalogram using different characterizing measures. In the EEG most seizures are easily recognized by their rhythmic high amplitude activity (cf. Figs. 5.2 and 5.4) reflecting the abnormal synchronization of a large number of neurons [32]. With this in mind, the question arises whether it is also possible to discriminate the intervals preceding seizures (pre-ictal periods) from the intervals far away from any seizure activity (inter-ictal periods). Provided that the analysis of the EEG would allow to reliably detect a pre-ictal state[2] in a prospective setting, new therapeutic possibilities (e.g., seizure prevention strategies) could be envisaged [29].

Therefore, it is not surprising to find a very rich and diverse literature dealing with the prediction of epileptic seizures. Starting from earliest approaches based on pattern recognition [173] and spike detection [40, 74] at first mostly univariate measures were employed, either linear [135, 28] or nonlinear [52, 30, 84, 93] in nature. Later efforts reporting the

[2]Note that there is a clear distinction between the term ‘pre-ictal interval’ and the notion of a ‘pre-ictal state’. While the former term defines the time interval before a seizure, the latter is used to describe a potential distinct state reflecting a disposition towards a transition to a seizure. Whenever there is a seizure, there has been a pre-ictal interval before by definition. The existence of the pre-ictal state, however, is still to be proven.

predictability of epileptic seizures by applying these two different kinds of univariate measures include Ref. [88] and Refs. [80, 105], respectively. It is only recently that bivariate [104, 51, 101] or multivariate [149] measures have been added to the wide range of approaches reportedly being able to detect a pre-ictal state. The current impact of this topic is stressed by recent controversies about the relevance of nonlinear approaches for the prediction of epileptic seizures [97, 100] and even more strikingly by studies raising doubts about the reproducibility of reported claims [26, 15, 72]. For an overview refer to Refs. [87, 89, 86].

Typically, in a study on the predictability of epileptic seizures first a certain characterizing measure is calculated from multi-channel EEG using a moving-window technique. The resulting measure profiles are then scanned for prominent features which can be related to the actual seizure times. These features might be drops or peaks (e.g., quantified as threshold crossings) or any other distinct pattern in the measure profile. In a second step the measures' capability to distinguish the pre-ictal from the inter-ictal interval is evaluated with a test statistics quantifying the occurrence of these features relative to the seizure times and resulting in some kind of performance value. If this performance is high, it might on the one hand reflect the existence of a pre-ictal state and the capability of the applied measure to detect it, but it might on the other hand also be due to statistical fluctuations or some (unknown) bias in the algorithm.

In the design of a seizure prediction algorithm there are many subtle points to be considered carefully. Typically the calculation of the measure as well as the later statistical evaluation involves the choice of certain parameters. In this context, much care needs to be taken to avoid in-sample optimization of these parameters. Certainly, what is true for a single measure holds also for a larger number of different measures. The application of a huge variety of measures to the EEG might yield a measure with seemingly good results just by chance (particularly on a limited database). Secondly, there are many degrees of freedom in the statistical evaluation. In the case of univariate measures often a best channel selection is performed, and for bivariate measures, which evaluate the dependencies between two channels, there are even more channel combinations to choose from. Finally, the same argument holds for different patients as well. Provocatively speaking, many (spurious) claims about the existence of a pre-ictal state might just be due to some 'best parameter', 'best measure', 'best channel' and/or 'best patient' selection.

Since usually these problems cannot be solved in the design of a seizure prediction statistics, the question arises how to interpret a non-zero performance value. This value might correctly reflect the existence of a detectable pre-ictal state, but it might also be the spurious result of statistical fluctuations. Therefore, to assess the performance yielded by a seizure prediction algorithm, a method to judge its statistical validity is needed. The result should be verified against some null hypothesis and its level of significance should be estimated. This can be achieved using the concept of surrogates [166, 155], in which the validity of a given test result is evaluated by applying the test not only to the original data but also to an ensemble of surrogate data generated by means of a Monte Carlo randomization. In this

case the null hypothesis H_0 to test against can be stated as follows: "The measure under investigation is not suited for seizure prediction." If this null hypothesis is fulfilled it might be due to two different reasons. Either a pre-ictal state does not exist (and thus there is no measure suited for seizure prediction) or a pre-ictal state does exist, but the measure is not able to detect it. On the other hand the null hypothesis can only be rejected if both inverse conditions are fulfilled: There are specific changes before a seizure and the measure is sensitive to these changes.

The performance of any seizure prediction algorithm crucially depends on whether the sequence of actual seizures is matched by some corresponding structure in the measure profiles. Therefore to test for statistical significance of a good performance by using the method of surrogates, any such structure should be destroyed by the randomization. Essentially, this can be done in two different ways. Andrzejak and colleagues [9] recently introduced the method of seizure time surrogates in which the seizure times are randomized, while the measure profiles are maintained. In this thesis the *method of measure profile surrogates* is proposed, a new and complementary approach, in which the seizure times are kept fixed and instead a constrained randomization of the measure profiles is performed using the method of simulated annealing [65, 64].

The concept of surrogates as a means to test a null hypothesis is applied equivalently in both methods: The seizure prediction algorithm is run using the original measure profiles (seizure times) and its performance is compared to the results of the same algorithm using an ensemble of measure profile surrogates (seizure time surrogates). Provided that a pre-ictal state exists and the prediction algorithm is able to detect it, its performance should be highest for the original measure profiles (seizure times). In this case the null hypothesis could be rejected at the level of significance determined by the number of measure profile surrogates (seizure time surrogates).

Both methods are reasonable statistical approaches to address the correspondence between measure profiles and seizure times, but the method of measure profile surrogates is the more natural choice: Usually, within the method of surrogates the property to test for is destroyed in the surrogates. And in the present case the object under investigation is the measure rather than the sequence of seizures. More specifically, the aim is to test the measure for its capability to extract information from the EEG that enables the prediction of the original seizures and not to test the sequence of seizures whether they resemble the measure profiles.

Within either of these methods there are certain properties of the original which should be preserved for the surrogates. In the case of seizure time surrogates it has been proposed to preserve the total number of seizures, the distribution of time intervals between consecutive seizures, and, as the case may be, any clustering of the seizures [9]. This has been achieved by a random permutation of the original seizure intervals. As indicated already in Ref. [9], this approach is applicable only if the number of seizures and hence the number of possible permutations is large enough to allow the generation of the number of surrogates needed

to obtain the desired significance. The number of possible permutations is even further diminished in the presence of recording gaps, since then permutations have to be discarded whenever one of the surrogate seizures falls into such a gap. To prevent a bias between the original and the surrogates, also ictal and post-ictal intervals as well as all other events known to possibly cause changes in the EEG have to be avoided (For the sake of brevity, throughout this thesis these intervals will also be referred to as recording gaps). But even when a sufficient number of permutations remain, much care has to be taken to ensure that the inter-ictal interval as well as any possible pre-ictal interval are equally well represented in the original and in all of the seizure time surrogates.

In the method of measure profile surrogates these issues are easily addressed, since the original seizure times are not changed at all. Rather they are correctly considered as given conditions based upon which the measure profiles are probed for their predictive performance. But also in this method there exist some constraints, i.e., properties which should be extracted from the original measure profile and imposed on the surrogate measure profiles. First of all, a suitable randomization should maintain all existing recording gaps. Furthermore, it is advisable not only to preserve the amplitude distribution but also to maintain essential parts of the autocorrelation function. The preservation of these features guarantees that, when regarded independently from the seizure times, the original as well the surrogate measure profiles can be considered as a possible original measure profile. The most important property that might remain different is the correspondence to the seizure times and this is exactly the property under investigation.

Despite the aforementioned advantages of the method of measure profile surrogates, in this thesis both methods are applied. This is due to the high computational cost of the method of measure profile surrogates which to date renders the application of this method to a very large database infeasible. Since the comparison of the bivariate measures regarding their capability to predict epileptic seizures is carried out by analyzing quasi-continuous long-term recordings from nine patients, performances yielded are statistically evaluated by means of the method of seizure times surrogates. Subsequently, the new method of measure profile surrogates is illustrated by exemplarily evaluating the predictive performance of two measures of synchronization analyzing recordings from one of these patients.

5.3 Statistical evaluation of the predictability of seizures

The final aim in the field of seizure prediction is the design of a prospective algorithm which is able to detect specific changes before an impending seizure when evaluating online the time profiles of a characteristic measure calculated in real-time. A successful implementation of such an algorithm is dependent on two important prerequisites. These are the existence of a pre-ictal state different from the inter-ictal state and the capability of the applied measure to detect it. Since pathological synchronization of neuronal ensembles is regarded as the main mechanism responsible for the generation of epileptic seizures [90] the bivariate measures of synchronization used in this thesis rank among the most prominent candidates to fulfill the latter demand. To evaluate which of these measures is best suited for the prospective detection of a pre-ictal state it is reasonable first of all to retrospectively compare their ability to distinguish between pre-ictal and inter-ictal intervals. Since the investigation on model systems did not lead to a reasonable exclusion of one or more measure and since the noise-to-signal ratio of the data to be analyzed is very low, all measures of synchronization and directionality will be probed for their predictive performance.

In case that no measure on its own would be capable to predict epileptic seizures with a sensitivity and specificity sufficient for a clinical application, it seems reasonable to test whether a combination of different measures could lead to a significant improvement in predictive performance (cf. Ref. [5]). However, before accomplishing this, it should be investigated to which extent different measures carry independent and non-redundant information. In the second part of this study this investigation is carried out for all bivariate measures analyzed in this thesis.

First a statistical approach is applied to measure profiles rendered from the analysis of quasi-continuous multi-day EEG recorded intracranially from nine patients (Section 5.3.1.1). The study is based on cleaned data due to a comprehensive preprocessing performed to identify and eliminate artifacts (Section 5.3.1.2). Amplitude distributions of pre-ictal and inter-ictal intervals are compared by means of Receiver-Operating-Characteristics without the use of any a posteriori knowledge. The analysis is performed with and without smoothing of the measure profiles, allowing different lengths of the pre-ictal interval and finally testing for both a pre-ictal decrease as well as an increase of the measures' values (Section 5.3.1.3). Since it is not known beforehand how a potential pre-ictal state manifests itself, different evaluation schemes are designed focussing either on global or local effects and using either a constant or an adaptive baseline (Section 5.3.1.4). As mentioned above, the obtained performance values are statistically validated by the use of seizure time surrogates (Section 5.3.1.5). Finally, in Section 5.3.1.6 the correlation between the different measures is quantified.

The whole study is an extension of an earlier study [103] in which the discriminative performance of nine bivariate measures of synchronization was evaluated by analyzing EEG

recordings from the first five patients of this study. In this earlier study a comprehensive comparison with univariate approaches was carried out as well. Differences between the extended results (Section 5.3.2) and the ones reported in Ref. [103] will be discussed in Section 5.3.3.

5.3.1 Methods

5.3.1.1 Data base and implementation of measures

This study evaluates the ability of the different measures of synchronization and directionality introduced in Chapter 3 to distinguish between pre-ictal and inter-ictal intervals. The degree of synchronization is quantified by symmetric measures like the linear cross correlation C_{max}, the mutual information I and the three different indices of phase synchronization with the phase of the time series being extracted either using the Hilbert transform (γ_{se}^H, γ_{cp}^H and γ_{cv}^H) or the wavelet transform (γ_{se}^W, γ_{cp}^W and γ_{cv}^W). Furthermore, the symmetrized variants of the nonlinear interdependencies S_s and H_s and the event synchronization Q are used, all of them also measuring the level of synchronization. From the latter approaches also three different measures of asymmetry are yielded, namely the nonlinear interdependencies S_a and H_a as well as the delay asymmetry q.

Measures were applied to quasi-continuous multi-day EEG recorded intracranially during the pre-surgical work-up from nine patients with mesial temporal lobe epilepsy (comprising 66 seizures in a recording time of 860 hours. Seizure onset times were determined by expert EEG readers using visual inspection.). The first five of these data sets have been provided from different epilepsy centers for the common data base established for the 1st International Workshop on Seizure Prediction held in April 2002 in Bonn, Germany [85]. These will be denoted as follows: A (Bonn, Germany), B (Florida, USA), C (Amsterdam, Holland), D (Kansas, USA) and E (Pennsylvania, USA). The remaining four data sets, like the first one, have been recorded in the Department of Epileptology, University of Bonn, Germany. They will be denoted as patients F, G, H and I.

For each measure and every patient time profiles are calculated using a moving window technique with non-overlapping segments of 4096 data points each. Depending on the sampling rate for the respective patient, the corresponding duration of these segments range from 17 to 20.5 s (To approximate a uniform segment duration for all patients, data of patient C are downsampled from 480 Hz to 240 Hz). This time frame can be regarded as a good compromise between a high temporal resolution and a statistical accuracy sufficient for the calculation of the different measures. To considerably reduce the amount of data and to be most sensitive to local effects the analysis is performed for neighboring channel combinations only[3]. More detailed information about the data sets and their acquisition can be found in Fig. 5.5 and Tab. 5.1.

[3]For patient C, who had a more distributed implantation scheme without apparent neighborhoods, com-

ID	C	Gender	Foc	L [h]	Sz	Ch	CC	SR [Hz]	Rs [Bit]	FS [Hz]
A	D	M	L	107	10	48	40	200	16	0.3-70
B	USA	M	L	65	15	32	25	200	10	0.1-70
C	HOL	M	L	20	3	32	31	480	12	0.1-100
D	USA	F	L	50	6	51	48	240	10	0.1-100
E	USA	F	L	69	17	81	69	200	12	0.5-70
F	D	F	L	190	1	56	48	200	16	0.3-70
G	D	M	R	112	6	20	18	200	16	0.3-70
H	D	M	L	141	5	48	40	200	16	0.3-70
I	D	F	L	106	3	48	39	200	16	0.3-70
$\sum$				860	66	416	358			

Table 5.1: Patient characteristics and recording parameters. Depicted are patient-ID, country of recording, gender, focal side, length of recordings L, number of seizures Sz as well as channels Ch and combinations of neighboring channels CC. Furthermore, the sampling rate SR, the AD resolution Rs and the bandpass filter settings FS are given.

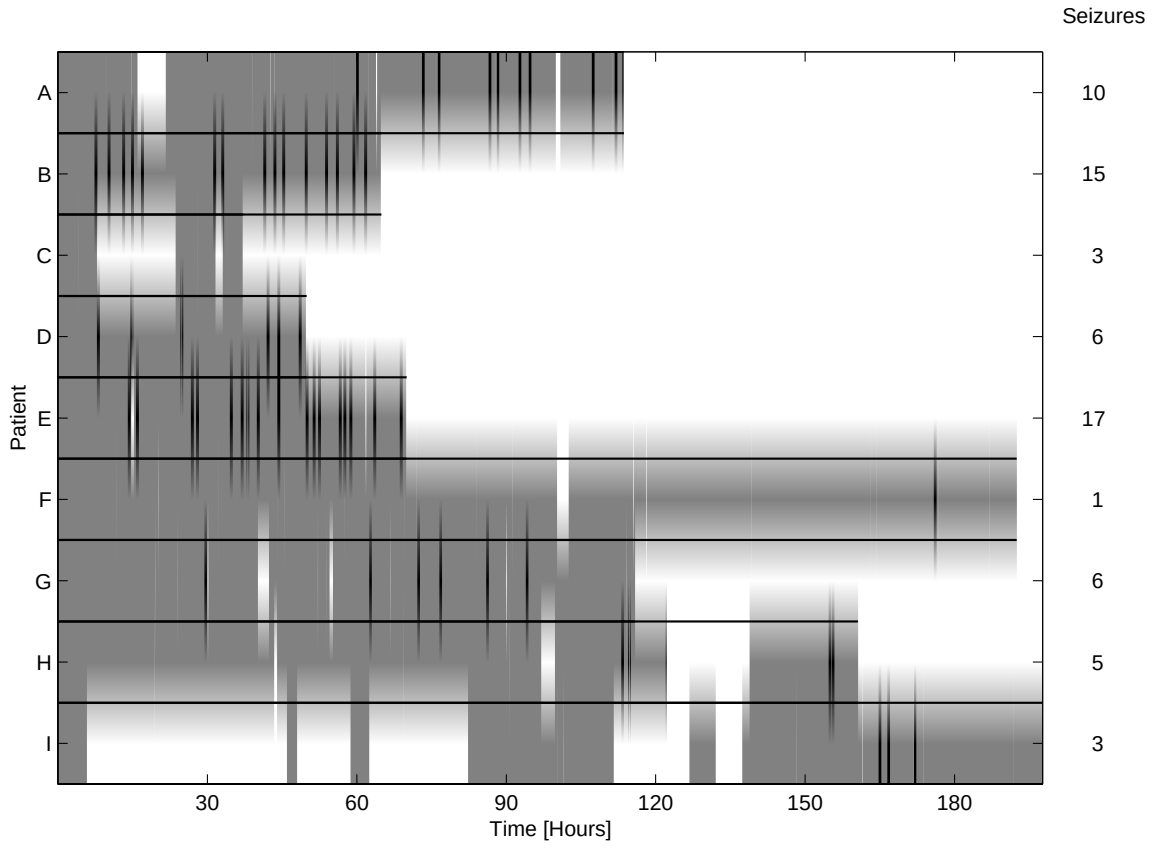


Figure 5.5: Schematic view of data base. Gray horizontal bars denote recording times, white blocks indicate recording gaps, black vertical lines denote seizures.

In addition to the steps of preprocessing already described in the first part of Section 4.1.2 certain parameters have to be adjusted for a meaningful application of the different measures to the EEG (cf. Chapter 3). Only cross correlation and the three indices of phase synchronization based on the Hilbert transform are free of parameters. The phase synchronization measures based on the wavelet transform were calculated with the center frequency set to $\omega_c = 3Hz$ and the number of cycles set to $nc = 3$ (chosen according to Refs. [61, 128]). The calculation of mutual information was performed without embedding (i.e., embedding dimension $m = 1$, no time delay τ needed) and with the number of nearest neighbors set to $k = 1$. The state space reconstruction as a prerequisite for the calculation of the nonlinear interdependence measures was carried out using an embedding dimension $m = 10$ and a time delay $d = 5$. Here a Theiler correction of $T = 50$ data points was used. The number of nearest neighbors was set to $k = 10$ (chosen in accordance with Ref. [128]). Finally, the events as basic ingredients for the calculation of event synchronization were defined as local maxima and minima and the method of fixed time lag was chosen.

5.3.1.2 Preprocessing: Elimination of artifacts

When measuring physiological data like the human electroencephalogram, it is a nontrivial task to guarantee a good quality of the data. Very often a preprocessing becomes necessary since EEG recordings typically are contaminated by various kinds of artifacts either caused by the recording system (e.g., power line interferences due to unsatisfactory shielding or insulation, data clipping due to insufficient input level adjustment) or by the patient himself (e.g., artifacts from eye movements or other muscle activity, or due to pulsative variations of the blood flow). In this study (cf. Ref. [103]), artifacts due to an external reference electrode as well as prominent electrocardiographic signals and movement artifacts found in the recording of patient E were suppressed by transforming the EEG to a common average reference scheme. Finally, to identify recording dropouts and data clippings, every analysis window from every patient was scanned for so-called plateaus: Any analysis window containing either more than 40 consecutive sampling points of identical value or more than 1000 data points in different plateaus was declared as artifact. If for a given channel combination the number of artifact windows thus identified exceeded 5% of all windows, the entire channel combination was discarded from the further analysis. This amounted to 3 discarded channel combinations for patient A, 2 for patient B, and 13 for patient E. For the measure profiles of the other channel combinations values of artifact windows were, according to the respective evaluation scheme, set to the median of the respective amplitude distributions (i.e., inter-ictal or pre-ictal). With this method any bias between inter-ictal and pre-ictal intervals due to artifacts is thoroughly avoided.

binations of successive channels were used.

5.3.1.3 Discriminating the pre-ictal from the inter-ictal interval using ROC-Curves

If a measure is capable to distinguish the pre-ictal interval from the inter-ictal interval, it should attain different amplitude values for these intervals. Therefore the potential predictive performance of different measures is evaluated by retrospectively comparing the amplitude distributions of the respective intervals for each measure (cf. Ref. [103]). There exists numerous different approaches to judge the dissimilarity of two amplitude distributions, e.g., the Kolmogorov-Smirnov test. Another well known approach is called Receiver-Operating-Characteristics (ROC) [44]. Within this statistics, a threshold for amplitude values is continuously shifted across these distributions, and the fraction of amplitude values of the first distribution below this threshold is plotted against the respective fraction of the second distribution. With respect to one of the two complementary hypotheses of separability (values from the pre-ictal distribution are generally lower (higher) than those from the inter-ictal distribution) this corresponds to plotting the sensitivity (ratio of true positives to total number of positives) against 1 minus the specificity (ratio of true negatives to total number of negatives). The capability of a measure to distinguish between the inter-ictal and the pre-ictal interval, i.e., its potential predictive performance, can then be quantified by the area between the resulting ROC-curve and the diagonal. Identical distributions lead to a zero area, while for distributions that are completely non-overlapping, ROC-values of 0.5 or -0.5 are attained, depending on which hypothesis is used for the definition of sensitivity and specificity. To cover the range from $[-1, 1]$ this area is renormalized by a factor of 2. Note that this definition differs from common practice in ROC-statistics where values between 0 and 1 are used. In Fig. 5.6 an illustration of ROC-curves is depicted.

In previous studies the predictive performance of a measure has been evaluated by quantifying the occurrence of distinct patterns, e.g., local drops or peaks parameterized by their width and depth relative to a given reference level (e.g., the mean value over all inter-ictal intervals), in the measures' profile relative to seizure onset [84, 30]. The approach to compare amplitude distributions using ROC-statistics is more robust since it does not depend on the choice of a reference level. Nevertheless, its application involves, as usual, the choice of certain algorithm parameters. If computationally feasible, one common practice in these cases is to try many different combinations of parameters and to choose the most successful one.

In this case parameters are necessary because it is not known beforehand which the prominent features are to be extracted from the measure profiles (e.g., drops or peaks), how long they last, and finally at what times before a seizure they occur. The first point is addressed by testing for both a pre-ictal decrease of amplitude values ($\downarrow$, in this thesis positive ROC-value by definition) as well as an increase ($\uparrow$, here negative ROC-value by definition), thereby judging ROC-values by their absolute value. To account for longer durations of drops or peaks, the analysis is performed not only using the unsmoothed measure profiles, but also after smoothing of the measure profiles with a backward moving average filter of size $d = 5$ min. Using a smoothing filter of size d is equivalent to applying a moving

window technique to the profiles and calculating the area under the profile in a window of length d. This area is smaller (larger) for windows that contain a local drop (peak). Therefore, the parameter d governs the minimum duration of drops and peaks since the smaller the duration of different drops and peaks, the more likely it is that they will cancel each other out. Subsequently, the profiles can be characterized by their level only, and simple thresholding with a continuously varied threshold can be used to compare the different distributions of drops and peaks in the inter-ictal and the pre-ictal interval as it is done using ROC-statistics. Finally, the discriminative test is performed allowing different lengths s of the pre-ictal interval. Since it is not computationally feasible to perform a statistical test for every possible duration of a presumed pre-ictal state, four different lengths are chosen to cover the range of anticipation times for different measures reported in the literature on seizure prediction: $s = 5$ min (cf. Refs. [135, 28]), $s = 30$ min (cf. Refs. [84, 93, 79, 76, 80]), $s = 120$ min, and $s = 240$ min (cf. Refs. [88, 51, 101, 100]). If a pre-ictal state were reflected by a decrease in values of a characterizing measure starting a

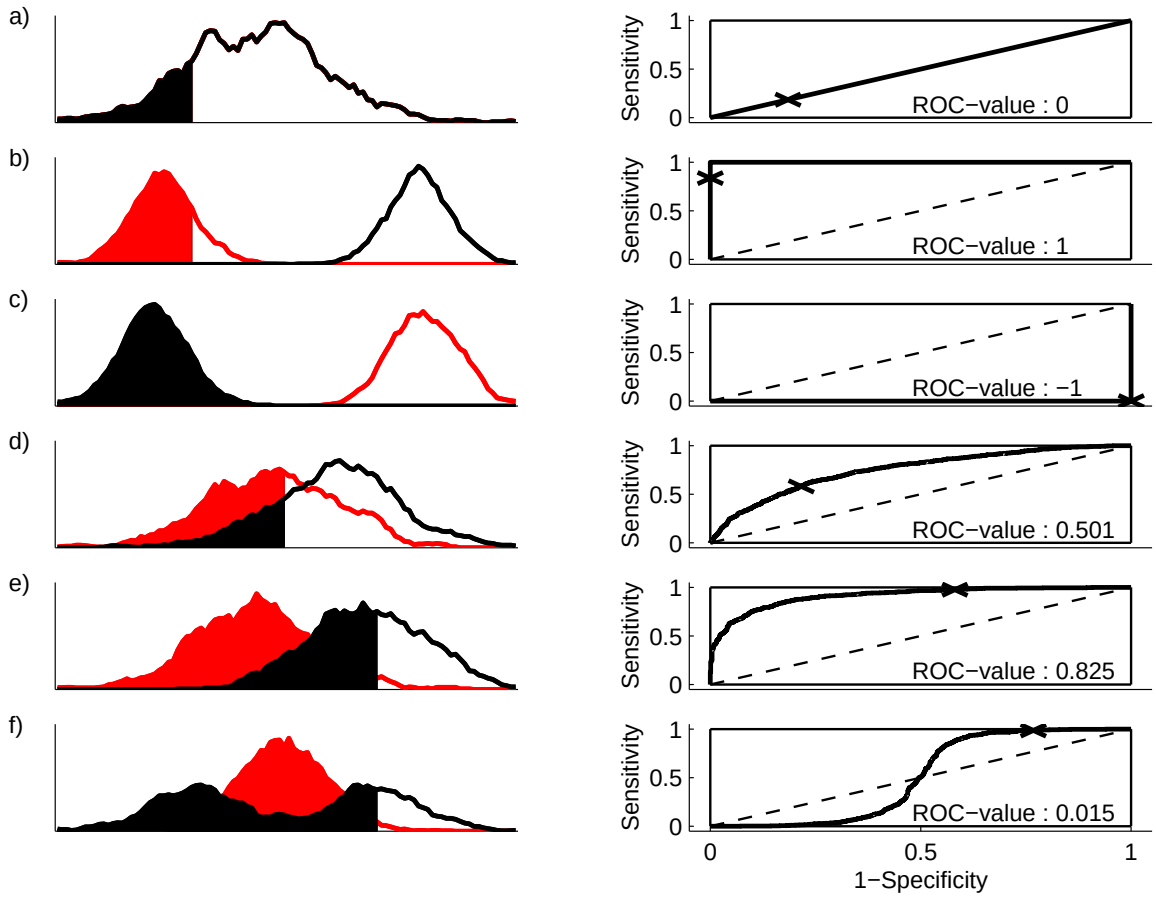


Figure 5.6: Illustration of Receiver-Operating Characteristics using the three special cases yielding neutral, minimum and maximum discrimination (a-c) as well as three further examples of more general distributions (d-f).

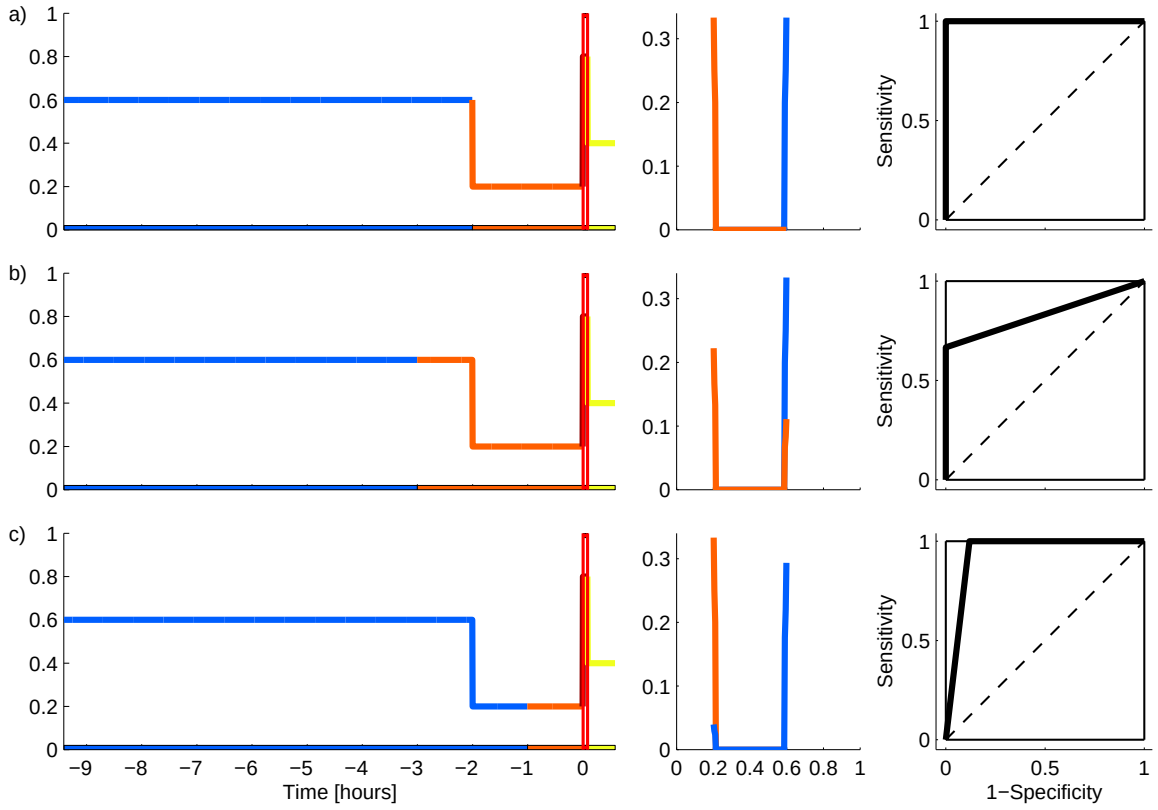


Figure 5.7: Ideal measure profile along with the corresponding amplitude distributions and ROC-curves for three different lengths of the pre-ictal interval: a) 2 hours, b) 3 hours, c) 1 hour. The actual pre-ictal decrease in the values of the measure profiles is assumed to start 2 hours before the seizure in all three cases.

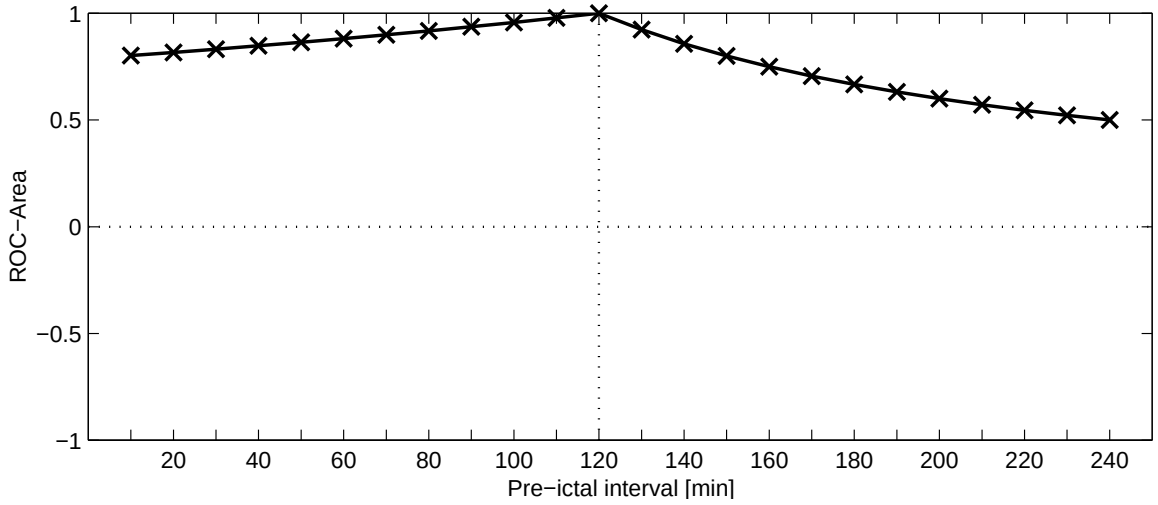


Figure 5.8: Influence of the length of the pre-ictal interval on the discriminative performance.

distinct time before seizure onset, then this effect would be best resolved by defining the pre-ictal interval according to this time. In this case an optimum discrimination between the inter-ictal and the pre-ictal amplitude distributions (cf. Fig. 5.7a) would be yielded. Selecting a pre-ictal interval of a larger (Fig. 5.7b) or smaller (Fig. 5.7c) length would result in an increased number of false negative or false positive classifications, respectively, and thus substantially decrease the discriminative performance (cf. Fig. 5.8). The slope on either side of the maximum performance is dependent on the ratio of the lengths of inter- and pre-ictal intervals.

Since the ictal and the post-ictal intervals are known to be associated with massive changes in the EEG, they are not included in either of the two amplitude distributions. More precisely, recording intervals lasting from seizure onset till 30 min after seizure termination are discarded from the analysis. If the time between two successive seizures is less than s + 30 min, the maximum amount of data available (i.e., from the end of the preceding seizures' postictal phase till seizure onset) is used instead.

5.3.1.4 Evaluation schemes

Four different evaluation schemes are applied (cf. Ref. [103]):

1. All inter-ictal $\Leftrightarrow$ All pre-ictal

   In the first and simplest evaluation scheme the distribution of all pre-ictal values from all combinations of neighboring channels and all seizures of a patient is compared to the respective distribution of all inter-ictal values. For this approach a high discrimination is attained only in case of a constant global effect, i.e., if either a pre-ictal increase or a decrease is encountered uniformly and on a similar level in all channel combinations and for all seizures.

2. Inter-ictal per channel combination $\Leftrightarrow$ Pre-ictal per channel combination

   For an actual seizure prediction a global effect would not be needed, a local effect would be sufficient as long as its spatial location is constant over successive seizures. And in fact, in most studies seizure precursors have been reported to occur only in certain distinct channels or channel combinations. Therefore in the second evaluation scheme the discriminative power of the different measures is evaluated for each channel combination separately. To account for a possible local effect the channel combination yielding the best performance is chosen. Therefore in this scheme a good predictive performance is obtained only if there exists a channel combination for which seizure precursors occur constantly and on a similar level for all seizures of a patient.

   For this evaluation scheme the discriminative test is illustrated by depicting an exemplary time profile (channel combination TR08-TR09 of patient A analyzed by the measure for phase synchronization γ_{cv}^{H}) with its two distributions and the resulting ROC-curve in Fig. 5.9.

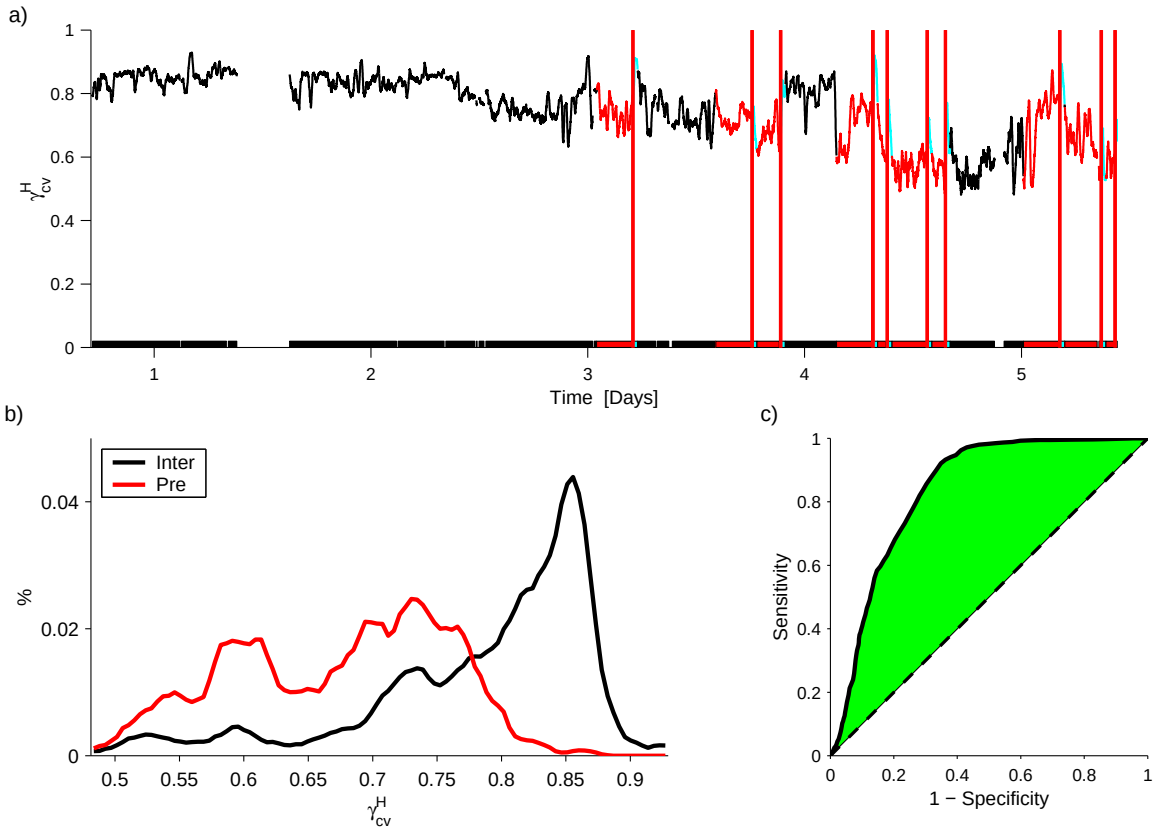


Figure 5.9: Illustration of the statistical evaluation based on ROC-curves: (a) Original measure profile of the best channel combination (TR08-TR09) of patient A for the index based on circular variance using the Hilbert Transform γ^H_{cv} obtained by choosing a 5 min smoothing filter and a pre-ictal interval of 240 min. Seizures are marked by vertical lines, pre-ictal and inter-ictal intervals are depicted in bright and dark color, respectively. (b) Distributions of values from all inter-ictal and all pre-ictal intervals. (c) Corresponding ROC-curve yielding the maximum performance value of 0.68.

3. Inter-ictal per channel combination $\Leftrightarrow$ Pre-ictal per channel combination and seizure

   In the third evaluation scheme for each channel combination the distribution of all inter-ictal values is successively compared to the distribution of the pre-ictal values of every seizure of a patient. Thus a performance value quantifying the discriminative power of a measure is obtained for every channel combination and each seizure. The distribution over the different seizures can show whether a measure might be capable of anticipating not all, but at least certain seizures of a patient. Still this would constitute a remarkable and useful achievement.

4. Inter-ictal per channel combination and seizure $\Leftrightarrow$ Pre-ictal per channel combination and seizure

   Also in the fourth evaluation scheme for each channel combination the discriminative

test is performed for the pre-ictal values of every seizure separately. But here this distribution is not compared to all inter-ictal values, but rather only to the inter-ictal values preceding the pre-ictal interval of the respective seizure. While the first test corresponds to a non-adaptive, constant reference level, this final scheme is adaptive in the sense that it accounts for possible slow changes in the dynamics resulting in slow baseline shifts (e.g., due to physiological variations).

The evaluation of this scheme becomes impossible in cases where the interval between two successive seizures is shorter than 35 min + s (30 min for the post-ictal interval of the first seizure, 5 min as the minimum length for the inter-ictal distribution, and with s as the pre-ictal interval of the second seizure). Whenever such a clustering of seizures occurs [45], the following relation holds: the longer the pre-ictal interval s, the more seizures have to be discarded from the analysis. While only one seizure has to be discarded for s = 5 min, this number amounts to 5 for s = 30 min, 24 for s = 120 min, and 35 for s = 240 min.

Another evaluation scheme (best channel combination per seizure) has been proposed in Ref. [51]. However, the practicability of this approach for a potential prospective implementation appears to be rather limited.

In each of the four evaluation schemes for every measure the respective performance value is averaged over patients. This average performance is used as the most important benchmark in the comparison of the different measures. But it also serves as a criterion for the optimum combination of parameters (smoothing, length of the pre-ictal interval, ROC-hypothesis of separability) for each measure. This way the comparison of the different measures is performed without the use of any a priori knowledge and all measures are treated equally since ”each one can choose its own optimum parameter combination”. In the first evaluation scheme only one over-all performance value is compared, from the second evaluation scheme onwards the performance value of the best channel combination is used for comparison. In the third and fourth evaluation scheme the criterion for the selection of the best channel combination is the median over all seizures of a patient.

5.3.1.5 Seizure times surrogates as a test for statistical validity

To assess the performances yielded by the different measures and to judge their statistical validity by assigning a level of significance (cf. Section 5.2), an ensemble of 19 seizure times surrogates is generated. In the original method of Ref. [9] each seizure times surrogate was obtained by a different random permutation of the intervals between subsequent seizures. The interval between the beginning of the recording and the first seizure was not permuted, this seizure was displaced using a random jitter. In the application of Ref. [103] this interval was included in the permutation scheme. In that study no level of significance was assigned to patient C, since the number of seizures and hence the number of possible permutations is not large enough to allow the generation of the 19 surrogates

needed to obtain the desired significance. In the extended group of patients analyzed in this study this problem would cover patients F and I as well. In order to include all patients in the estimation of significance, here a different randomization scheme is applied. The intervals between seizures are no longer permuted but rather the seizures themselves are placed on new positions using a constrained randomization. Clusters of seizures (defined as consecutive seizures for which the interval in between is less than the maximum length of the pre-ictal interval, e.g., 240 min) are displaced as a whole. Furthermore, the first four hours at the beginning of the recording are prohibited in order to guarantee a complete pre-ictal interval for the very first seizure. Subsequently, for each evaluation scheme and every measure the seizure prediction algorithm is applied to the respective time profiles using the different seizure time surrogates instead of the original seizure onset times. In each of the different evaluation schemes the null hypothesis can be rejected with a significance level of $p = 0.05$, if highest performance values are yielded for the original seizure times.

5.3.1.6 Correlations between the different measures

To investigate to which extent the different measures of synchronization and directionality carry independent and non-redundant information, their correlation is estimated on the entire database analyzed in this Section. For this aim, correlation coefficients are determined from all 358 channel combinations and all 153548 windows from all nine patients.

5.3.2 Results

First a short overview and some general guidelines are given. Subsequently, the results of the four different evaluation schemes are presented.

These results can be seen as projections of a high-dimensional space since there are many different degrees of freedom involved (in detail: evaluation schemes, measures, patients, channel combinations, seizures, order of smoothing filter, length of the pre-ictal interval, ROC-hypotheses of separability, seizure time sequences). These degrees of freedom are related as follows: In four different evaluation schemes the predictive performance of fourteen bivariate measures is evaluated by analyzing nine patients. Each of these patients was implanted with an individual number of electrodes resulting in an individual number of channel combinations. Every patient had a certain number of seizures during the recording. In each evaluation scheme the respective pre-ictal intervals are distinguished from the respective inter-ictal intervals using ROC-statistics. Every analysis is carried out using two different smoothing filters and four different values for the length of the pre-ictal interval. Furthermore, both a pre-ictal increase as well as a pre-ictal decrease are considered. Finally, every discriminative test is performed 20 times, one time for the original sequence of seizures and 19 times for the different seizure time surrogates.

In the following for each evaluation scheme the mean of the respective performance value over patients is displayed to allow the comparison of the different measures. Furthermore, each and every time the results yielded for the single patients are depicted using for each measure the combination of parameters that lead to the respective optimized performance value. This combination is specified as well.

5.3.2.1 First evaluation scheme: All inter-ictal $\Leftrightarrow$ All pre-ictal

In the first evaluation scheme the distribution of all pre-ictal values from all neighboring channel combinations and all seizures of a patient is tested against the respective distribution of all inter-ictal values. For each measure the maximum out of 16 performance values (two different smoothing filters, four different lengths of the pre-ictal interval, pre-ictal increase / decrease) is chosen. Results of this evaluation scheme are depicted in Fig. 5.10.

The performance values obtained by averaging over patients range from 0.008 for the index based on circular variance using the Hilbert Transform γ_{cv}^{H} to 0.064 for event synchronization. The performances of the anti-symmetric measures are among those of the symmetric measures. All of these values are non-negative due to the selection of the hypothesis of separability, pre-ictal increase ($\uparrow$) or pre-ictal decrease ($\downarrow$), with the highest absolute ROC-value. Nevertheless, no measure is able to clearly discriminate the pre-ictal from the inter-ictal interval in this evaluation scheme since all performance values are quite low and none of them proves to be significant when tested by the method of seizure times surrogates.

For the measures of synchronization as well as for the measures of directionality a pre-ictal increase is observed about twice as often as a decrease. The corresponding parameters d and s that yield the maximum performance show non-uniform values, too.

These results show that no measure is able to show a distinct separation between the distribution of all pre-ictal values from all combinations of neighboring channels and all seizures and the respective distribution of all inter-ictal values. There is not a single combination of parameters for which a significant global effect (i.e., an over-all increase or an over-all decrease of pre-ictal values) can be observed. Since there are two steps of averaging involved in this evaluation scheme (over channel combination and over seizures) all features that might be predictive of seizures either do not occur globally or do not occur constantly (or neither of both). If there are local effects (in space or in time) these are either too weak to leave any remarkable impact in the average or get cancelled out by pronounced opposite effects. The possibility of predictive features occurring only in some channel combinations is tested in the next evaluation scheme.

5.3.2.2 Second evaluation scheme: Inter-ictal per channel combination $\Leftrightarrow$ Pre-ictal per channel combination

In the second evaluation scheme a new degree of freedom is introduced, since the discriminative power of the different measures is evaluated for each neighboring channel combina-

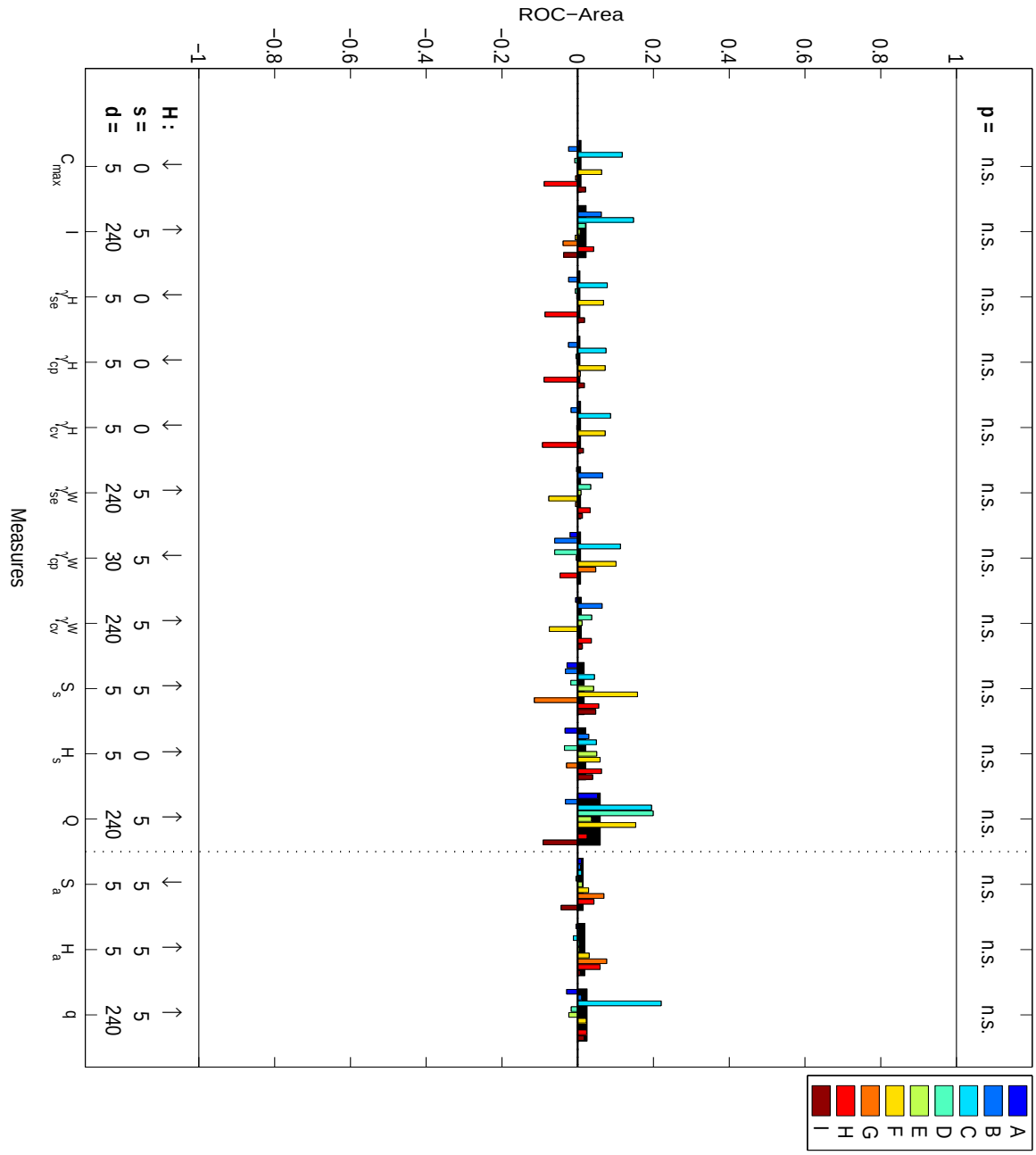


Figure 5.10: Comparison of measures for the first evaluation scheme (All inter -ictal $\Leftrightarrow$ All pre-ictal). For each measure the average performance value is depicted by a wide black bar in the background. In front of these the values obtained for the single patients are displayed using the combination of parameters d and s and the respective ROC-hypotheses ($\downarrow$ for a pre-ictal decrease, $\uparrow$ for an increase) that yielded the maximum performance. These values are stated at the bottom of the figure. At the top the levels of significance of the performance values are shown. Whenever the p-values is larger than 0.05, the respective result is marked as non-significant (n.s.). The dashed line separates the measures of synchronization from the measures of directionality.

tion separately. Therefore, for each patient and measure the single ROC-value used in the first evaluation scheme is now replaced by a distribution of ROC-values. An example how to obtain the ROC-value from a measure profile has already been shown in Fig. 5.9. In Fig. 5.11 for the same parameter combination, the same measure and the same patient the pre-ictal and inter-ictal distributions from all neighboring channel combinations of the two depth electrodes of the left and the right hemisphere are shown exemplarily. For most channel combinations the pre-ictal and inter-ictal amplitude distributions turn out to be almost indistinguishable, only for some channel combinations a high degree of discrimination can be obtained.

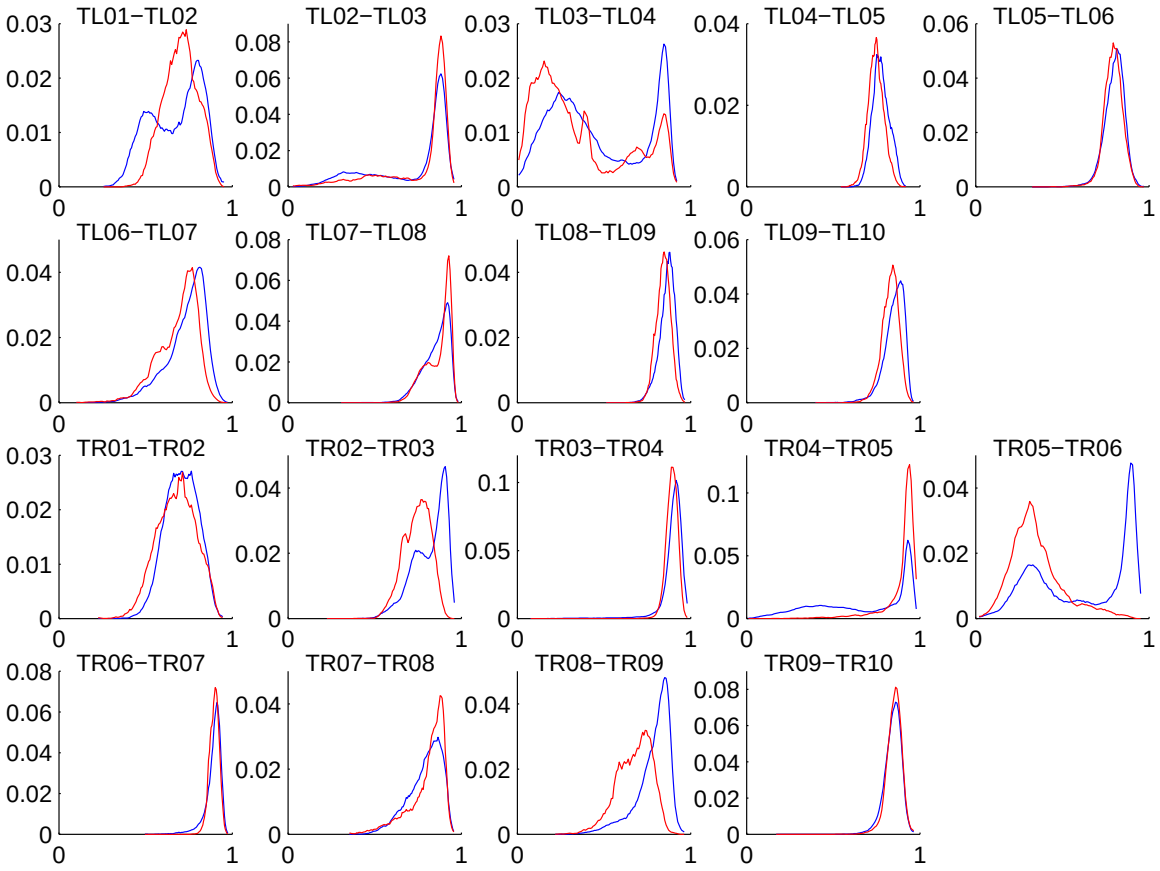


Figure 5.11: Exemplary inter-ictal (blue) and pre-ictal (red) distributions (smoothed) of γ_{cv}^{H} for each channel combination on the left and right hemisphere of patient A (second evaluation scheme). A 5 min smoothing filter and a pre-ictal interval of 240 min is used.

For each measure the ROC-value of the channel combination with the highest discrimination (in this example TR08-TR09) is averaged over patients. The average performance value is obtained by choosing the parameter combination that yields the highest mean value. These values are depicted in Fig. 5.12 along with the performance values obtained for the single patients' channel combinations with maximum, minimum, and median discrimination. The performance values for the second evaluation scheme range from 0.51

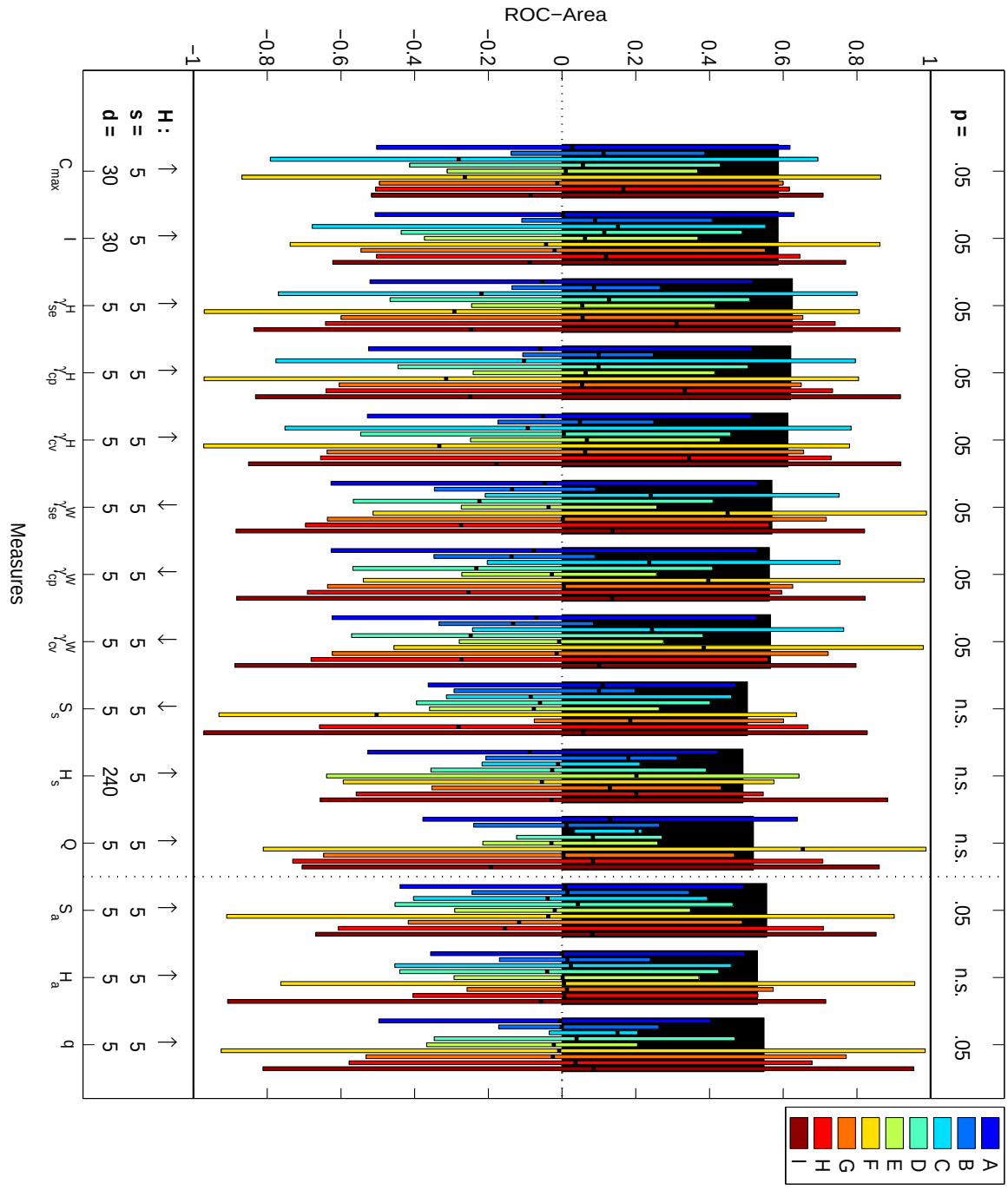


Figure 5.12: Same as Fig. 5.10, but this time for the second evaluation scheme (Inter-ictal per channel $\Leftrightarrow$ Pre-ictal per channel). For each measure and patient, the maximum and the minimum of the distribution of performance values over different channel combinations are depicted by the range of vertical bars. The little black square on each bar denotes the median of the respective distribution. For each measure the over-all performance value (represented by a black bar in the background) is obtained by averaging over the performances of the best channel combination for each patient.

for the nonlinear interdependence H_s to 0.64 for the index based on conditional probability using the Hilbert Transform γ^H_{cp}. Again the anti-symmetric measures rank between the symmetric measures. In this scheme for all measures maximum performances are obtained using a 5 min smoothing. Results for the length of the pre-ictal interval and the ROC-hypothesis of separability are still not consistent. For 10 out of 14 measures performance values reach statistical significance.

Whereas it has been impossible to find a global effect in the first evaluation scheme, in this scheme for most measures and patients distinct local effects can be observed in certain channel combinations (cf. the example given in Fig. 5.11). Due to the new degree of freedom performance values are much higher than in the first evaluation scheme and, furthermore, mostly prove to be significant. Since all pre-ictal phases are included in the discriminative test, these effects either occur constantly and on a similar level for all seizures of a patient or alternatively effects before certain seizures are so pronounced that they surpass opposite effects preceding other seizures. These results appear quite promising, nevertheless there are some effects raising doubts about their actual usefulness. When comparing the performance of different patients it is remarkable that the most discriminative channel combinations can be found for patients with one (patient F) or only a few seizures (patients E and I, cf. Tab. 5.1). This effect appears to be due to the patients' different statistical fluctuations (i.e., different number of seizures and different number of channel combinations). Statistically, it is more likely to find a channel combination with a (seemingly) good discrimination for a patient with only one seizure than for a patient with many seizures. In the latter case there is a high chance that different or even opposite effects in the pre-ictal intervals of different seizures cancel each other out. For the same reason it is not surprising that for most measures the maximum discrimination is yielded for a pre-ictal length of 5 min. The shorter the over-all pre-ictal length (i.e., the size of the first distribution of values), the higher the probability to find a good discrimination between the two distributions just by chance. The fact that the ROC-hypothesis of separability for which maximum performances are obtained differs among measures casts further doubt on the usefulness of these results.

5.3.2.3 Third evaluation scheme: Inter-ictal per channel combination $\Leftrightarrow$ Pre-ictal per channel combination and seizure

In the third evaluation scheme for each channel combination the distribution from the pre-ictal intervals of the single seizures are tested separately for their overlap with the distribution from the inter-ictal interval. The discriminative test is performed for each channel combination and every single seizure of a patient and thus for each measure and each patient a two-dimensional distribution of performance values is yielded.

The results of this evaluation scheme for the combination of parameters (smoothing filter: 5 min, pre-ictal interval:240 min) used in Figs. 5.9 and 5.11 are exemplarily depicted in

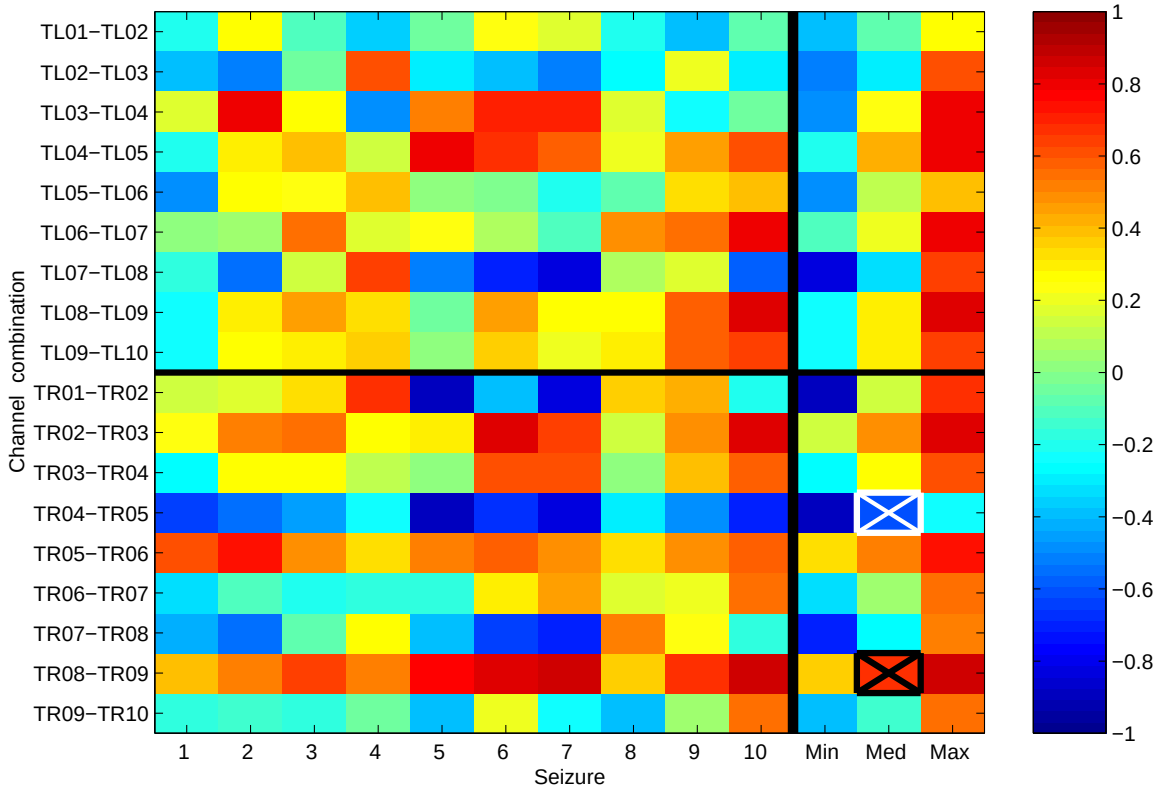


Figure 5.13: Color-coded ROC-values of the third evaluation scheme obtained by applying γ_{cv}^{H} to patient A (smoothing filter $s = 5min$, length of the pre-ictal interval $d = 240min$. In the first ten columns for each channel combination the ROC-values for the ten seizures of this patient are shown. The last three columns depict minimum, median, and maximum of these values. The two channel combinations with the highest median performance for the hypothesis of a pre-ictal decrease (increase) are marked in black (white). In this example the performance value is equal to the median decrease of channel combination TR08-TR09 which is more pronounced than the median increase in channel combination TR04-TR05.

Fig. 5.13, again for patient A and γ_{cv}^{H}. Here the median values are used as a criterion for the best channel combination, i.e., the channel combination is selected for which the pre-ictal intervals of half of the seizures can be distinguished best from the respective inter-ictal intervals.

The performances of the different measures are depicted in Fig. 5.14 along with the single patients' distributions of the median performances in all channel combinations (i.e., the distribution color-coded in the last but one column in Fig. 5.13). Again maximum, minimum, and median channel combination are depicted (the former two are the ones marked by black and white crosses in the aforementioned column of that figure).

The average performance values range from 0.62 for the nonlinear interdependence H_s to 0.76 for the mutual information I and are thus substantially higher than the performance

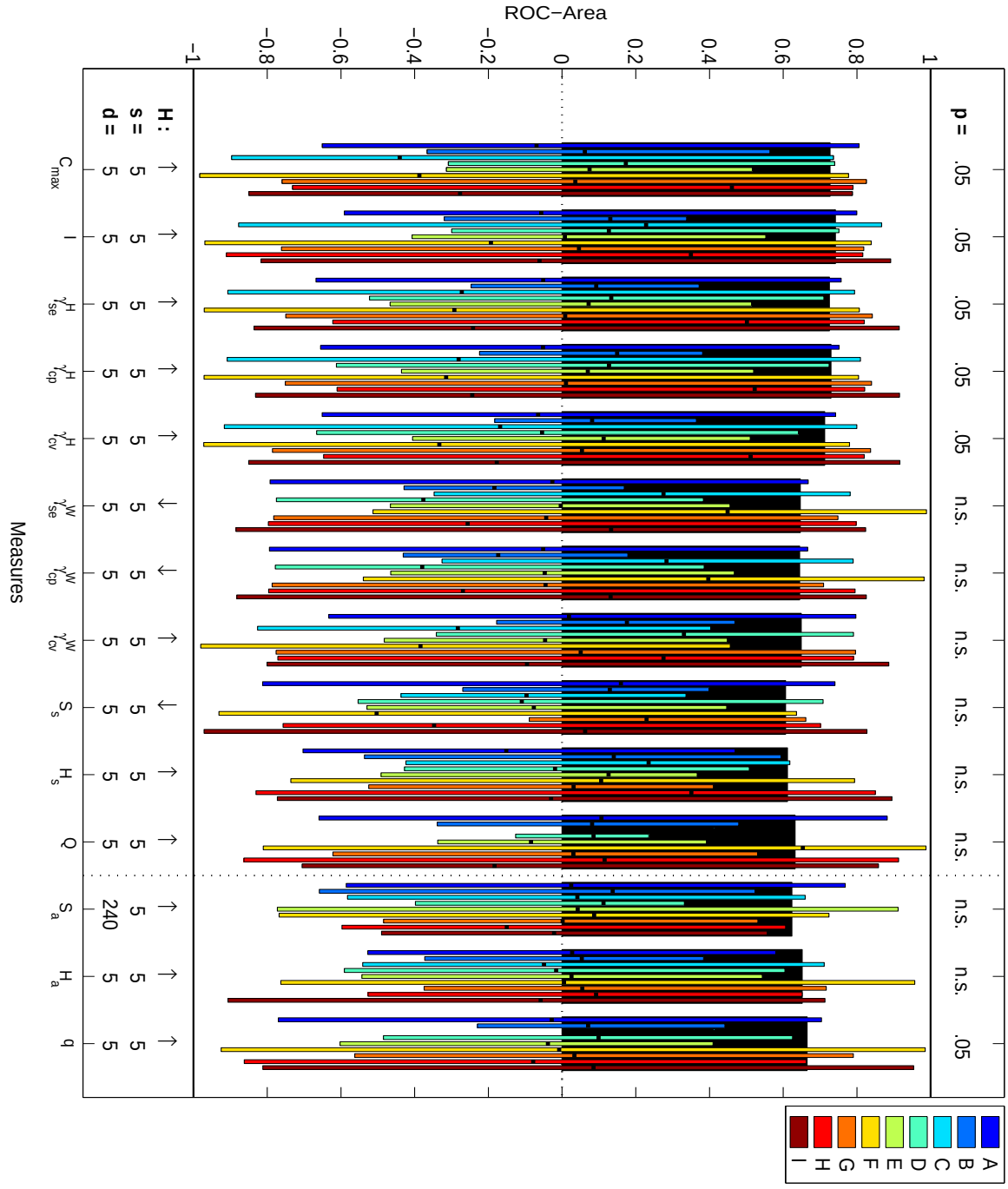


Figure 5.14: Same as Fig. 5.12, but this time for the third evaluation scheme (Inter-ictal per channel combination $\Leftrightarrow$ Pre-ictal per channel combination and seizure). For each measure the over-all performance (represented by wide black bars in the background) is obtained by averaging over the performances of the best channel combinations from each patient. Again for each patient the distribution over the performances of different channel combinations is shown, but this time the performance of a channel combination is defined as the median performance of the different seizures of the respective patient.

values obtained in the second evaluation scheme. Also in this scheme the ROC-values obtained for the anti-symmetric measures rank in between. All measures, except for the nonlinear interdependence S_s, show highest performances when the hypothesis of a pre-ictal increase is used. While for all measures maximum performances are yielded when a 5 min smoothing filter is applied, the optimal length of the pre-ictal interval varies between 5 min and 240 min. Only cross correlation, mutual information, and the Hilbert indices of phase synchronization as well as the delay asymmetry show significant performances.

For the same evaluation scheme in Fig. 5.15 a different projection of the two-dimensional performance values (cf. Fig. 5.13) is used to show the actual discriminative performance of the best channel combination regarding the distribution over all seizures. Using the optimum parameter combination for each measure and the best channel combination (with respect to the median seizure performance, cf. Fig. 5.14) for every patient, the respective distribution of performance values for all seizures is represented by its maximum, minimum, and median value (there is no such distribution for patient F with only one seizure, cf. Tab. 5.1). The median values of these distributions correspond to the maximum values of the distributions displayed in Fig. 5.14. While for all measures and patients the maximum and also the median of the distribution is positive, for many patients (in particular for those with a high number of seizures) the minimum is negative. Thus, while for the majority of seizures a certain discrimination of the inter-ictal and the pre-ictal intervals can be achieved based on either a pre-ictal decrease or increase in values, there are also seizures of some patients for which the opposite effect is observed.

The reason for the higher performances (when compared to the second evaluation scheme) appears to be similar to the one mentioned above in the discussion of the second evaluation scheme. In the third scheme the different seizures of a patient are introduced as a further degree of freedom (although the median performance and not the performance obtained for the seizure with maximum discrimination is chosen as a criterion). Thus for every channel combination the distribution of all pre-ictal values is split up into several pre-ictal distributions, one for each seizure of a patient. These distributions are now compared against the same inter-ictal distributions as before. Therefore, it is likely to yield performances equal to or better than in the previous scheme since there effects before the single seizures might cancel each other out. Results can become better due to the optimization over more performance values with higher statistical fluctuations introduced by the new degree of freedom (cf. Fig. 5.13). Only if there was a constant effect on a similar level over all seizures, results would stay the same (of course they are the same for patient F with only one seizure). This view is further supported by the observation that in this evaluation scheme the performance over different patients is more uniform than before, the observed dependency on the different number of seizures has disappeared. Also the combination of parameters for which maximum performances are yielded is much more consistent among the different measures. All measures, except for the nonlinear interdependence S_s, show a pre-ictal increase, either in synchronization or in asymmetry. In contrast to this uniformity, the length of the pre-ictal interval still shows conflicting results. Only the measures of phase synchro-

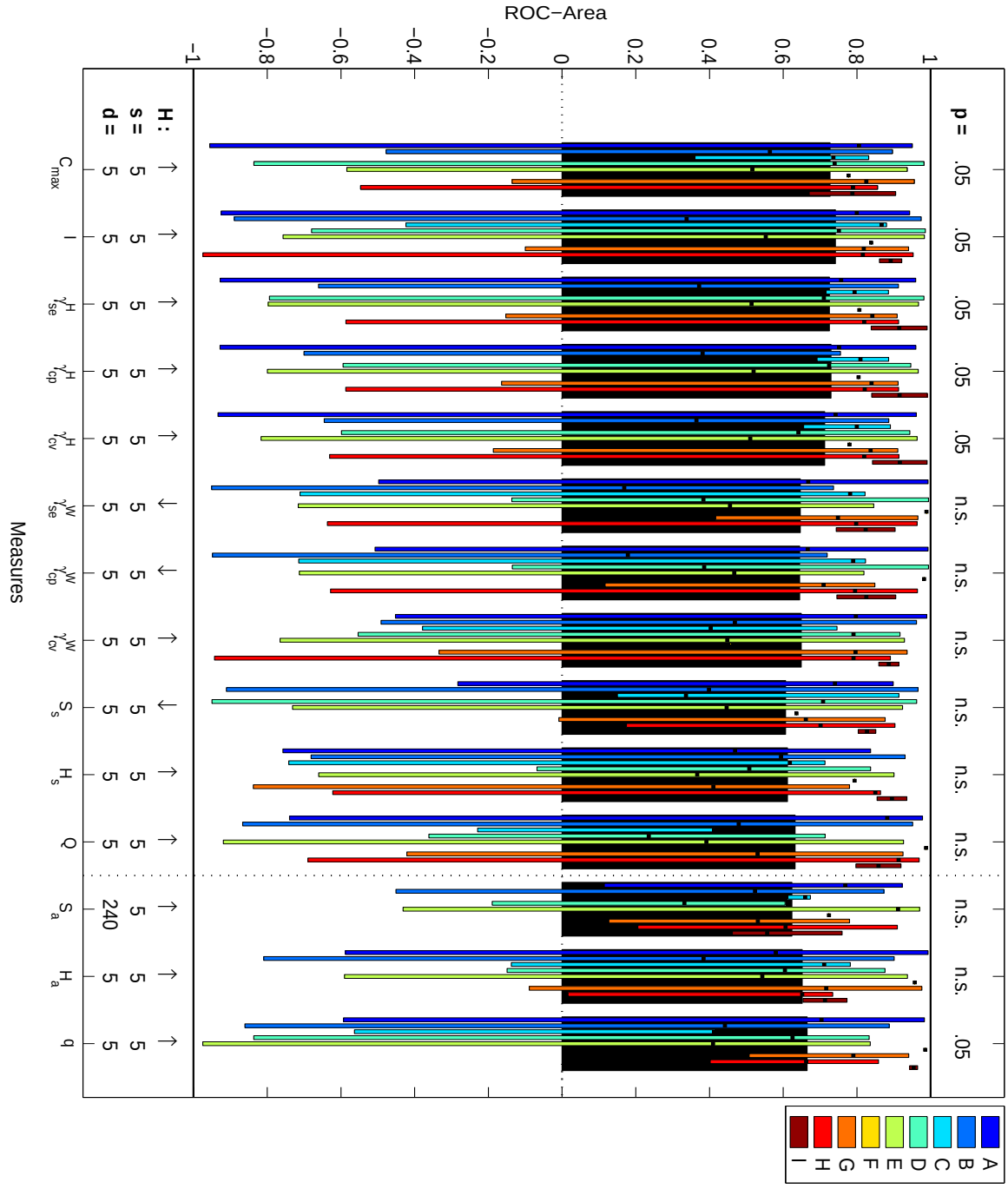


Figure 5.15: Third evaluation scheme: Minimum, maximum, and median performance values for different seizures using the best channel combination (with respect to the performance of the median seizure) of each patient and the best parameter combination for each measure (cf. Fig. 5.14). The median values of the distributions (denoted by the little black squares inside each bar) correspond to the maximum values of the distributions displayed in Fig. 5.14. Thus also the mean values over patients (again represented by the black bars in the background) are the same.

nization based on the wavelet transform have not only switched their ROC-hypothesis but also shifted to the long pre-ictal length. Looking at the distributions over different seizures, again a distinct dependence on the number of seizures can be observed. Whereas for patients with one or just a few seizures very often all pre-ictal intervals can be discriminated with one hypothesis, this is rarely possible for the other patients. For these mostly there exist some seizures for which the opposite ROC-hypothesis is favored. Furthermore, since many of the seizures investigated in this study occur in clusters (cf. Fig. 5.5), the question arises whether there is any outstanding importance of the leading seizure. In the comparison of the first seizure with the succeeding seizures of a patient no such difference can be found. Finally, neither in this nor in the other evaluation schemes any correlation between the location of the channel combinations showing the highest discriminative power based on either ROC-hypothesis and the focal region of each patient can be seen. Maximum performances are obtained about as often in the focal as in the non-focal hemisphere.

5.3.2.4 Fourth evaluation scheme: Inter-ictal per channel and seizure $\Leftrightarrow$ Pre-ictal per channel and seizure

Also in the fourth and last evaluation scheme the discriminative test is performed for the pre-ictal intervals of each seizure separately. But this time for each channel combination the distributions of these intervals are compared to the distributions of the preceding inter-ictal intervals. Results for this evaluation scheme are displayed in Figs. 5.16 and 5.17.

As can be seen from Fig. 5.16 average performance values for the measures of synchronization are again higher than those of the second evaluation scheme ranging from 0.59 for the nonlinear interdependence S_s to 0.77 for the index based on Shannon entropy using the Hilbert Transform γ_{se}^{H}. In comparison to the third evaluation scheme performance values are higher for some measures but lower for other measures. Once more no fundamental difference between the symmetric and the anti-symmetric measures can be observed. For all measures a 5 min smoothing filter leads to highest performance values, and for all measures, except for the event synchronization Q, maximum performances are obtained for a pre-ictal length of 5 min. The ROC-hypothesis of separability varies from measure to measure. Only the symmetric and the anti-symmetric variant of event synchronization reach statistically significant performance values.

In Fig. 5.17 again for each measures' best parameter combination and each patients' best channel combination the distributions of different seizures is depicted. Also in the fourth evaluation scheme none of the two hypotheses of either a pre-ictal decrease or an increase in values is valid for all seizures of a patient.

The comparison of the third and the fourth evaluation scheme shows that quite the same results are yielded no matter whether the pre-ictal distributions of most seizures are discriminated from the preceding intervals only or from the whole inter-ictal interval. Thus it appears as if accounting for the variability of the inter-ictal intervals by the choice of an

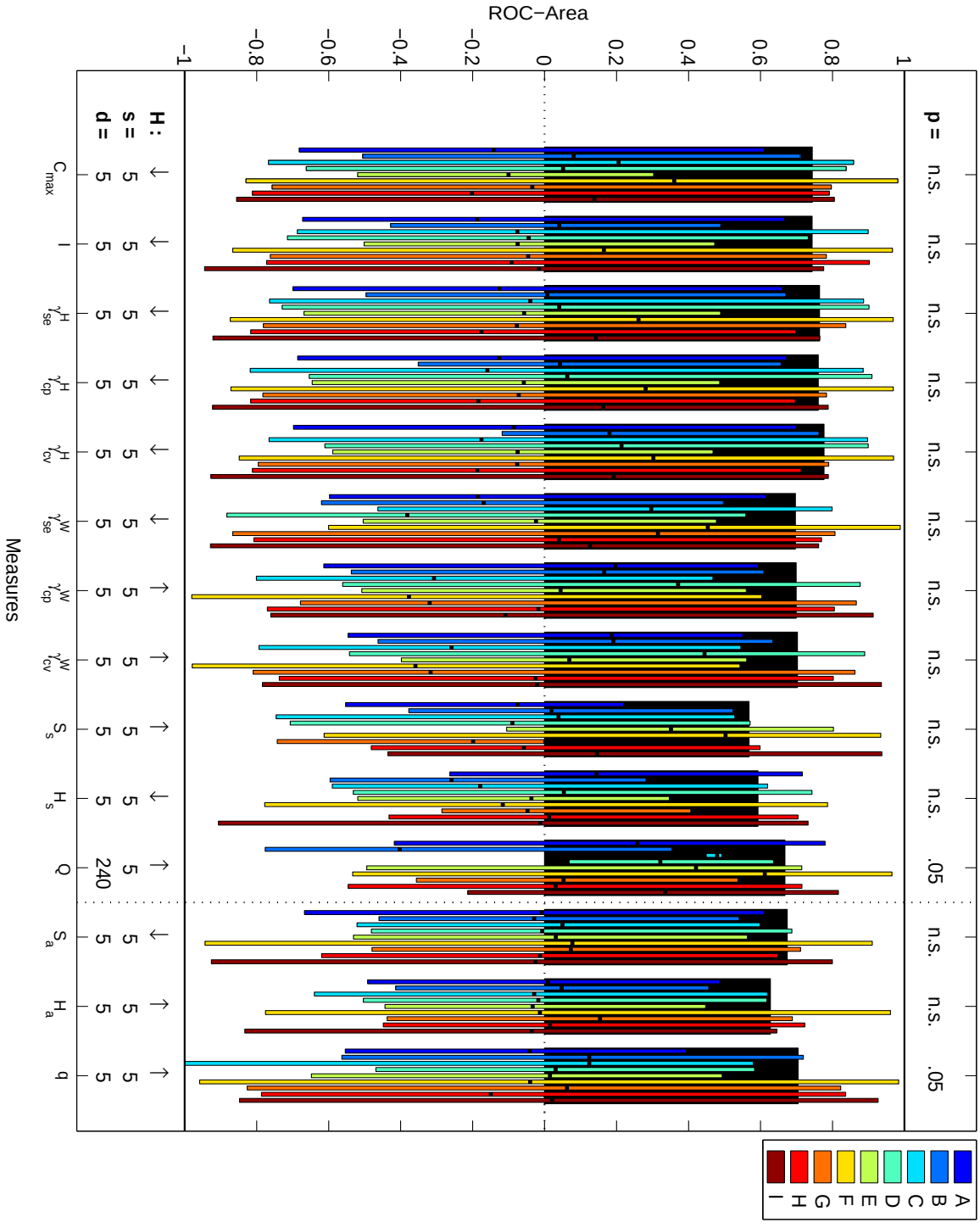


Figure 5.16: Same as Fig. 5.14, but this time for the fourth evaluation scheme (Inter-ictal per channel and seizure $\Leftrightarrow$ Pre-ictal per channel and seizure). Here the pre-ictal interval of each seizure is tested only against the preceding inter-ictal interval.

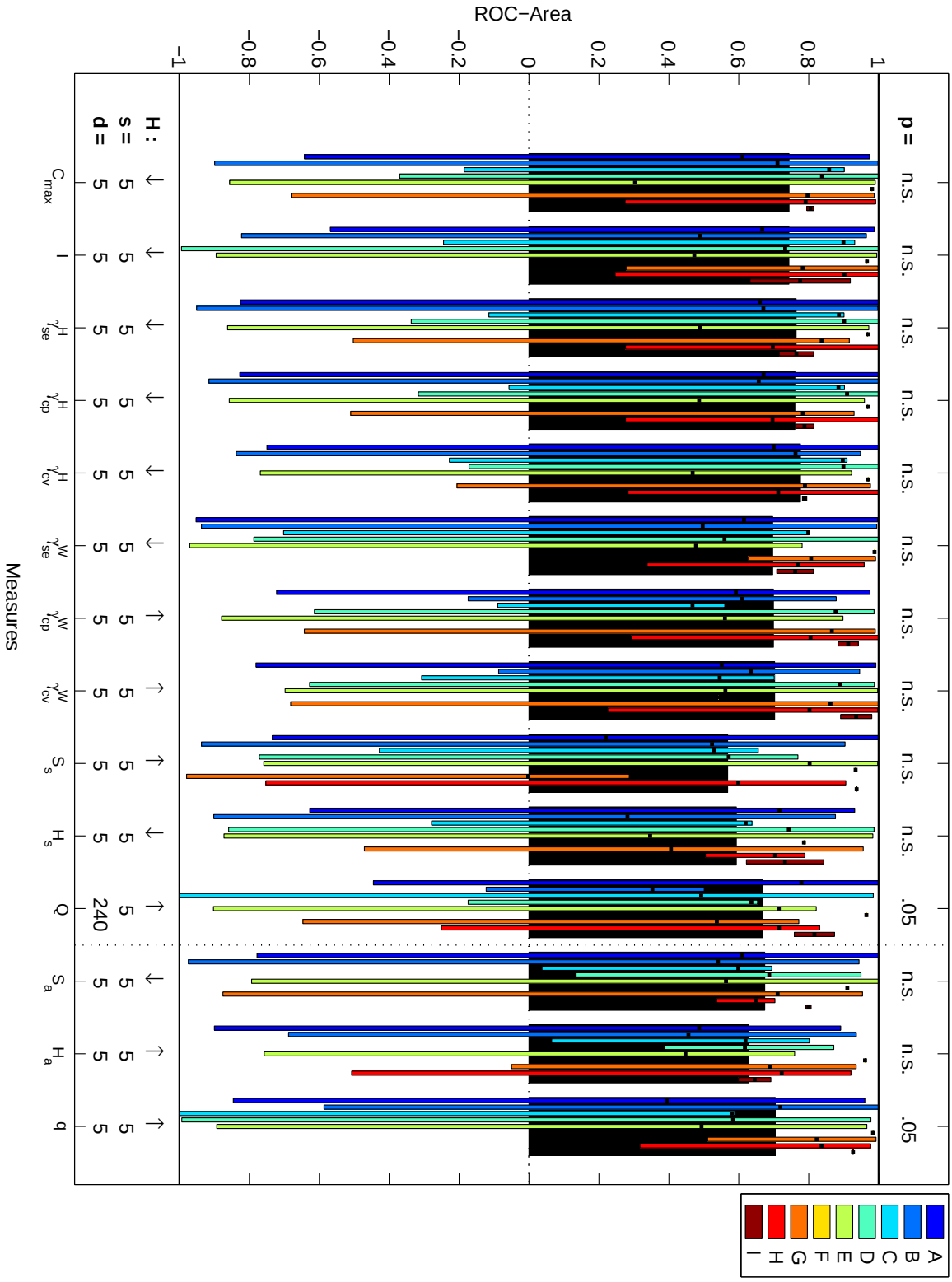


Figure 5.17: Same as Fig. 5.15, but this time for the fourth evaluation scheme.

adaptive baseline does not lead to a considerable improvement. In this evaluation scheme for all measures except for the event synchronization Q maximum performances are obtained for a pre-ictal length of 5 min. In this case the sizes of the two distributions are the smallest of all evaluation schemes and the high statistical fluctuations associated with this might be responsible for the preference of this pre-ictal length. For one half of the measures a pre-ictal increase, for the other half a pre-ictal decrease is observed. Summing up the results of the third and fourth evaluation scheme it seems that in some channel combinations there are distinct effects before seizures, but only rarely these effects occur constantly over all seizures.

5.3.2.5 Correlations between the different measures

In Fig. 5.18 the pairwise correlation coefficients between all fourteen measures are depicted based on the entire database analyzed in this Section.

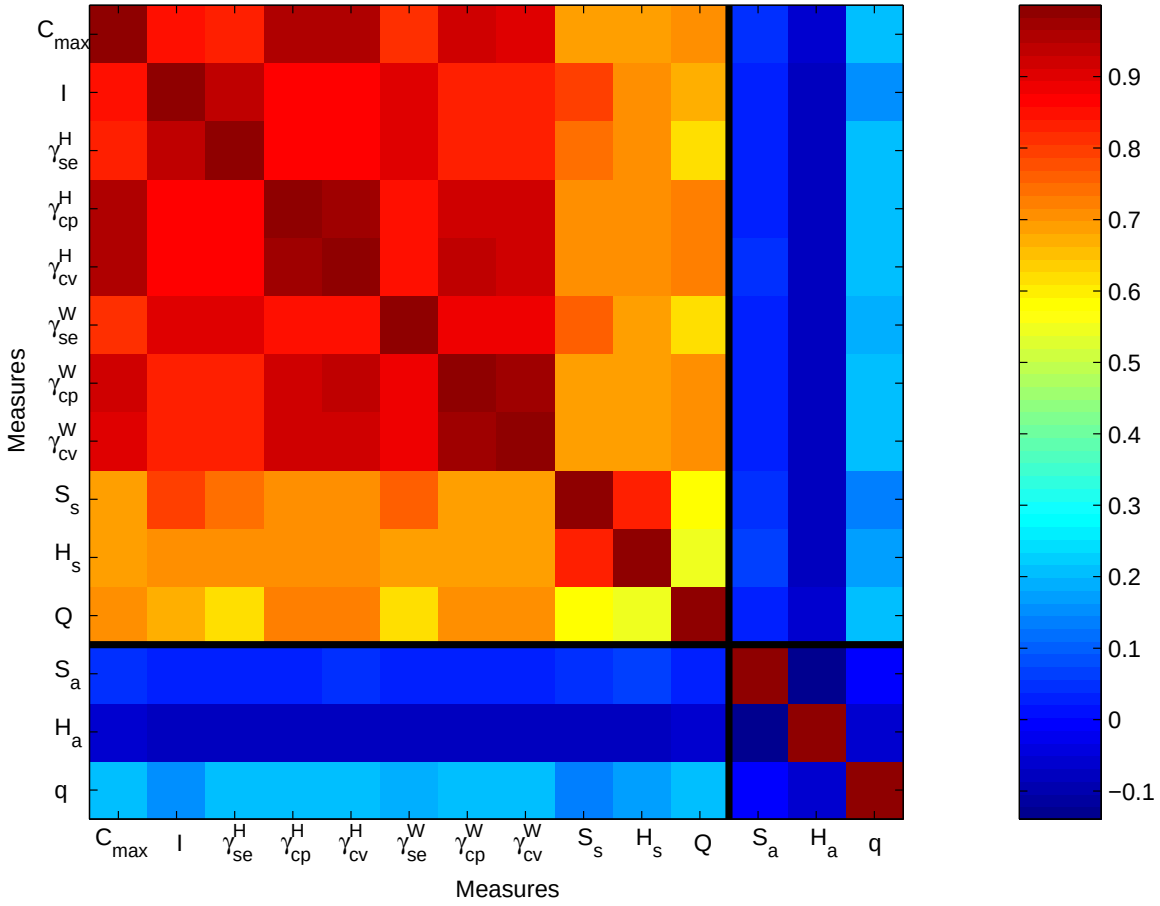


Figure 5.18: Correlation coefficients between the eleven measures of synchronization as well as the three measures of directionality. Values are determined using all 358 channel combinations and all 153548 windows from all nine patients.

The two different groups of measures for synchronization and measures for directionality can be recognized easily. The latter are rather independent from the former, but, quite surprisingly, also from one another. For the anti-symmetric variants of the two nonlinear interdependencies even the highest anti-correlation (-0.14) of all pairs of measures is observed. Among the measures of synchronization the two nonlinear interdependencies and event synchronization differ most from the other measures (but with the minimum value still as high as 0.55). Highest correlations (up to 0.996) can be observed between the two indices based on conditional probability, the two indices based on circular variance and the cross correlation. Also mutual information and the two indices based on Shannon entropy as well as the two nonlinear interdependencies form subclusters.

5.3.3 Discussion

Comparing the performance of the different measures over all evaluation schemes (the first scheme does not contribute substantially due to the weak results of all measures) a ranking over three different groups could be recognized. Maximum and most significant performances were obtained for the three indices of phase synchronization based on the Hilbert transform, closely followed by the mutual information and the cross correlation as the only linear measure evaluated. Intermediate performances were rendered for the three indices of phase synchronization based on the wavelet transform along with event synchronization and all three measures of asymmetry. The weakest and least significant discrimination was yielded by the symmetrized versions of the nonlinear interdependencies. Neglecting small fluctuations this tendency remained quite stable over the last three evaluation schemes. Concerning the different types of synchronization (cf. Section 2.3), measures of phase synchronization yield better results than measures for generalized synchronization. This can be regarded as an indication that different concepts for the quantification of synchronization are not equally well suited for the detection of spatio-temporal changes associated with the generation of epileptic seizures in the brain. This is in good agreement with the results of Ref. [103]. In this study for the subgroup of the first five patients also a comprehensive comparison with twenty-one univariate measures was carried out. These included twelve linear approaches like the statistical moments and the power spectral bands as well as nine nonlinear approaches like an estimate of an effective correlation dimension, the largest Lyapunov exponent and estimates for determinism and entropy. The latter were applied twice, purely and with a surrogate correction to account for specific nonlinear properties. In this comparison substantial differences between univariate and bivariate measures were observed [103]. While univariate measures appeared to be sensitive to changes occurring shortly before a seizure in relation to the period immediately preceding these changes, bivariate measures were found to reflect changes in dynamics on a longer time scale starting hours before a seizure. Both among the univariate and the bivariate approaches linear measures performed equally good as or even better than nonlinear measures [103].

In the present study for all but five measures in the first evaluation scheme the 5 min

smoothing was preferred against no smoothing. Again, this was consistent with the results of Ref. [103] and the explanation given there holds here as well: Smoothing of the measure profiles leads to a decrease in width of both amplitude distributions (the outliers are captured) while their mean values are left unchanged. Thus if the mean value of the pre-ictal distribution differs from the mean value of the inter-ictal distribution, the overlap of the two distributions is decreased by the smoothing. The separability is increased, and higher ROC-values are obtained. If the time profiles of a measure are not correlated, the effect of smoothing is even higher and so is the increase in the discriminative performance.

For the remaining two parameters, the length of the pre-ictal interval and the ROC-hypothesis of separability, a combined interpretation seems to be appropriate. Starting with the three measures of directionality a rather uniform picture could be observed. They showed an increase of pre-ictal values in 10 out of 12 cases and a pre-ictal length of 5 min in 9 out of 12 cases (however, only 5 of these cases proved to be significant). A rather obvious explanation could be the spreading of hypersynchronous activity during the generation of the seizure which naturally results in a certain directionality. Thus the different measures might indicate a driver-responder relationship, e.g., between the focal region and its surrounding. In this study mostly only the focal hemisphere is known (cf. Tab. 5.1) while the exact delineation of the focus is not given (and sometimes this is very difficult if not impossible to achieve anyway). Therefore, the verification of this hypothesis was accomplishable only to a certain extent. In this regard, as in the case of the measures of synchronization also for the measures of directionality a distinct relation to the focal region of each patient could not be seen. Maximum performances were obtained about as often in the focal as in the non-focal hemisphere. Furthermore, because of the short length of the predominant pre-ictal interval and due to ambiguities in the definition of seizure onset the question arises whether the observed effect is indeed pre-ictal or whether it is due to the first stages of seizure generation perceived in the last part of this pre-ictal interval. On the other hand, if the effect would be pre-ictal and if it could successfully be used for a prospective detection of a pre-ictal state, the prediction horizon could possibly still be large enough to allow the successful application of seizure prevention techniques [29]. However, in all these cases the corresponding performance values were surpassed by the (significant) performances of the best measures of synchronization. For these measures, as well as for the aforementioned, only the shortest and the longest possible length of the pre-ictal interval were observed. Again a pre-ictal length of 5 min was found predominantly. For these cases the accompanying ROC-hypothesis was almost uniformly distributed showing as much increases as decreases of synchronization. This can be interpreted as another indication for the high statistical fluctuations associated with this length. For the remaining 11 out of 44 trials the most successful length was 240 min and in all these cases the accompanying ROC-hypothesis was a pre-ictal increase of synchronization.

This is in contrast to the results yielded in Refs. [104, 101, 100] in which long-lasting decreases of synchronization before seizures were reported. In these studies a possible explanation was already given by the hypothesis that a decrease of synchronization is observed

when a recording site becomes ”torn out” of its physiological state of synchronization with a neighboring recording site in one direction and is forced into synchronization with some slowly expanding region of pathologically synchronized neuronal tissue emanating from another direction. According to this hypothesis decreases and increases of synchronization go along with each other and can be interpreted as two different local effects of a more global phenomenon. This point of view was further supported by the observation that an increase in synchronization is indeed frequently found to be associated with a decrease in synchronization in an adjacent channel combination, a phenomenon already reported for the subgroup of patients [103].

Discrepancies between the results of that study and the present ones are due to effects obtained for the latter four patients which were not included before. Some of these changes are very distinct, in particular in the second and third evaluation scheme for which in the earlier study almost exclusively an increase of synchronization using a pre-ictal length of 240 min was reported. This might be due to the different statistical properties of the two subgroups (cf. Tab. 5.1): The average number of seizures of this early collective of patients is higher than 10, while this number is smaller than 4 for the new patients included here.

To account for these non-uniformities the design of the study could be adapted as follows: In the comparison of the different measures the average has been performed over patients to account for the best clinical perspective. All patients have been weighted equally although due to the different number of seizures the analyzed data of the single patients differ a lot in their statistical properties. Thus a better estimate of the performance of a measure could be achievable by weighting different patients according to their different numbers of seizures. This would correspond to averaging over seizures instead of averaging over patients. Also the different implantation schemes involving different numbers of channels and channel combinations could be taken into account in the weighting, since the optimization concerned with this additional degree of freedom is very important in the last three evaluation schemes.

Compared to the investigation on model systems in Chapter 4, in this study it is even more difficult to objectively assess the performance of different measures, since it is not known beforehand whether there really exist specific changes of synchronization or directionality before seizures. If such changes are indicated by one measure but not by another measure, it is hard to say on a theoretical basis which measure is right. Therefore, in this study measures are judged rather pragmatically by their predictive performance using the simple criterion ”Which is the measure that most reliably yields information possibly useful for diagnostic purposes?”. As reported above, according to this criterion the three indices of phase synchronization based on the Hilbert transform, mutual information and cross correlation have to be regarded as most promising.

However, much more research has to be done to evaluate the question whether the statistical performances of these measures are sufficient to yield an acceptable performance in an algorithmic implementation. In case that no measure on its own is able to fulfill this requirement, a combination of different measures could be more promising as long as these

measures render complementary or at least non-redundant information. As for the bivariate measures analyzed in this thesis, correlation coefficients showed that at least measures of synchronization and measures of directionality carry independent information. Among these groups only the measures of the latter were independent from one another. Thus combining different measures of synchronization seems to be reasonable only to a certain extent. Event synchronization and the nonlinear interdependencies were the only measures that differed considerably from the other.

However, besides the bivariate approach reflecting the interaction between different regions of the brain there are two further concepts to track different kinds of spatio-temporal variations in the EEG, namely univariate and multivariate approaches. Applying univariate measures to an EEG signal aims at quantifying the state of the respective region within the brain with respect to a certain property (e.g., dimension). Multivariate approaches (e.g., based on spatial embedding techniques [147] or independent component analysis (ICA) [50]) can extract global information from a multichannel EEG. With respect to the predictive performance, in Refs. [103, 99] substantial differences between univariate and bivariate measures were observed. Since so far rarely a multivariate analysis was carried out, a comprehensive comparison with this approach is still missing. Nevertheless, the assumption that also this approach might yield non-redundant results seems reasonable when regarding the underlying conceptual differences. Thus combining several measures from these different approaches appears to be a promising idea worth trying.

In any case, the development of new approaches to predict epileptic seizures should go along with statistical validation. As announced before, this issue is further addressed in the following section.

5.4 The method of measure profile surrogates

In this study the new method of measure profile surrogates introduced in Section 5.2 is illustrated by exemplarily evaluating the predictive performance of two measures of synchronization, the *index based on circular variance using the Hilbert Transform* and *event synchronization* (cf. Section 3). These measures are calculated from the quasi-continuous EEG recording of patient A (cf. Section 5.3). The seizure prediction statistics applied to the resulting measure profiles and their surrogates is taken from the methodology of Section 5.3.1.3. Again amplitude distributions of pre-ictal and inter-ictal intervals are compared using ROC-curves. The remainder of this study is organized as follows: First the data (Section 5.4.1.1), the measures (Section 5.4.1.2) and the seizure prediction statistics (Section 5.4.1.3) used to demonstrate the new method of measure profile surrogates are described. This method is introduced in Section 5.4.1.4. In Section 5.4.2 the results of the exemplary application are shown, before the conclusions are drawn in Section 5.4.3.

5.4.1 Methods

5.4.1.1 Data

The analysis was carried out using quasi-continuous multi-channel EEG recorded from an epilepsy patient over five days during which the patient had ten epileptic seizures (patient A in Section 5.3, for recording parameters cf. Tab. 5.1). The EEG was recorded prior to and independently from the design of this study during the pre-surgical work-up [31]. Using two implanted depth electrodes each equipped with 10 separate contacts (denoted as TL01,...,TL10 and TR01,...,TR10), the EEG was measured directly within the brain (cf. Fig. 5.1). The EEG contains one major and two minor recording gaps. In addition to the ten ictal and post-ictal intervals (defined from seizure onset until 30 min after seizure termination), four other events known to be associated with changes in the EEG (three sub-clinical seizures and one period of hyperventilation) took place during the acquisition.

5.4.1.2 Measures

From these data two bivariate measures of synchronization were calculated using a moving window technique with non-overlapping segments of 20.48 s corresponding to $N = 4096$ data points. In order to focus on local synchronization effects, in this study only the 18 neighboring channel combinations (TL01-TL02,...,TL09-TL10 and TR01-TR02,...,TR09-TR10) were analyzed. Measures comprise the index based on circular variance using the Hilbert Transform γ_{cv}^{H} (sometimes also termed mean phase coherence R) as a measure for phase synchronization (cf. Section 3.3) and *event synchronization* Q (cf. Section 3.5). Details about their practical implementation can be found in Section 5.3.1.1. For the sake

of brevity, in the following the index based on circular variance using the Hilbert Transform will be referred to as 'phase synchronization' γ_{cv}^{H}.

5.4.1.3 Seizure Prediction Statistics

The seizure prediction statistics to be applied to the original as well as to the surrogate measure profiles is straightforward, simply comparing amplitude distributions of pre-ictal and inter-ictal intervals using the renormalized form of Receiver-Operating-Characteristics (ROC) introduced in Section 5.3.1.3. It is applied to each channel separately according to the second evaluation scheme "Inter-ictal per channel $\Leftrightarrow$ Pre-ictal per channel" described in Section 5.3.1.4 and illustrated in Fig. 5.9.

To investigate the effect of best parameter selection on the statistical validity of the results, the same optimization as already used in Section 5.3 is carried out. For the original and the surrogates the best combination of the two parameters (smoothing and length of the pre-ictal interval) is selected along with the most successful ROC-hypothesis of separability (cf. Section 5.3.1.3). Exactly the same optimization is applied to the original and the surrogates in order to avoid any bias which could fool the interpretation of an acceptance or a rejection of the null hypothesis.

The evaluation of this algorithm is carried out twice for each measure, first regarding each channel combination separately and second after selecting the best channel combination. Thus in the first evaluation scheme there are $2 * 2 * 4 = 16$ different values to choose from for each channel combination. Accordingly, in the second scheme the final performance value for each measure is chosen as the maximum of $16 * 18 = 288$ different values.

5.4.1.4 Measure Profile Surrogates

To test against a certain null hypothesis via a constrained randomization of time series is a well known concept within the framework of nonlinear time series analysis [55, 155]. The original algorithm [166] as well as a number of expansions or refinements [125, 154] are each designed to impose specific constraints on the surrogates and thus to address one particular null hypothesis. In contrast to these standard approaches the method of simulated annealing [152] provides a rather universal means for generating random time series with a wide variety of possible constraints and therefore allows testing of almost arbitrary null hypotheses. Furthermore, the standard algorithms act in the Fourier domain and therefore can produce artifacts because of their implicit assumption of periodic continuation. The resulting edge effect is due to the fact that when preserving the amplitude spectrum, according to the Wiener-Khinchin theorem only the 'periodic' sample autocorrelation function is maintained. In contrast, the method of simulated annealing acts in the time domain and thus is able to preserve the original autocorrelation function. Simulated annealing is also clearly

superior when it comes to the constrained randomization of data with recording gaps. Coping with these gaps is a non-trivial problem for Fourier-based randomization schemes. To treat each segment independently is not a good approach since it is desirable to preserve autocorrelations between different data sets as well. Interpolation schemes might offer a solution for quasi-continuous data sets, but become unfeasible when confronted with long recording gaps. Again, the method of simulated annealing offers a better approach since in the time domain the missing values due to the recording gaps can be set to zero and thus can be neglected in the autocorrelation function.

Simulated annealing (for an overview see [172]) as a method for combinatorial minimization with false minima was introduced in Ref. [56] and was first applied to the generation of surrogates from time series by Schreiber [155]. In short, constraints are specified in terms of a cost function which is then minimized among all possible permutations of the original measure profile. This cost function can be interpreted as the energy E of a thermodynamic system which is annealed slowly towards the global minimum. In this process, starting from an initial random permutation of the original measure profile, randomly chosen pairs of values are exchanged repeatedly until a desired accuracy (i.e., a sufficiently low value of the cost function) is reached. In each iteration step the cost function is updated and depending on the present "temperature" T the exchange is accepted with probability

$$p(\Delta E, T) = \begin{cases} e^{-\frac{\Delta E}{T}} & \Delta E > 0 \\ \\ 1 & \Delta E \leq 0. \end{cases} \tag{5.1}$$

Exchanges with increasing energy are also accepted with non-zero probability to allow escaping from local minima. Whenever a certain number of either tested or accepted exchanges has been performed, the temperature is slowly decreased according to some cooling scheme (e.g., $T_{new} = T_{old} \times \alpha$ with $1 > \alpha \gg 0$).

In the application of this method the three different constraints mentioned in Section 5.2 can easily be imposed on the measure profile surrogates. First of all, recording gaps are preserved by excluding the missing values in the gaps from the permutation scheme. Since all surrogates are permutations of the original measure profile, the amplitude distribution is maintained by construction. The last constraint is the approximate preservation of the autocorrelation function

$$C(\tau) = \frac{1}{N-\tau} \sum_{n=0}^{N-\tau-1} x_{n+\tau} x_n \qquad \tau \geq 0. \tag{5.2}$$

This constraint is formulated in the cost function

$$E = \sum_{\tau=1}^{N-1} \omega_\tau \left| C^{Surr}(\tau) - C^{Ori}(\tau) \right| \tag{5.3}$$

with weights here defined as

$$\omega_\tau = \begin{cases} \frac{1}{\tau} & \tau \leq \tau_{max} \\ 0 & \text{else.} \end{cases} \tag{5.4}$$

A proper choice of these weights is essential. First, they offer the possibility to define the part of the autocorrelation function that should be preserved. This crucially depends on the original autocorrelation function. Four typical examples for the measures and the patient analyzed are depicted in Fig. 5.19. While the autocorrelation function of most channel combinations decays rather fast and does not show any long range correlations, some channel combinations clearly seem to reflect the circadian rhythm resulting approximately in a 24 h periodicity. This different behavior can be judged as an essential property of the individual measure profiles worth to be preserved. To guarantee this, for each measure profile the maximum time lag τ_{max} is set to 4600 windows, thereby ensuring that the first 26 h of the autocorrelation function (given a window length of 20.48 s) are maintained. Without such peculiarities present, a reasonable choice could have been the first zero crossing of the original autocorrelation function.

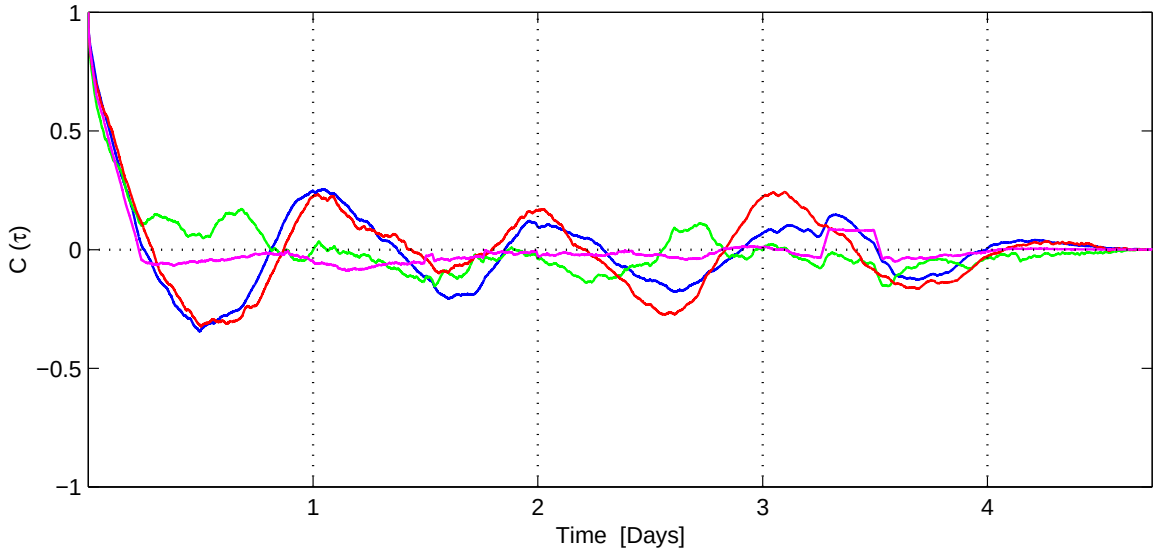


Figure 5.19: Four exemplary autocorrelation functions of original measure profiles for the phase synchronization γ_{cv}^{H}.

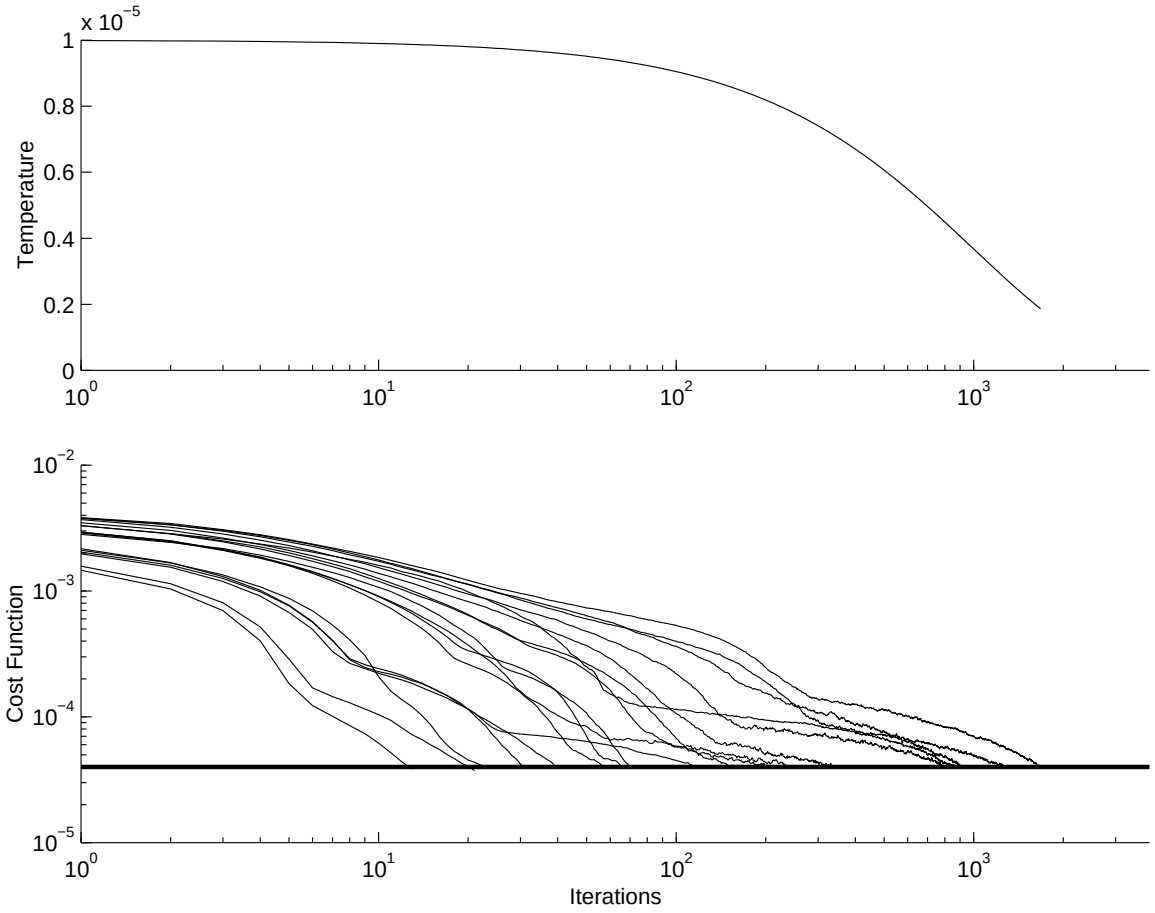


Figure 5.20: Temperature as well as the various cost functions for the 18 different measure profiles of the phase synchronization γ_{cv}^{H} versus the number of iterations (logarithmic scale).

The second issue to be considered when choosing appropriate weights is the computational cost. Typically the number of iterations needed to reach the desired precision is quite large (cf. Fig. 5.20) and in each iteration step an update of the cost function has to be performed. Fortunately this only requires the calculation of those terms of the autocorrelation function to which the two values of the exchanged pair actually contribute. These can further be reduced by setting every other weight to zero. Given the smoothness of the autocorrelation function, the omitted terms are then adjusted automatically. To avoid periodicity artifacts the very first weights are not set to zero. In order to give higher importance to small lags, the remaining terms are weighted by $1/\tau$. Many further possibilities to reduce the high computational cost can be found in Refs. [152, 155].

Using this method of simulated annealing for each measure profile from every channel combination, an ensemble of 19 different measure profile surrogates is generated. Subsequently, the seizure prediction algorithm is applied to the original as well as to the measure profile surrogates. As stated already in Section 5.4.1.3, the evaluation of the algorithm

is carried out twice. Since each measure profile surrogate is generated by a constrained randomization of a single measure profile from one channel combination, in the first evaluation scheme the performance of the two synchronization measures is compared for each channel combination separately. For the original measure profile as well as for each of the 19 surrogates exactly the same optimization is performed, thereby choosing the one out of 16 different combinations of parameters (two different smoothing filters, four different lengths of the pre-ictal interval; pre-ictal increase / decrease) that yields the maximum performance. In the second evaluation scheme for each measure the best channel combination is selected additionally. Here each measures' final performance value is thus chosen as the maximum value out of a set of 288 different possibilities. In each of the two schemes the respective null hypothesis can be rejected with a significance level of $p = 0.05$, if highest values are yielded for the original measure profiles.

Both evaluation schemes test the general null hypothesis H_0 "The measure under investigation is not suited for seizure prediction.". But actually they can be regarded as conceptually different tests with different extended null hypotheses, since they are not based on the same assumptions. Looking at the single channel combinations corresponds to testing for a possible predictive feature consisting of a significantly high number of local effects. Selecting the best channel combination, on the other hand, is aiming at prediction by a maximum local effect. Apart from these two, many other evaluation schemes are conceivable [103]. Averaging over all channel combinations, to name one further example, would test for a global effect. In fact, the choice of an evaluation scheme for the surrogate test constitutes a new degree of freedom which has to be considered carefully. The respective scheme could, in principle, also be incorporated in the null hypothesis, e.g., the extended null hypothesis for the second scheme H_0^{II} could read "The measure under investigation is not suitable to find maximum local effects predictive of epileptic seizures.".

5.4.2 Results

For an exemplary channel combination the original measure profile of the phase synchronization γ_{cv}^{H} as well as four surrogates are depicted in Fig. 5.21. By construction all measure profiles are identical in certain characteristic properties (i.e., the recording gaps, the amplitude distribution and the autocorrelation function up to the maximum time lag) and in this respect each of them can be regarded as a possible original measure profile. However, the surrogates can clearly be distinguished from the original measure profile as well as from one another by the temporal distribution of drops, peaks and quasi-plateaus. The variety among the surrogates clearly demonstrates that the imposed constraints leave enough degrees of freedom for the randomization and do not overspecify the surrogates.

The remaining and most crucial question is whether the original measure profile stands out from the surrogates with respect to its correspondence of the seizure times. To answer this question, the seizure prediction statistics is applied to the original measure profiles as well

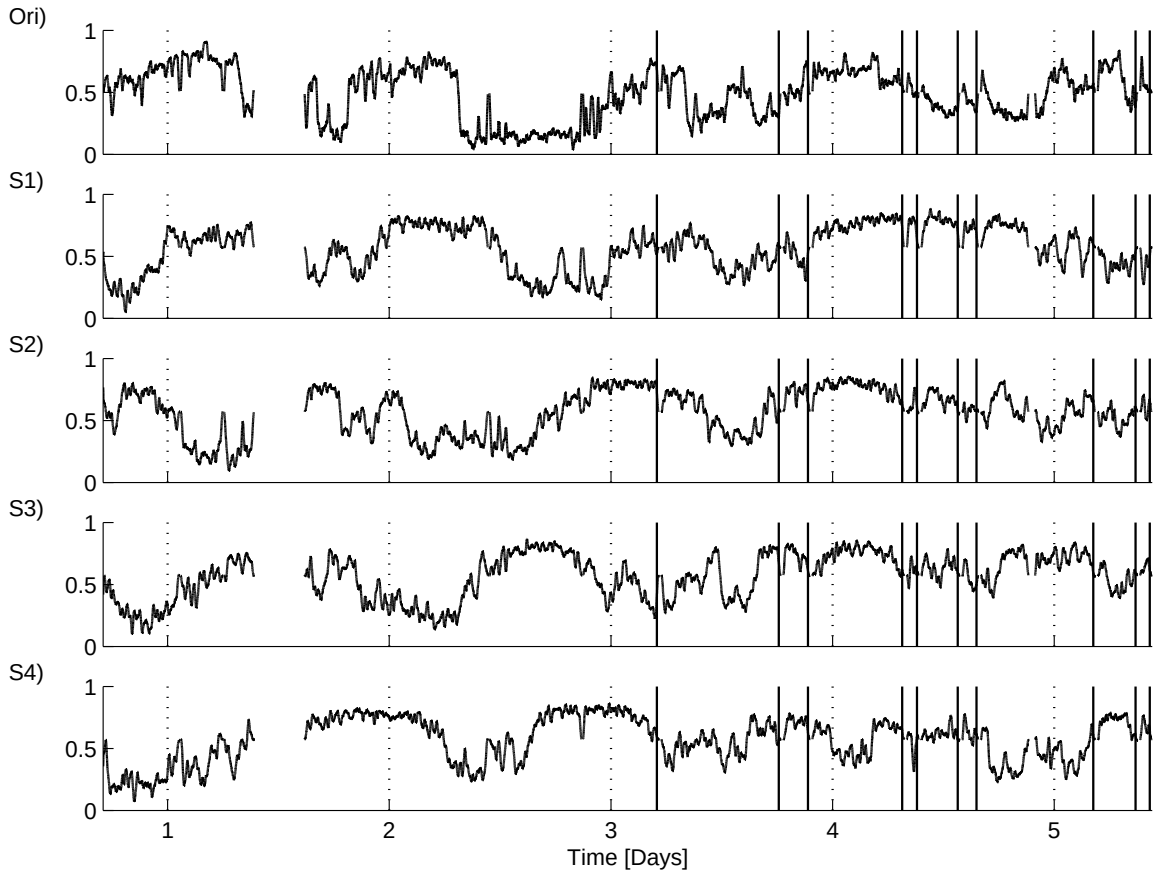


Figure 5.21: Original measure profile of the phase synchronization γ_{cv}^{H} for channel combination TL01-TL02 and four exemplary surrogates, all of them smoothed using a 5 min moving-average filter. Seizures are marked by solid vertical lines.

as to their surrogates. In the first evaluation scheme each channel combination is regarded separately performing exactly the same optimization for the original measure profile as well as for each of the 19 surrogates. The resulting performance values are shown in Fig. 5.22 for the phase synchronization γ_{cv}^{H} and in Fig. 5.23 for the event synchronization Q. Signed ROC-values are depicted to indicate whether a pre-ictal decrease or increase of synchronization is observed for the respective profiles. In order to show the rank of the original performance inside the distribution of the values obtained for the surrogate measure profiles, all performances are sorted by their absolute value.

When considering the performances obtained for the original measure profiles only, highly non-uniform results can be observed. For most channel combinations ROC-values close to zero are obtained reflecting that pre-ictal and inter-ictal amplitude distributions are almost indistinguishable. But for some channel combinations (e.g., TR02-TR03, TR05-TR06 and TR08-TR09) high ROC-values indicating a considerable degree of discrimination between these distributions can be observed, no matter which of the two measures is used. This might correctly reflect the existence of a pre-ictal state which can be detected using either

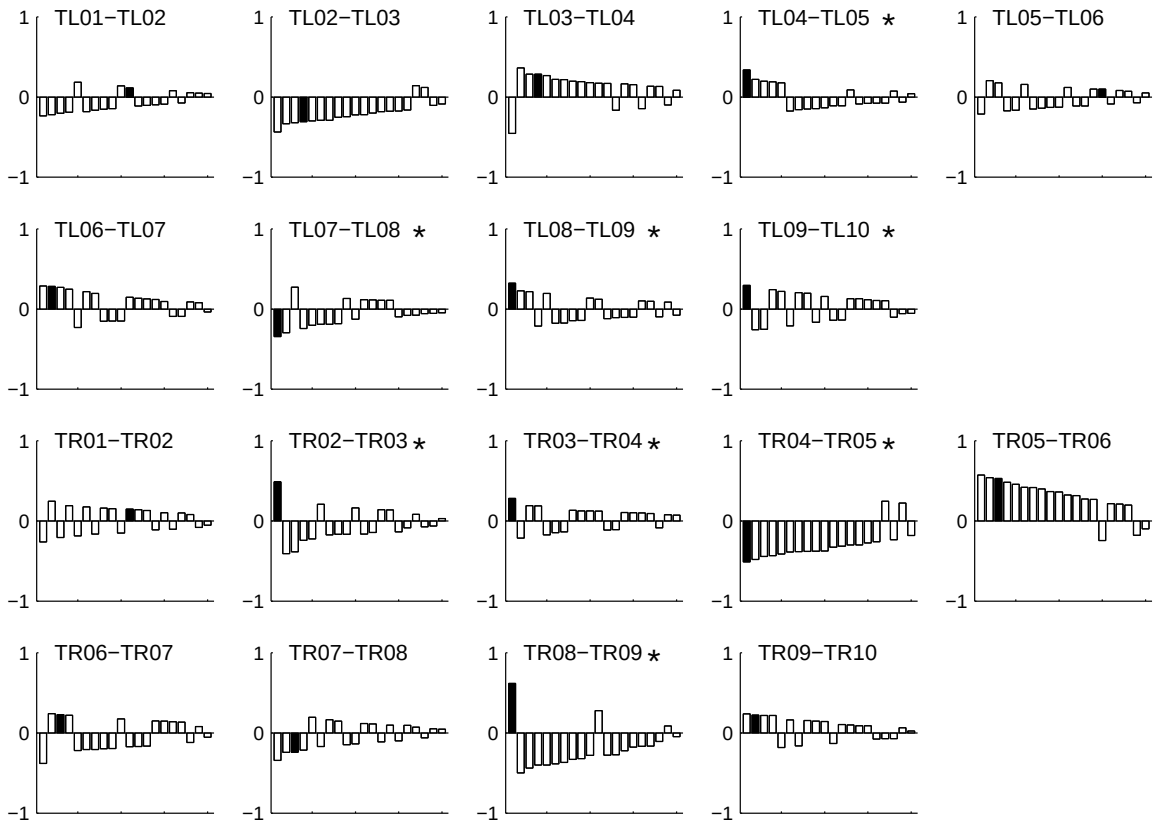


Figure 5.22: Performance values for the first evaluation scheme (parameter optimization is performed for each channel combination separately) of the phase synchronization γ_{cv}^{H} for the original measure profiles (highlighted by filled bars) and the surrogates. For each channel combination of the right and left depth electrode signed ROC-values are depicted, sorted by their absolute value. Asterisks mark channel combinations yielding maximum performance for the original measure profile.

measure, but it could also be the spurious result of statistical fluctuations.

This ambiguity can be resolved by the method of measure profile surrogates. First of all, the information gathered by the surrogates is non-redundant to the information of the original performance values. This can be seen, e.g., when turning the attention to the results of event synchronization in channel combinations TL04-TL05 and TR04-TR05 (cf. Fig. 5.23). In the channel combination from the left hemisphere the absolute performance value obtained for the original measure profile is quite low, but still larger than all values yielded by the surrogates, whereas in the right channel combination a higher absolute performance value is observed, which, however, does not prove to be significant.

In contrast to the high consistency in the two measures' ROC-values regarding the original measure profiles only, qualitatively different results are obtained in the comparison of the performances yielded for the original measure profiles with the ones observed for the surrogates. For the phase synchronization results seem to be significant for 8 out of 18 channel

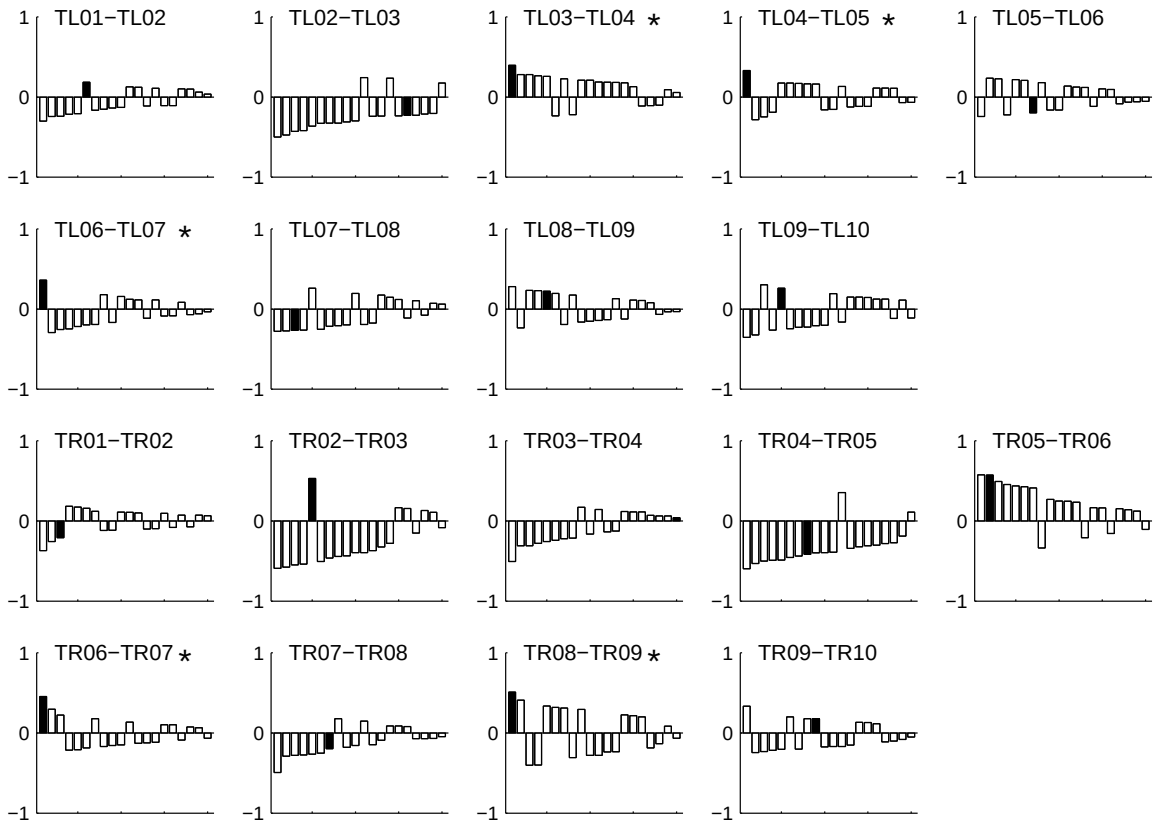


Figure 5.23: Same as Fig. 5.22, but for event synchronization Q.

combinations (TL04-TL05, TL07-TL08, TL08-TL09, TL09-TL10, TR02-TR03, TR03-TR04, TR04-TR05 and TR08-TR09). For event synchronization in 5 channel combinations (TL03-TL04, TL04-TL05, TL06-TL07, TR06-TR07 as well as TR08-TR09) highest absolute ROC-values are obtained for the original measure profiles.

If a hypothesis test with a nominal size p is performed q times, the likelihood P to get at least r rejections by chance is given by

$$P = \sum_{k=r}^{q} \binom{q}{k} p^k (1-p)^{q-k}. \tag{5.5}$$

Here a one-sided test with 19 surrogates (hence $p = 0.05$) is performed for $q = 18$ different channel combinations. This yields probabilities $P(r \geq 8) \approx 10^{-6}$ for the phase synchronization and $P(r \geq 5) \approx 10^{-3}$ for event synchronization. The calculation of these values of significance is based on the implicit assumption that measure profiles from different channel combinations can be regarded as independent. To verify this assumption empirically, the correlation coefficients of all combinations of measure profiles are estimated (cf. Fig. 5.24). The majority of values is close to zero and only rarely a distinct dependence is observed (most prominent the anti-correlations between channel combinations TL02-TL03

and TL03-TL04 as well as channel combinations TR04-TR05 and TR05-TR06 (cf. Ref. [102]). Furthermore, as can be seen from Figs. 5.22 and 5.23, also the performance values obtained for the original measure profiles do not seem to show any kind of clustering for values from neighboring channel combinations. But even when a slight reduction in the number of independent channel combinations is taken into account, the effect remains that the number of channel combinations to show significant ROC-values by itself seems to be significant. Thus apparently the corresponding null hypothesis H_0^I "The measure is not suitable to find a significant number of local effects predictive of epileptic seizures." can be rejected for both measures.

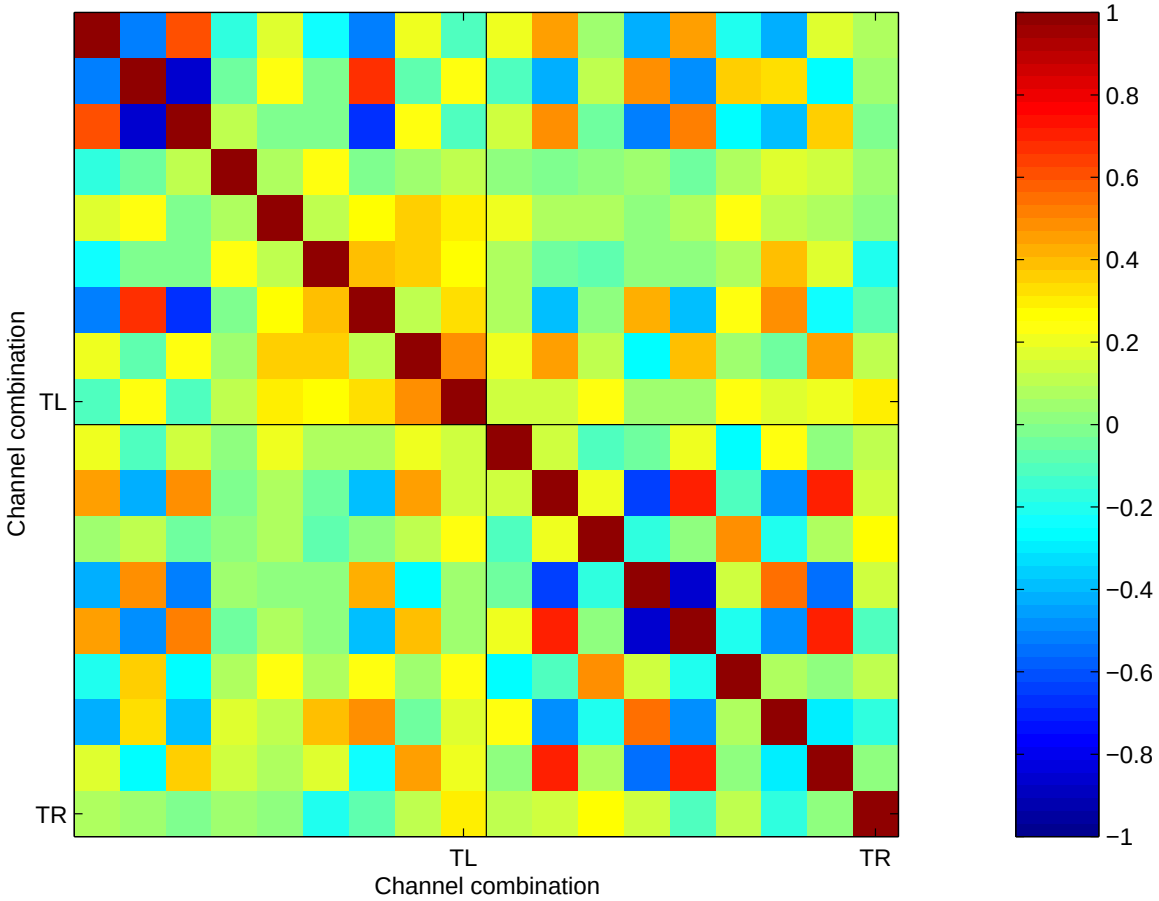


Figure 5.24: Correlation coefficients of the 18 channel combinations (9 from the left and 9 from the right hemisphere) for the phase synchronization γ_{cv}^H. Values are determined using all 18735 time windows of the patient.

When the surrogate test is performed for each channel combination separately, the mean phase coherence already seemed to show a slightly higher level of statistical validity. This difference becomes more striking and even leads to a principal distinction in significance in the second evaluation scheme. Here for each measure and for the original as well as for the 19 surrogates, the channel combination with the highest performance is chosen.

The resulting distributions of over-all performance values are shown in Fig. 5.25. While for the phase synchronization results prove to be significant rendering the highest over-all performance value for an original measure profile, this time the corresponding null hypothesis H_0^{II} can not be rejected for event synchronization. Here the performance value of the best original measure profile falls into the distribution obtained for the ensembles of surrogates.

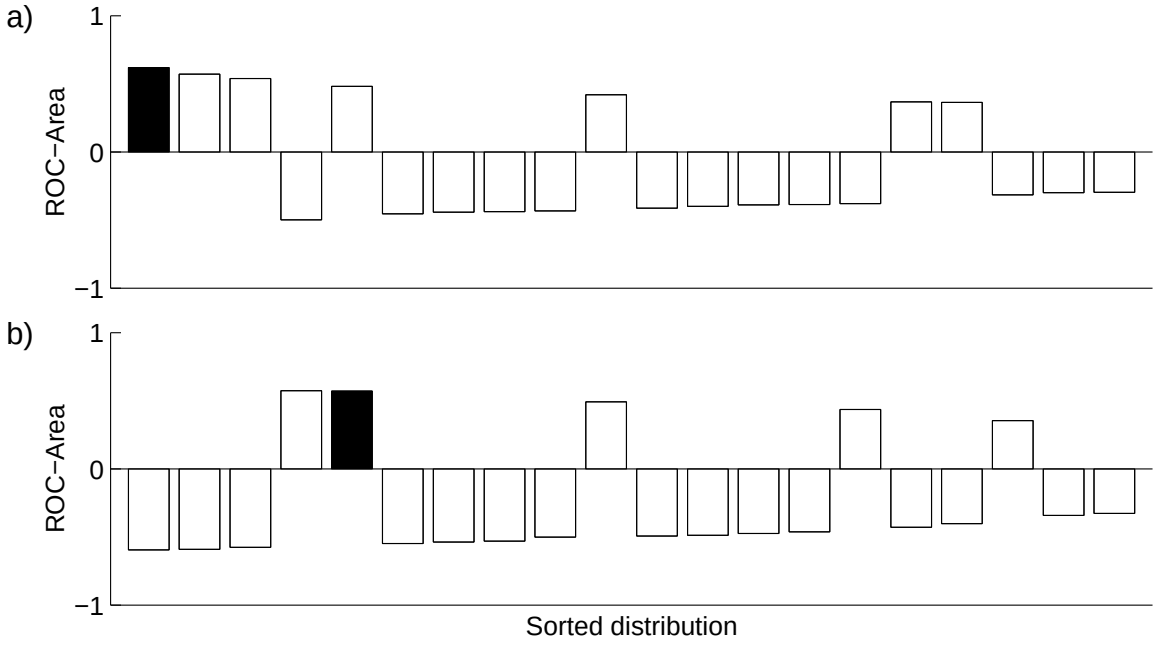


Figure 5.25: ROC-values of the second evaluation scheme (best channel selected) for the original measure profiles (highlighted by a filled bar) and the surrogates, again sorted by their absolute value: (a) phase synchronization γ_{cv}^{H}, (b) Event synchronization Q.

A closer look on the results obtained for event synchronization in Figs. 5.23 and 5.25b reveals that in the second evaluation scheme the best performance yielded by the original measure profile of channel combination TR05-TR06 is surpassed by performances obtained from surrogate measure profiles from other channel combinations, namely TR04-TR05 once and TR02-TR03 twice. This effect is due to the fact, that here an ensemble surrogate test is performed. For each measure the best performance yielded by the entirety of the 18 different original measure profiles is compared to the maximum performance values of 19 surrogate ensembles. These surrogate ensembles preserve the properties of the ensemble of original measure profiles as a whole, since they consist of 18 surrogate measure profiles each of which individually substitutes one of the original measure profiles. When the over-all optimization from the second evaluation scheme is now applied to the original as well as to the surrogate ensembles, it thus can happen that the channel combination yielding the best performance is not the same for the original measure profiles and the surrogate ensembles. This effect is required to investigate the statistical validity of the optimization procedure performed, in this case the best channel selection.

5.4.3 Discussion

Within the method of measure profile surrogates, results obtained from a seizure prediction algorithm are tested against the basic null hypothesis H_0 "The measure under investigation is not suited for the prediction of epileptic seizures.". To demonstrate this approach two different evaluation schemes have been used to investigate the predictive performance of two measures of synchronization, namely the phase synchronization and event synchronization, by means of a straightforward seizure prediction statistics. Measure profile surrogates have been generated by a constrained randomization of the original measure profiles. In the first evaluation scheme the significance of the measures' original performance values has been tested for each channel combination separately resulting in a higher number of significant values for the phase synchronization. Finally, after choosing the best channel combination for each measure in the second scheme an ensemble surrogate test has been performed. Here only the phase synchronization has reached a significant performance value. Thus for event synchronization only null hypothesis H_0^I, for the phase synchronization both null hypotheses H_0^I and H_0^{II} could have been rejected.

A method to statistically validate the performance of epileptic seizure prediction algorithms (such as the proposed method of measure profile surrogates or, if computationally infeasible, alternatively the method of seizure time surrogates [9]) should be applied whenever there is the slightest chance of any 'in-sample' over-optimization. This is the general case since so far rarely a sufficient amount of data is available to perform a proper 'out-of-sample' study, where the recordings are divided into a training set on which all algorithm parameters are adjusted and a test set on which the performance of the algorithm is evaluated.

The method of measure profile surrogates is suited to serve the need for statistical validation of seizure prediction results. On the other hand, in the application of this method there might be some caveats and pitfalls, too (e.g., a hidden bias between the original profiles and the surrogates). Therefore also the results obtained with this method should be interpreted with care and jumping to conclusions too quickly should thoroughly be avoided. In particular, the additional degree of freedom introduced in the choice of a suitable null hypothesis should always be considered. Furthermore, whenever a null hypothesis is rejected, it is always very important to keep in mind that the complementary hypothesis is very comprehensive and might include many different reasons that are possibly responsible for this rejection.

Concerning the practical implementation of this method, in some cases the computational cost can be lowered by simplifying the randomization scheme. Some characterizing measures from time series analysis (e.g., an effective correlation dimension evaluated for seizure prediction in Refs. [84, 15] or the degree of nonlinear determinism applied in Ref. [9]) show measure profiles with a distinct ceiling effect. For these measures, most values lie at the upper or lower end of the definition range, and only rarely sparse deviations (i.e., drops or peaks) can be observed. In such cases the method of simulated annealing does not

seem to be appropriate. A suitable randomization of the original measure profile could be achieved by performing a random shuffle of these deviations instead.

The application of the proposed method of measure profile surrogates is not restricted to the problem of seizure prediction. In principle it is rather universal and can be used for the statistical validation of the performance of time-resolved measures in many other detection and prediction problems. The only requirement is that a finite number of observables is measured and from their analysis certain circumscribed events are to be detected or predicted. Thus many other applications are also conceivable.

Regarding the particular application considered in this study it is important to keep in mind that it was not the aim to prove or disprove the existence of a pre-ictal state, but rather to supply a new and general means to reliably evaluate the statistical validity of the performance of a seizure prediction algorithm. In future applications, measure profile surrogates can be used as a powerful tool to distinguish between measures and algorithms unsuited for the prediction of epileptic seizures, and more promising approaches.

Chapter 6

Summary and Outlook

Synchronization is a frequently occurring phenomenon of high prominence in many scientific and technical disciplines [123]. To gain a better understanding of the dynamical systems involved, the reliable estimation of the synchronization between measured time series is of utmost importance. In this thesis fourteen different measures of synchronization and directionality derived from various theoretical frameworks were compared to each other with respect to their capability to detect and quantify dependencies in dynamical systems. The measures of synchronization comprised the linear cross correlation, the mutual information and six different indices of phase synchronization (either based on Hilbert or on wavelet transform) as well as symmetrized versions of the nonlinear interdependencies and the event synchronization. The anti-symmetrized versions of the last three approaches were used to characterize possible directionalities between two systems.

In the first part of this thesis the measures of synchronization were applied to three different coupled model systems in order to evaluate whether the analysis of these systems can contribute useful information for the decision which measure of synchronization is most suitable for an application to field data. This aim was addressed twofold by comparing measures on the one hand with respect to their capability to reflect the strength of coupling and on the other hand regarding their robustness against noise. In both parts of the study a measure of order was employed as an indicator for a non-monotonic dependence on the coupling strength. Regarding the first aim, the question which measure is best suited to distinguish between different coupling strengths could not be answered in general, since the implicit assumption that an increase of coupling necessarily leads to an increase of all kinds of synchronization does not hold rigorously for every model system. Furthermore, in the comparison of measures rather inconsistent results were obtained for different systems. Therefore, from this part of the analysis no obvious and objective criterion to prefer or to exclude different measures could be derived.

The second part of the study, the dependence on the degree of noise, remained unaffected from these caveats and here indeed an assessment of the different measures could be carried out. All measures were quite robust against white noise with slight advantages for

cross correlation and the phase synchronization indices based on circular variance. As for the robustness against iso-spectral noise, for some model systems and some measures, in particular for the nonlinear interdependencies, a (spurious) synchronization between the contaminating noise-signals could be observed. This should be taken into account when choosing a suitable measure for an application which might be contaminated by a considerable amount of noise.

To conclude, the analysis of coupled model systems is not able to replace a thorough and comprehensive test application on the field data to be analyzed. It is certainly useful to render additional information regarding certain properties of the measures. But the decision which measure to apply to a certain task should in most cases be made rather pragmatically by choosing the measure which most reliably yields valuable information in test applications.

Such a test application was carried out in the second part of this thesis in which the different measures of synchronization and directionality were applied to electroencephalographic time series recorded from epilepsy patients. In particular the challenging task of predicting epileptic seizures was addressed. The given advice to pragmatically use the measure which most reliably yields valuable information is only reasonable when the evaluation of the performances is carried out with much care and paying a great deal of attention to the statistical validation.

Therefore, in this thesis a new approach to address this issue, the method of measure profile surrogates [65, 64], was introduced and compared against the existing method of seizure times surrogates [9]. Many advantages of this method could be shown (e.g., the more natural approach and the easier handling of the constraints), but due to one major disadvantage, its high computational cost, it could not be used in the comparison of the bivariate measures regarding their predictive performance. Nevertheless, in the restricted study on one patient and two measures the new method proved useful in distinguishing between measures and algorithms unsuited for the prediction of epileptic seizures and more promising approaches. In conclusion, measure profile surrogates could be established as a powerful tool to statistically validate epileptic seizure predictions. But since this method is rather universal and can easily be adapted to validate the performance of time-resolved measures in many other detection and prediction problems, and given the fast progress in computer technology, the method of measure profile surrogates can be expected to play a key role in many future applications.

In the comparison of all bivariate measures (cf. Ref. [103]) a statistical evaluation of the predictability of epileptic seizures was carried out analyzing measure profiles rendered from quasi-continuous intracranial EEG recordings. In four different evaluation schemes amplitude distributions of intervals preceding seizures and intervals far away from any seizure activity were compared using Receiver-Operating-Characteristics (ROC). As judged by the pragmatic criterion introduced above, different concepts for the quantification of synchronization proved to be not equally suited for the detection of temporal

changes associated with the generation of epileptic seizures in the human brain. While highest predictive performance values were obtained for the measures of phase synchronization based on the Hilbert transform, the symmetrized versions of the nonlinear interdependencies were least competitive. The three measures of directionality as well as the only linear measure, the cross correlation, ranked between these extremes.

Much more research has to be done to evaluate the question whether the highest performances yielded allow to render an acceptable performance in an algorithmic and prospective implementation. In case that no measure on its own is able to fulfill this requirement, a combination of different measures could be more promising as long as these measures render complementary or at least non-redundant information. For the bivariate measures analyzed in this thesis considerable redundancies between certain measures of synchronization could be observed. Thus a combined use of these measures appears to be reasonable only up to a certain extent. However, because of conceptional differences the combined application with uni- or multivariate approaches (e.g., based on independent component analysis) can be regarded as rather promising. Also the use of bi- or multivariate surrogate techniques [125, 155, 27, 156, 6] could contribute to a better characterization of the spatio-temporal variations in the electroencephalogram of epilepsy patients. Finally, a further exploitation of the directionality of interaction (e.g., using recently proposed methods based on phase synchronization [138, 137, 159]) could lead to a better understanding of synchronization phenomena in the human brain.

Appendix A

Nonlinear deterministic systems

A.1 Hénon map

The Hénon map was the first published example of a strange attractor with less than two dimensions [47]. It can be considered as a two-dimensional extension of the well-known logistic map [96]. Its equations read

$$\begin{aligned} x_1' &= a - x_1^2 + b\, x_2 \\ x_2' &= x_1 \end{aligned} \tag{A.1}$$

For the parameters $a = 1.4$ and $b = 0.3$ an exemplary short segment of the first component, the corresponding power spectrum and the attractor are depicted in Fig. A.1.

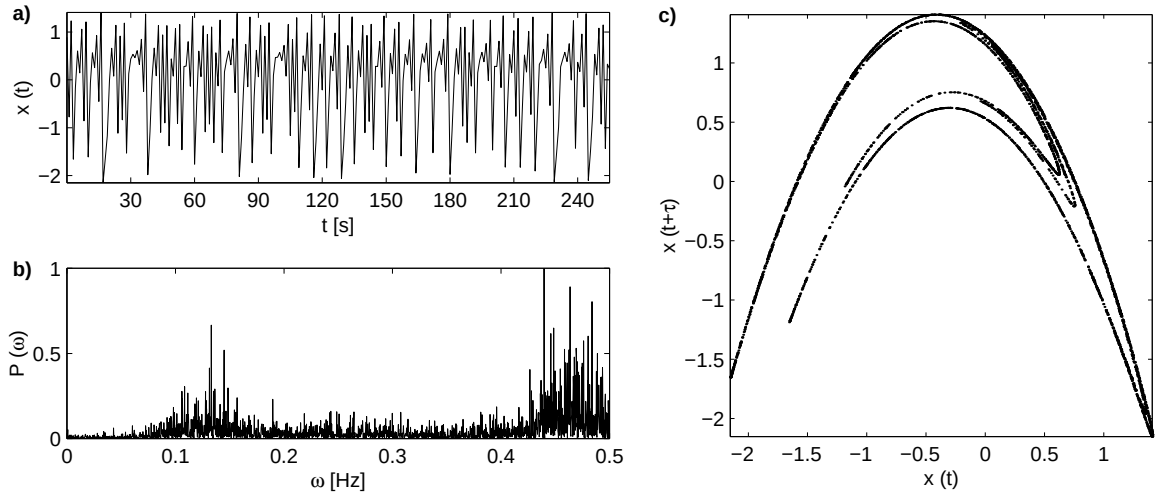


Figure A.1: Hénon map. a) Excerpt of an exemplary time series of the first component. b) Normalized power spectrum. c) Attractor.

A.2 Lorenz system

The differential equations introduced by Lorenz as a simplified model for thermal convection both in the atmosphere and in the fluid of the Rayleigh-Bénard experiment constitute the first and most famous example of a nonlinear dynamical system exhibiting chaotic behavior [91]. The strange attractor with its butterfly shape became one of the most seen symbols of chaos. The equations of the Lorenz systems are

$$\begin{aligned} \frac{dx_1}{dt} &= \sigma(-x_1 + x_2) \\ \frac{dx_2}{dt} &= r\ x_1 - x_2 - x_1\ x_3 \\ \frac{dx_3}{dt} &= x_1\ x_2 - b x_3. \end{aligned} \tag{A.2}$$

In Fig. A.2 an exemplary segment of the first component, the corresponding power spectrum and a projection of the attractor onto the $(x_1 x_3)$-plane are shown for the parameters $\sigma = 10$, $r = 28$ and $b = \frac{8}{3}$.

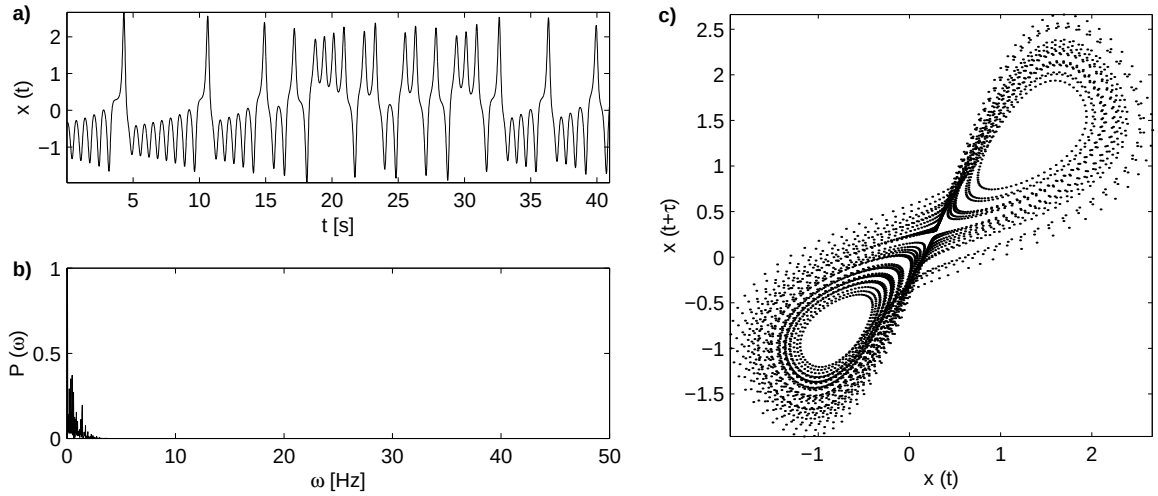


Figure A.2: Lorenz system. a) Exemplary time series of the first component. b) Normalized power spectrum. c) Projection of the attractor onto the $(x_1 x_3)$-plane.

A.3 Rössler system

The differential equations of the Rössler system represent the simplest example of a dynamical system with a strange attractor [143]. While the first two equations describe a linear damped oscillator, the only nonlinear term $x_1 x_3$ is contained in the third one:

$$\begin{aligned} \frac{dx_1}{dt} &= -\omega\{x_2 + x_3\} \\ \frac{dx_2}{dt} &= \omega\{x_1 + a\ x_2\} \\ \frac{dx_3}{dt} &= \{b + x_3(x_1 - c)\}. \end{aligned} \tag{A.3}$$

For the parameters $a = 0.15$, $b = 0.2$ and $c = 10$ an exemplary segment of the first component, the corresponding power spectrum and a projection of the attractor onto the $(x_1 x_2)$-plane are depicted in Fig. A.3. The characteristic folding mechanism is clearly visible.

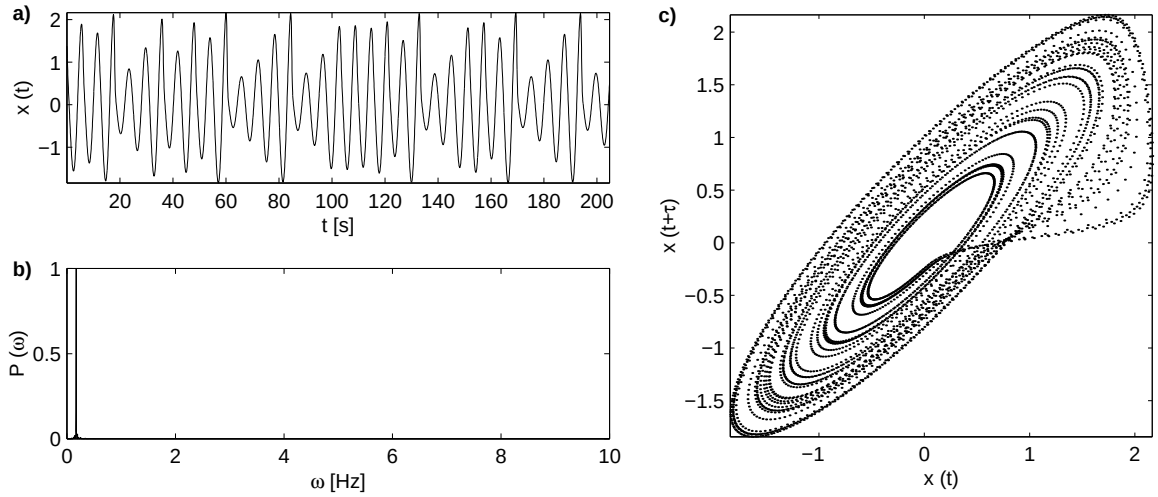


Figure A.3: Rössler system. a) Exemplary time series of the first component. b) Normalized power spectrum. c) Projection of the attractor onto the $(x_1 x_2)$-plane.

Appendix B

Stochastic signals

B.1 White noise

The most commonly used stochastic signal is a sequence of random numbers with either a uniform amplitude distribution in a certain interval (e.g., [0,1]) or a Gaussian amplitude distribution. In the asymptotic case of infinitely long signals all frequencies contribute equally, leading to a homogenous frequency spectrum. In analogy to the spectrum of visible light such signals are called white noise. In Fig. B.1 an exemplary segment of Gaussian white noise, the corresponding power spectrum and a state space portrait are shown.

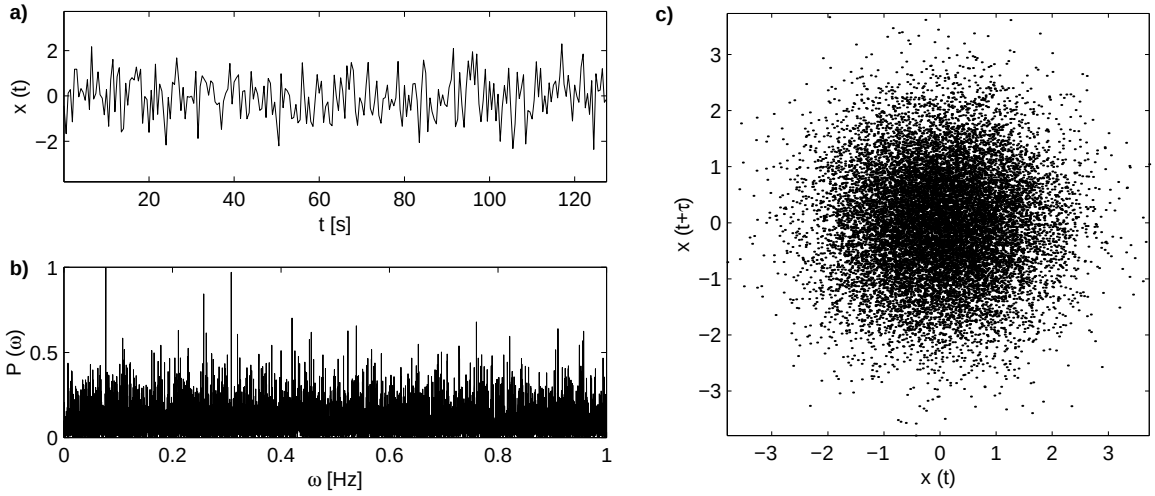


Figure B.1: Gaussian white Noise. a) Excerpt of an exemplary time series. b) Normalized power spectrum. c) State space portrait.

B.2 Iso-spectral noise

Iso-spectral noise is characterized by the same power spectrum as the original time series and is generated by means of phase-randomized surrogates first introduced in Ref. [166]. The generation of these surrogates consists of three steps: Fourier transform, randomization of phases and inverse Fourier transform. In Fig. B.2 an example of a time series from the Lorenz system and its phase-randomized surrogate is depicted. They look very different from each other although their power spectrums are indistinguishable.

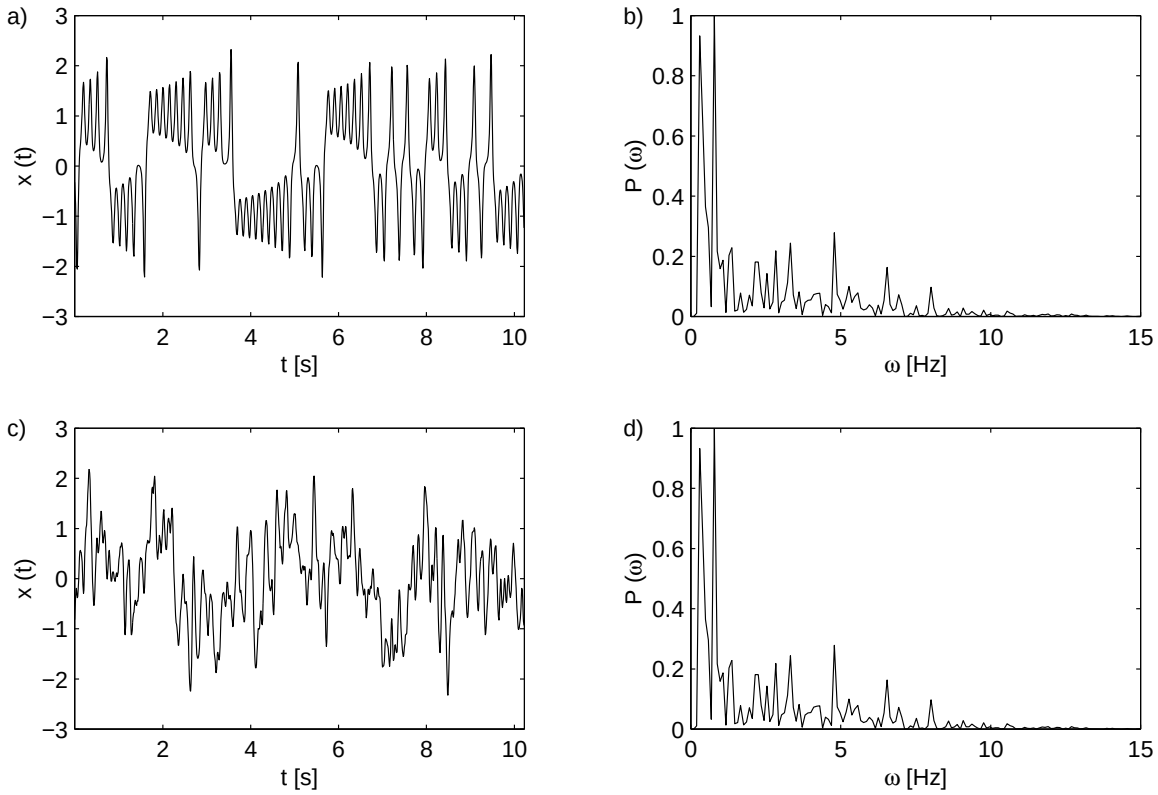


Figure B.2: Iso-spectral surrogate: Exemplary time series (Sampling rate: 100 Hz) of the first component of the Lorenz system (a) with its normalized power spectrum (b). The phase-randomized surrogate (c) looks completely different but its power spectrum is completely indistinguishable (d).

Bibliography

[1] H. D. I. Abarbanel, N. Rulkov, and M. Sushchik. Generalized synchronization of chaos: The auxiliary system approach. *Phys. Rev. E*, 53:4528, 1996.

[2] V. S. Afraimovich, N. N. Verichev, and M. I. Rabinovich. Stochastic synchronization of oscillation in dissipative systems. *Radiophys. Quantum Electron.*, 29:795, 1986.

[3] R. B. Aird. The importance of seizure-inducing factors in the control of refractory forms of epilepsy. *Epilepsia*, 24:567, 1983.

[4] R. G. Andrzejak. Anteile nichtlinearer Determinismen in Zeitreihen: Theorie und Simulation sowie Anwendung auf hirnelektrische Aktivität von Patienten der prächirurgischen Epilepsiediagnostik. Master's thesis, Department of Physics, University of Bonn, Germany, 1997.

[5] R. G. Andrzejak. *Epilepsie als eine nichtlinear deterministische Dynamik: Eine Untersuchung hirnelektrischer Aktivität mit Methoden der linearen und nichtlinearen Zeitreihenanalyse*. PhD thesis, Department of Physics, University of Bonn, Germany, 2001.

[6] R. G. Andrzejak, A. Kraskov, H. Stögbauer, F. Mormann, and T. Kreuz. Bivariate surrogate techniques: Necessity, strengths, and caveats. *Phys. Rev. E (in press)*, 2003.

[7] R. G. Andrzejak, T. Kreuz, F. Mormann, G. Widmann, C. E. Elger, and K. Lehnertz. Improved characterization of neuronal dynamics by focusing on nonlinearity. *(submitted)*, 2003.

[8] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Phys. Rev. E*, 64:061907, 2001.

[9] R. G. Andrzejak, F. Mormann, T. Kreuz, C. Rieke, A. Kraskov, C. E. Elger, and K. Lehnertz. Testing the null hypothesis of the nonexistence of a preseizure state. *Phys. Rev. E*, 67:010901, 2003.

[10] R. G. Andrzejak, G. Widman, K. Lehnertz, P. David, and C. E. Elger. The epileptic process as nonlinear deterministic dynamics in a stochastic environment: An evaluation on mesial temporal lobe epilepsy. *Epilepsy Res.*, 44:129, 2001.

[11] J. F. Annegers. The epidemiology of epilepsy. In E. Wyllie, editor, *The treatment of epilepsy: Principles and practice*, page 165. Williams and Wilkins, Baltimore, 1996.

[12] E. V. Appleton. The automatic synchronization of triode oscillator. *Proc. Cambridge Phil. Soc. (Math. and Phys. Sci.)*, 21:231, 1922.

[13] J. Arnhold. *Nichtlineare Analyse raum-zeitlicher Aspekte der hirnelektrischen Aktivität von Epilepsiepatienten*. PhD thesis, Department of Physics, University of Wuppertal, Germany, 2000.

[14] J. Arnhold, K. Lehnertz, P. Grassberger, and C. E. Elger. A robust method for detecting interdependences: Application to intracranially recorded EEG. *Physica D*, 134:419, 1999.

[15] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. U. Voss, J. Timmer, and A. Schulze-Bonhage. How well can epileptic seizures be predicted? An evaluation of a nonlinear method. *Brain (in press)*, 2003.

[16] J. S. Barlow. Methods of analysis of nonstationary EEGs with emphasis on segmentation techniques: A comparative review. *J. Clin. Neurophysiol.*, 2:267, 1985.

[17] N. N. Bartlett. On the theoretical specification and sampling properties of autocorrelated time-series. *J. R. Stat. Soc.*, B8:27, 1946.

[18] H. Berger. Über das Elektroenkephalogramm des Menschen. *Arch Psychiat Nervenkrankh*, 87:527, 1929.

[19] G. E. P. Box and G. M. Jenkins. *Time series analysis: Forecasting and control. Revised Ed.* Holden-Day, San Francisco, 1993.

[20] D. S. Broomhead and G. P. King. Extracting qualitative dynamics from experimental data. *Physica D*, 20:217, 1986.

[21] L. Callenbach, P. Hänggi, S. J. Linz, J. A. Freund, and L. Schimansky-Geier. Oscillatory systems driven by noise: frequency and phase synchronization. *Phys. Rev. E*, 65:051110, 2002.

[22] T. L. Carroll and L. M. Pecora. Cascading synchronized chaotic systems. *Physica D*, 67:126, 1993.

[23] B. W. Colder, C. L. Wilson, R. C. Frysinger, L. C. Chao, R. M. Harper, and J. Engel Jr. Neuronal synchrony in relation to burst discharge in epileptic human temporal lobes. *J. Neurophysiol.*, 65:2496, 1996.

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

[25] G. A. Darbellay and I. Vajda. Estimation of the information by an adaptive partitioning of the observation space. *IEEE Trans. Inform. Th.*, 45:1315, 1999.

[26] W. De Clerq, P. Lemmerling, S. van Huffel, and W. van Paesschen. Anticipation of epileptic seizures from standard EEG recordings. *Lancet*, 361:970, 2003.

[27] K. T. Dolan and A. Neiman. Surrogate analysis of coherent multichannel data. *Phys. Rev. E*, 65:026108, 2002.

[28] R. B. Duckrow and S. S. Spencer. Regional coherence and the transfer of ictal activity during seizure onset in the medial temporal lobe. *Electroencephalogr. Clin. Neurophysiol.*, 82:415, 1992.

[29] C. E. Elger. Future trends in epileptology. *Curr. Opin. Neurol.*, 14:185, 2001.

[30] C. E. Elger and K. Lehnertz. Seizure prediction by nonlinear time series analysis of brain electrical activity. *Eur. J. Neurosci.*, 10:786, 1998.

[31] J. Engel Jr, editor. *Surgical Treatment of the Epilepsies*. Raven Press, New York, 1993.

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

[33] L. Fabiny, P. Colet, and R. Roy. Coherence and phase dynamics of spatially coupled solid-state lasers. *Phys. Rev. A*, 47:4287, 1993.

[34] J. Fell, P. Klaver, K. Lehnertz, T. Grunwald, C. Schaller, C. E. Elger, and G. Fernández. Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. *Nature Neurosci.*, 4:1259, 2001.

[35] J. Fell, J. Röschke, and C. Schäffner. Surrogate data analysis of sleep electroencephalograms reveals evidence for nonlinearity. *Biol. Cybern.*, 75:85, 1996.

[36] A. Fraser and H. Swinney. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, 33:1134, 1986.

[37] H. Fujisaka and T. Yamada. Stability theory of synchronized motion in coupled dynamical systems. *Prog. Theor. Phys.*, 69:32, 1983.

[38] D. Gabor. Theory of communication. *Proc. IEEE London*, 93:429, 1946.

[39] E. S. Goldensohn and D. P. Purpura. Intracellular potentials of cortical neurons during focal epileptogenic discharges. *Science*, 139:840, 1963.

[40] J. Gotman, J. Ives, P. Gloor, A. Olivier, and L. Quesney. Changes in inter-ictal EEG spiking and seizure occurrence in humans. *Epilepsia*, 23:432, 1982.

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

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

[43] R. Gray. *Entropy and Information Theory*. Springer Verlag, New York, 1990.

[44] J. A. Hanley and B. J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143:29, 1982.

[45] S. R. Haut, C. Swick, K. Freeman, and S. Spencer. Seizure clustering during epilepsy monitoring. *Epilepsia*, 43:711, 2002.

[46] J. F. Heagy, T. L. Caroll, and L. M. Pecora. Synchronous chaos in coupled oscillator systems. *Phys. Rev. E*, 50:1874, 1994.

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

[48] B. Hjorth. EEG analysis based on time domain properties. *Electroenceph Clin Neurophysiol*, 29:306, 1970.

[49] C. Huygens. *Horoloquium Oscilatorium*. Paris, 1673.

[50] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent component analysis*. John Wiley and Sons, New York, 2001.

[51] L. D. Iasemidis, P. Pardalos, J. C. Sackellares, and D. S. Shiau. Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures. *J. Combinatorial Optimization*, 5:9, 2001.

[52] L. D. Iasemidis, J. C. Sackellares, H. P. Zaveri, and W. J. Williams. Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures. *Brain Topogr.*, 2:187, 1990.

[53] K. K. Jerger, T. I. Netoff, J. T. Francis, T. Sauer, L. Pecora, S. L. Weinstein, and S. J. Schiff. Early seizure detection. *J. Clin. Neurophysiol.*, 18(3):259, 2001.

[54] H. Kantz. Quantifying the closeness of fractal measures. *Phys. Rev. E*, 49:5091, 1994.

[55] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge Univ. Press, Cambridge, UK, 1997.

[56] S. Kirkpatrik, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671, 1983.

[57] L. Kocarev and U. Parlitz. General approach for chaotic synchronization with applications to communication. *Phys. Rev. Lett.*, 74:5028, 1995.

[58] K. Krajsek. Charakterisierung der Dynamik räumlich ausgedehnter Systeme mit der multivariaten spatio-temporalen Zustandsraumkonstruktion. Master's thesis, Department of Physics, University of Bonn, Germany, 2003.

[59] A. Kraskov. Personal communication.

[60] A. Kraskov, T. Kreuz, R. G. Andrzejak, H. Stögbauer, W. Nadler, and P. Grassberger. Extracting phases from aperiodic signals. *(submitted)*, 2003.

[61] A. Kraskov, T. Kreuz, R. Quian Quiroga, P. Grassberger, F. Mormann, K. Lehnertz, and C. E. Elger. Phase synchronization using continuous wavelet transform of the EEG for interictal focus localization in mesial temporal lobe epilepsy. *Epilepsia*, 42(7):43, 2001.

[62] A. Kraskov, H. Stögbauer, and P. Grassberger. Estimating mutual information. *(submitted)*, 2003.

[63] T. Kreuz. Symbolische Dynamik in nichtlinearen Modellsystemen und Zeitreihen hirnelektrischer Aktivität. Master's thesis, Department of Physics, University of Bonn, Germany, 1999.

[64] T. Kreuz, R. G. Andrzejak, A. Kraskov, F. Mormann, H. Stögbauer, C. E. Elger, P. Grassberger, and K. Lehnertz. Time profile surrogates: A new method to validate the performance of seizure prediction algorithms. *Epilepsia*, 44(7):231, 2003.

[65] T. Kreuz, R. G. Andrzejak, F. Mormann, A. Kraskov, H. Stögbauer, C. E. Elger, K. Lehnertz, and P. Grassberger. Measure profile surrogates: A method to validate the performance of epileptic seizure prediction algorithms. *(submitted)*, 2003.

[66] T. Kreuz, A. Kraskov, R. G. Andrzejak, F. Mormann, C. Rieke, P. Grassberger, K. Lehnertz, and C. E. Elger. Seizure prediction: Quantifying the performance of measures in distinguishing pre-ictal from inter-ictal states. *Epilepsia*, 43(7):48, 2002.

[67] T. Kreuz, A. Kraskov, R. Quian Quiroga, P. Grassberger, R. G. Andrzejak, F. Mormann, C. Rieke, K. Lehnertz, and C. E. Elger. The capability of different interdependence measures to predict epileptic seizures. *Epilepsia*, 42(7):39, 2001.

[68] T. Kreuz, K. Lehnertz, P. David, and C. E. Elger. Symbolic dynamics: Reducing the information content of the EEG for interictal focus localization in mesial temporal lobe epilepsy. *Epilepsia*, 41(7):212, 2000.

[69] T. Kreuz, R. Quian Quiroga, P. Grassberger, K. Lehnertz, and C. E. Elger. Interdependencies in intracranial EEG recordings of epilepsy patients: A comparison of different measures. *Epilepsia*, 42:49, 2001.

[70] J. P. Lachaux, E. Rodriguez, M. Le Van Quyen, A. Lutz, J. Martinerie, and F. J. Varela. Studying single-trials of phase-synchronous activity in brain. *Int. J. Bifurcation Chaos Appl. Sci. Eng.*, 10:2429, 2000.

[71] J. P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.*, 8:194, 1999.

[72] Y. Lai, M. A. F. Harrison, M. G. Frey, and I. Osorio. Inability of Lyapunov exponents to predict epileptic seizures. *Phys. Rev. Lett.*, 91:068102, 2003.

[73] J. Laidlaw, A. Richens, and J. Oxley. *A Textbook of Epilepsy*. Churchill Livingstone, New York, 1988.

[74] H. H. Lange, J. P. Lieb, J. Engel Jr., and P. H. Crandall. Temporo-spatial patterns of pre-ictal spike activity in human temporal lobe epilepsy. *Electroencephalogr. Clin. Neurophysiol.*, 56:543, 1983.

[75] M. Le Van Quyen, C. Adam, M. Baulac, J. Martinerie, and F. J. Varela. Nonlinear interdependencies of EEG signals in human intracranially recorded temporal lobe seizures. *Brain Res.*, 792:24, 1998.

[76] M. Le Van Quyen, C. Adam, J. Martinerie, M. Baulac, S. Clémenceau, and F. J. Varela. Spatio-temporal characterization of non-linear changes in intracranial activities prior to human temporal lobe seizures. *Eur. J. Neurosci.*, 12:2124, 2000.

[77] M. Le Van Quyen, J. Foucher, J. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, and F. J. Varela. Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *J. Neurosci. Meth.*, 111:83, 2001.

[78] M. Le Van Quyen, J. Martinerie, C. Adam, and F. J. Varela. Nonlinear analyses of interictal EEG map the brain interdepences in human focal epilepsy. *Physica D*, 127:250, 1999.

[79] M. Le Van Quyen, J. Martinerie, M. Baulac, and F. J. Varela. Anticipating epileptic seizure in real time by a nonlinear analysis of similarity between EEG recordings. *Neuroreport*, 10:2149, 1999.

[80] M. Le Van Quyen, J. Martinerie, V. Navarro, P. Boon, M. D'Havé, C. Adam, B. Renault, M. Baulac, and F. J. Varela. Anticipation of epileptic seizures from standard EEG recordings. *Lancet*, 357:183, 2001.

[81] K. Lehnertz. *Nichtlineare Zeitreihenanalysen intrakraniell registrierter hirnelektrischer Aktivität: Charakterisierung der räumlich-zeitlichen Dynamik des primären epileptogenen Areals von Patienten mit Schläfenlappenepilepsie*. PhD thesis, Department of Physics, University of Bonn, Germany, 1997.

[82] K. Lehnertz, R. G. Andrzejak, J. Arnhold, T. Kreuz, F. Mormann, C. Rieke, G. Widman, and C. E. Elger. Nonlinear EEG analysis in epilepsy: Its possible use for interictal focus localization, seizure anticipation, and prevention. *J. Clin. Neurophysiol.*, 18:209, 2001.

[83] K. Lehnertz, J. Arnhold, P. Grassberger, and C. E. Elger. *Chaos in Brain?* World Scientific, Singapore, 2000.

[84] K. Lehnertz and C. E. Elger. Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity. *Phys. Rev. Lett.*, 80:5019, 1998.

[85] K. Lehnertz and B. Litt. The first international collaborative workshop on seizure prediction: Summary and data descriptions. *(submitted)*, 2003.

[86] K. Lehnertz, F. Mormann, T. Kreuz, R. G. Andrzejak, C. Rieke, P. David, and C. E. Elger. Seizure prediction by nonlinear EEG analysis. *IEEE Eng. Med. Biol.*, 22:57, 2003.

[87] B. Litt and J. Echaux. Prediction of epileptic seizures. *Lancet Neurol*, 1:22, 2002.

[88] B. Litt, R. Esteller, J. Echaux, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, M. Dichter, and G Vachtsevanos. Epileptic seizures may begin hours in advance of clinical onset: A report of five patients. *Neuron*, 30:51, 2001.

[89] B. Litt and K. Lehnertz. Seizure prediction and the preseizure period. *Curr Opin Neurol*, 15:173, 2002.

[90] F. H. Lopes da Silva. EEG analysis: Theory and practice. In E. Niedermayer and F. H. Lopes da Silva, editors, *Electroencephalography, basic principles, clinical applications and related fields*, page p. 1097. Urban and Schwarzenberg, 3rd Ed. (Williams and Wilkins, Baltimore), 1993.

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

[92] K. V. Mardia. *Probability and mathematical statistics: Statistics of directional data*. Academy Press, London, 1972.

[93] J. Martinerie, C. Adam, M. Le Van Quyen, M. Baulac, S. Clemenceau, B. Renault, and F. J. Varela. Epileptic seizures can be anticipated by non-linear analysis. *Nature Med.*, 4:1173, 1998.

[94] H. Matsumoto and C. Ajmone-Marsan. Cortical cellular phenomena in experimental epilepsy: Interictal manifestations. *Exp. Neurol.*, 9:286, 1964.

[95] H. Matsumoto and C. Ajmone-Marsan. Cortical cellular phenomena in experimental epilepsy: Ictal manifestations. *Exp. Neurol.*, 9:305, 1964.

[96] R. M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459, 1976.

[97] P. E. McSharry, L. E. Smith, and L. Tarassenko. Prediction of epileptic seizures: Are nonlinear methods relevant? *Nature Medicine*, 9(3):241, 2003.

[98] F. Mormann. Synchronisationsphänomene in synthetischen Zeitreihen und Zeitreihen hirnelektrischer Aktivität. Master's thesis, Department of Physics, University of Bonn, Germany, 1998.

[99] F. Mormann. *Synchronization phenomena in the human epileptic brain*. PhD thesis, Department of Physics, University of Bonn, Germany, 2003.

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

[101] F. Mormann, T. Kreuz, R. G. Andrzejak, P. David, K. Lehnertz, and C. E. Elger. Epileptic seizures are preceded by a decrease in synchronization. *Epilepsy Res.*, 53:173, 2003.

[102] F. Mormann, T. Kreuz, R. G. Andrzejak, C. Rieke, A. Kraskov, P. David, C. E. Elger, and K. Lehnertz. Preictal state detection in continuous intracranial EEG recordings based on decresed phase synchronization: Problems and pitfalls. *Epilepsia*, 43(7):121, 2002.

[103] F. Mormann, T. Kreuz, C. Rieke, R. G. Andrzejak, A. Kraskov, P. David, C. E. Elger, and K. Lehnertz. On the predictability of epileptic seizures. *(submitted)*, 2003.

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

[105] V. Navarro, J. Martinerie, M. Le Van Quyen, S. Clemenceau, C. Adam, M. Baulac, and F Varela. Seizure anticipation in human neocortical partial epilepsy. *Brain*, 125:640, 2002.

[106] T. I. Netoff and S. J. Schiff. Decreased neuronal synchronization during experimental seizures. *J. Neurosci.*, 22:7297, 2002.

[107] E. Niedermeyer. *The Epilepsies - Diagnosis and Management*. Urban and Schwarzenberg, Baltimore, 1990.

[108] G. Osipov, A. Pikovsky, M. Rosenblum, and J. Kurths. Phase synchronization effects in a lattice of nonidentical Rössler oscillators. *Phys. Rev. E*, 55:2353, 1997.

[109] H. Osterhage. Synchronisation und Interdependenz: Eine vergleichende Untersuchung an Modellsystemen und Zeitreihen hirnelektrischer Aktivität. Master's thesis, Department of Physics, University of Bonn, Germany, 2003.

[110] M. Palus. Detecting phase synchronization in noisy systems. *Phys. Lett. A*, 227:301, 1997.

[111] M. Palus, V. Komárec, Z. Hrncír, and K. Strebová. Synchronization as adjustment of information rates: Detection from bivariate time series. *Phys. Rev. E*, 63:046211, 2001.

[112] M. Palus and A. Stefanovska. Direction of coupling from phases of interacting oscillators: An information-theoretic approach. *Phys. Rev. E*, 67:055201, 2003.

[113] P. Panter. *Modulation, noise, and spectral analysis*. McGraw-Hill, New York, 1965.

[114] U. Parlitz, L. Junge, W. Lauterborn, and L. Kocarev. Experimental observation of phase synchronization. *Phys. Rev. E*, 54:2115, 1996.

[115] K. Pawelzik and H.G. Schuster. Generalized dimensions and entropies from a measured time series. *Phys. Rev. A*, 35:481, 1987.

[116] L. M. Pecora and T. L. Carroll. Synchronization in chaotic systems. *Phys. Rev. Lett.*, 64:821, 1990.

[117] L. M. Pecora, T. L. Carroll, G. A. Johnson, D. J. Mar, and J. F. Heagy. Fundamentals of synchronization in chaotic systems, concepts and applications. *Chaos*, 7:520, 1997.

[118] E. Pereda, R. Rial, A. Gamundi, and J. Gonzalez. Assessment of changing interdependencies between human electroencephalograms using non-linear methods. *Physica D*, 148:147, 2001.

[119] D. W. Peterman, M. Ye, and P. E. Wigen. High frequency synchronization of chaos. *Phys. Rev. Lett.*, 74:1740, 1995.

[120] A. S. Pikovsky. On the interaction of strange attractors. *Z. Phys. B: Condens Matter*, 55:149.

[121] A. S. Pikovsky. Phase synchronization of chaotic oscillations by a periodic external field. *Sov J Commun Technol Electron*, 30:85, 1985.

[122] A. S. Pikovsky, M. G. Rosenblum, and J. Kurths. Synchronization in a population of globally coupled chaotic oscillators. *Europhys. Lett.*, 34:165, 1996.

[123] A. S. Pikovsky, M. G. Rosenblum, and J. Kurths. *Synchronization.* A *universal concept in nonlinear sciences*. Cambridge Univ. Press, Cambridge, UK, 2001.

[124] W. H. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical recipes in Pascal:* T*he art of scientific computing*. Cambridge Univ. Press, Cambridge, UK, 1989.

[125] D. Prichard and J. Theiler. Generating surrogate data for time series with several simultaneously measured variables. *Phys. Rev. Lett.*, 73:951, 1994.

[126] K. Pyragas. Continuous control of chaos by self-controlling feedback. *Phys. Lett. A*, 170:421, 1992.

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

[128] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger. Performance of different synchronization measures in real data: A case study on electroencephalographic signals. *Phys. Rev. E*, 65:041903, 2002.

[129] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger. Reply to: Comment on 'performance of different synchronization measures in real data: A case study on electroencephalographic signals'. *Phys. Rev. E*, 67:063902, 2003.

[130] R. Quian Quiroga, T. Kreuz, and P. Grassberger. Event synchronization: A simple and fast method to measure synchronicity and time delay patterns. *Phys. Rev. E*, 66:041904, 2002.

[131] A. Renyi. *Probability Theory*. North Holland, Amsterdam, 1971.

[132] C. Rieke. *Nichtstationarität in dynamischen* S*ystemen*. PhD thesis, Department of Physics, University of Bonn, Germany, 2003.

[133] C. Rieke, F. Mormann, R. G. Andrzejak, T. Kreuz, P. K. A. David, C. E. Elger, and K. Lehnertz. Discerning nonstationarity from nonlinearity in seizure-free and pre-seizure EEG recordings from epilepsy patients. *IEEE Eng. Med. Biol.*, 50:634, 2003.

[134] C. Rieke, K. Sternickel, R. G. Andrzejak, C. E. Elger, P. David, and K. Lehnertz. Measuring nonstationarity by analyzing the loss of recurrence in dynamical systems. *Phys. Rev. Lett.*, 88:244102, 2002.

[135] Z. Rogowski, I. Gath, and E. Bental. On the prediction of epileptic seizures. *Biol. Cybern.*, 42:9, 1981.

[136] E. Rosa Jr., W. B. Pardo, C. M. Ticos, J. A. Walkenstein, and M. Monti. Phase synchronization of chaos in a plasma discharge tube. *Int. J. Bifurc. Chaos*, 10:2551, 2000.

[137] M. G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, and R. Mrowka. Identification of coupling direction: Application to cardiorespiratory interaction. *Phys Rev E*, 65:041909, 2002.

[138] M. G. Rosenblum and A. S. Pikovsky. Detecting direction of coupling in interacting oscillators. *Phys Rev E*, 64:045202, 2001.

[139] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths. Phase synchronization of chaotic oscillators. *Phys. Rev. Lett.*, 76(11):1804, 1996.

[140] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths. From phase to lag synchronization in coupled chaotic oscillators. *Phys. Rev. Lett.*, 78:4193, 1997.

[141] M. G. Rosenblum, A. S. Pikovsky, J. Kurths, C. Schaefer, and P. A. Tass. Phase synchronization: from theory to data analysis. In F. Moss and S. Gielen, editors, *Handbook of biological physics*, page 297. Elsevier Science, Amsterdam, 2001.

[142] A. G. Rossberg, K. Bartholom, and J. Timmer. Data driven optimal filtering for phase and frequency of noisy oscillations: Application to vortex flowmetering. *(submitted)*, 2003.

[143] O. E. Rössler. An equation for continuous chaos. *Phys. Lett. A*, 57:397, 1976.

[144] R. Roy and K. S. Thornburg. Experimental synchronization on chaotic lasers. *Phys. Rev. Lett.*, 72:2009, 1994.

[145] N. F. Rulkov, M. M. Sushchik, L.S. Tsimring, and H. D. I. Abarbanel. Generalized synchronization of chaos in directionally coupled chaotic systems. *Phys. Rev. E*, 51(2):980, 1995.

[146] N. F. Rulkov, L.S. Tsimring, and H. D. I. Abarbanel. Tracking unstable orbits in chaos using dissipative feedback control. *Phys. Rev. E*, 50:314, 1994.

[147] T. Sauer, J.A. Yorke, and M. Casdagli. Embeddology. *J. Stat. Phys.*, 65:579, 1991.

[148] S. J. Schiff, P. So, T. Chang, R. E. Burke, and T. Sauer. Detecting dynamical interdependence and generalized synchrony throug mutual prediction in a neural ensemble. *Phys. Rev. E*, 54:6708, 1996.

[149] K. Schindler, R. Wiest, M. Kollar, and F. Donati. EEG analysis with simulated neuronal cell models helps to detect pre-seizure changes. *Clin. Neurophysiol.*, 113:604, 2002.

[150] A. Schmitz. *Erkennung von Nichtlinearitäten und wechselseitigen Abhängigkeiten in Zeitreihen*. PhD thesis, Department of Physics, University of Wuppertal, Germany, 2000.

[151] A. Schmitz. Measuring statistical dependence and coupling of subsystems. *Phys. Rev. E*, 62:7508, 2000.

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

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

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

[155] T. Schreiber and A. Schmitz. Surrogate time series. *Physica D*, 142:346, 2000.

[156] J. Schumacher. Charakterisierung von Nichtlinearität in gekoppelten Modellsystemen und Zeitreihen hirnelektrischer Aktivität mit multivariaten stochastischen Kontrollsignalen. Master's thesis, Department of Physics, University of Bonn, Germany, 2002.

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

[158] D. A. Smirnov and R. G. Andrzejak. (personal communication).

[159] D. A. Smirnov and B. P. Bezruchko. Estimation of interaction strength and direction from short and noisy time series. *(submitted)*, 2003.

[160] C. J. Stam and B. W. van Dijk. Synchronization likelihood: An unbiased measure of generalized synchronization in multivariate data sets. *Physica D*, 163:236, 2002.

[161] F. Takens. Detecting strange attractors in turbulence. In D. A. Rand and L. S. Young, editors, *Dynamical Systems and Turbulence*, volume 898 of *Lecture Notes in Mathematics*, page 366. Springer-Verlag, Berlin, 1980.

[162] D. Y. Tang, R. Dykstra, M. W. Hamilton, and N. R. Heckenberg. Experimental evidence of frequency entrainment between coupled chaotic oscillations. *Phys. Rev. E*, 57(3):3649, 1998.

[163] P. A. Tass, M. G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H. J. Freund. Detection of n:m phase locking from noisy data: Application to magnetoencephalography. *Phys. Rev. Lett.*, 81(15):3291, 1998.

[164] C. A. Tassinari, G. Rubboli, and R. Michelucci. Reflex epilepsy. In M. Dam and L. Gram, editors, *Comprehensive Epileptology*, page 233. Raven Press, New York, 1990.

[165] J. Theiler. Spurious dimensions from correlation algorithms applied to limited time-series data. *Phys. Rev. A*, 34:2427, 1986.

[166] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer. Testing for nonlinearity in time series: The method of surrogate data. *Physica D*, 58:77, 1992.

[167] R. D. Traub and R. K. Wong. Cellular mechanism of neuronal synchronization in epilepsy. *Science*, 216:745, 1982.

[168] B. van der Pol. Forced oscillations in a circuit with non-linear resistance. *Phil. Mag.*, 3:64, 1927.

[169] B. van der Pol and J. van der Mark. The heartbeat considered as a relaxation oscillation, and an electrical model of the heart. *Phil. Mag.*, 6:763, 1928.

[170] F. J. Varela. Resonant cell assemblies: A new approach to cognitive functions and neuronal synchrony. *Biol. Res.*, 28:81, 1995.

[171] F. J. Varela, J. P. Lachaux, E. Rodriguez, and J. Martinerie. The brain web: Phase synchronization and large-scale integration. *Nature Rev. Neurosci.*, 2:229, 2001.

[172] R. V. V. Vidal, editor. *Applied Simulated Annealing*. Springer Verlag, Berlin, 1993.

[173] S. S. Viglione and G. O. Walsh. Epileptic seizure prediction. *Electroencephalogr. Clin. Neurophysiol.*, 39:435, 1975.

[174] J. Wegner. Zeitreihenanalyse synthetischer Daten sowie hirnelektrischer Aktivität mit Hilfe des größten Lyapunov-Exponenten. Master's thesis, Department of Physics, University of Bonn, Germany, 1998.

[175] H. Whitney. Differentiable manifolds. *Ann. Math.*, 37:645, 1936.

[176] P. Wolf. *Epileptic seizures and syndromes*. John Libbey, London, 1994.

[177] A. R. Wyler and A. A. Ward. Epileptic neurons. In J. S. Lockard and A. A. Ward, editors, *Epilepsy, A Window to Brain Mechanisms*. Raven Press, New York, 1992.

[178] M. Zaks, E. Park, M. Rosenblum, and J. Kurths. Alternating locking ratios in imperfect phase synchronization. *Phys. Rev. Lett.*, 82:4228, 1999.

[179] Z. Zheng and G. Hu. Generalized synchronization versus phase synchronization. *Phys. Rev. E*, 62:7882, 2000.

[180] Z. Zheng, X. Wang, and M. C. Cross. Transitions from partial to complete generalized synchronizations in bidirectionally coupled chaotic oscillators. *Phys. Rev. E*, 65:56211, 2002.

Danksagung

Zunächst möchte ich Prof. Dr. P. Grassberger, Prof. Dr. C. E. Elger und Priv. Doz. Dr. K. Lehnertz dafür danken, daß sie es mir ermöglicht haben, diese interdisziplinäre Doktorarbeit am John von Neumann Institute for Computing im Forschungszentrum Jülich und in der Klinik für Epileptologie der Universität Bonn durchzuführen. Besonders dankbar bin ich ihnen dafür, dass sie es mir ermöglicht haben, meine Forschungsergebnisse auf zahlreichen nationalen und internationalen Konferenzen zu präsentieren und dort Kontakte zu vielen interessanten Wissenschaftlern zu knüpfen.

Meinem Doktorvater Prof. Dr. P. Grassberger danke ich besonders für sein ständiges Interesse, die stets einladend offene Bürotür, sowie für die vielen lehrreichen Gespräche und Diskussionen, die zum Gelingen dieser Doktorarbeit maßgeblich beigetragen haben.

Mein besonderer Dank gilt weiterhin Priv. Doz. Dr. K. Lehnertz für die hervorragende Betreuung und Unterstützung während dieser Arbeit.

Den Kollegen am John von Neumann Institute for Computing und an der Klinik für Epileptologie danke ich für das ausgezeichnete Arbeitsklima sowie für viele interessante und weiterführende Diskussionen. Dabei sind auf der Bonner Seite Volker Hadamschek, Hannes Osterhage, Dr. Christoph Rieke und Robert Sowa und in Jülich Harald Stögbauer und Dr. Rodrigo Quian Quiroga besonders hervorzuheben. Bei Alexander Kraskov in Jülich möchte ich mich sehr für seine kollegiale Zusammenarbeit und ständige Hilfsbereitschaft bedanken. Für viele interessante Diskussionen bezüglich physikalischer und medizinischer Fragestellungen und die fortdauernde freundschaftliche Zusammenarbeit in der Bonner Epilepsieklinik bin ich Dr. Florian Mormann sehr dankbar. Sowohl in Bonn als auch in Jülich war schliesslich Dr. Ralph G. Andrzejak ein besonders ausdauernder und verlässlicher Kollege, Diskussionspartner und Freund.

Allen Mitarbeitern der Epilepsieklinik und des Forschungszentrums möchte ich meinen Dank für die ständige Hilfsbereitschaft bei allen auftretenden Problemen aussprechen. Dabei möchte ich Frau Helga Frank in Jülich besonders hervorheben.

Meinen Freunden, meinen Geschwistern und insbesondere meinen Eltern danke ich sehr für die kontinuierliche Unterstützung in allen Lebenslagen, ohne die dies alles gar nicht möglich gewesen wäre.

Bonn, im September 2003

Thomas Kreuz