TU
WIEN

TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

DISSERTATION

Novel Methods For Epileptic Onset Zone Detection

ausgeführt zum Zwecke der Erlangung des akademischen
Grades eines Doktors der Naturwissenschaften
unter der Leitung von

Em.o.Univ.-Prof. Dipl.-Ing. Dr. Manfred Deistler
E105 - Institut für Wirtschaftsmathematik

eingereicht an der Technischen Universität Wien
bei der Fakultät für Mathematik und Geoinformation

von
Dipl.-Ing. Christoph Flamm
Matrikelnummer: 0125607
Rauchenwarth Nr.18
2320 Rauchenwarth

_____   _____
Datum   Unterschrift

Deutsche Kurzfassung

In dieser Arbeit stellen wir zwei Methoden vor zur Detektion der Anfallsursprungszone von epileptischen Anfällen basierend auf ECoG Daten. Neben einer Einführung in Epilepsie stellen wir die Grundlagen für die vorgestellten Methoden bereit. Nach einer gründlichen Analyse der Methoden wird deren Anwendung auf wirkliche Daten präsentiert.

Epilepsie ist eine weit verbreitete neurologische Krankheit, die sich durch wiederkehrende unprovozierte Anfälle äußert. In dieser Arbeit werden wir uns mit einer bestimmten Epilepsieform beschäftigen, der Temporallappen-Epilep-sie. In diesem speziellen Fall gehen die epileptischen Anfälle von einer umschriebenen Region im Temporallappen des Gehirns aus, der sogenannten *Anfallsursprungszone* oder auch *Fokus*. Eine chirurgische Resektion der Anfallsursprungszone kann zu einer Heilung der Anfälle führen. Die genaue Lage der Anfallsursprungszone wird durch eine visuelle Analyse der EEG Daten, oder besser von ECoG Daten (die direkt vom Gehirn abgeleitet werden), ermittelt.
Die beiden vorgestellten Methoden sollen den Ärzten diese visuelle Analyse vereinfachen. Dadurch soll eine Optimierung des postoperativen Outcome erzielt werden.

Die erste vorgestellte Methode basiert auf der kausalen Analyse der ECoG Daten, diese Methode wird auch in Flamm et al. (2012a) beschrieben. In der vorliegenden Arbeit verwenden wir vorwiegend das Konzept der Granger Kausalität, die Grundidee dieses Konzepts basiert auf der Vohersagbarkeit der Daten. Die Neuheit und Besonderheit der vorgestellten Methode liegt in der Anwendung der Granger Kausalität auf Faktormodelle. Da die ECoG Daten viele gleiche Bewegungen zeigen, eignen sich Faktormodelle gut zu deren Beschreibung.
Wir geben eine Einführung in Faktormodelle und graphische Modelle sowie eine Einführung in Kausalität, siehe auch Flamm et al. (2012b). Basierend auf diesen mathematischen Gebieten werden wir eine kausale Untersuchungsmethode vorstellen, die wir *Einfluss-Analyse* nennen. Wir werden die Eigenschaften der Einfluss-Analyse diskutieren und sie auf die Daten eines Patienten anwenden. Die Anfallsursprungszone ist bei dieser Methode das Gebiet mit den einflussreichsten Elektroden.
Die mit dieser Methode errechnete Anfallsursprungszone deckt sich mit der An-

fallsursprungszone aus der visuellen Analyse der Ärzte.
Die Einfluss-Analyse ist der wichtigste Teil dieser Arbeit.

Die zweite beschriebene Methode ist praktischer als die Erste, weil sie auf der
Segmentierung der Daten basiert. Die zweite Methode wird auch in Graef et al.
(2012a) beschrieben.
Die ECoG Daten sind nicht stationär, das bedeutet die Eigenschaften der Daten
ändern sich über die Zeit. Wir unterteilen die Daten also in Segmente, in denen
sie gleichbleibende Eigenschaften haben. Wir verwenden dazu das sogenannte
Band-Leistungs-Maß, welches auf den physiologischen Frequenzbändern des men-
schlichen Gehirns basiert. Im nächsten Schritt klassifizieren wir den epileptischen
Charakter jedes einzelnen Segments. Rhythmische ϑ-Wellen sind charakteristisch
für Anfälle von Temporallappen-Epilepsie Patienten, deshalb werden Segmente,
die rhythmische ϑ-Aktivität zeigen, als epileptisch klassifiziert. Durch die Kom-
bination der beiden Schritte wird der Beginn der epileptischen Aktivität pro Kanal
festgestellt als der Beginn des ersten epileptischen Segments des jeweiligen Kanals.
Die Anfallsursprungszone enthält die Kanäle, die die erste epileptische Aktivität
zeigen.
Die Anwendung dieser Methode liefert ebenfalls Ergebnisse, die in guter Überein-
stimmung mit der visuellen Analyse der Ärzte sind.
Das Band-Leistungs-Maß ist der zweite wichtige Beitrag dieser Arbeit.

Zusammenfassend ist zu erwähnen, dass beide vorgestellten Methoden vielver-
sprechende Ergebnisse liefern. Die durch die Methoden errechneten Anfallsur-
sprungszonen stimmen mit der Anfallsursprungszone aus der visuellen Analyse
der Ärzte überein.

Abstract

In this thesis we present two mathematical methods for the detection of the epileptic seizure onset zone based on the analysis of ECoG data. We give a step by step introduction to epilepsy and provide the background information for our methods. The two methods are discussed in detail followed by an application to real world data.

Epilepsy is a common neurological disease which is characterized by recurring unprovoked seizures. A common sub-type of this disease is temporal lobe epilepsy. In this special case the epileptic seizures emanate from a circumscribed area in the temporal lobe, the so called *focus* or *seizure onset zone*. A surgical removal of the seizure onset zone should render the patient seizure-free. The clinicians determine the exact area of this focus by a visual inspection of EEG or preferably ECoG data (which are recorded directly from the cortex).
The main aim of the two presented methods is to assist the clinicians in this visual analysis in order to increase the chance of a seizure-free surgical outcome.

The first presented methodology is based on the casual analysis of the ECoG data, it is also described in Flamm et al. (2012a). The causality concept we will use for this analysis is Granger causality, which is based on the predictability of the data. The particularity of the proposed method is the application of Granger causality to factor models. This model class is used because it is well-fit for the ECoG data which show co-movement.
We give an introduction to factor models and graphical models as well as an introduction to causality, which is also described in Flamm et al. (2012b). Based on these mathematical topics we propose our methodology, called *influence analysis*, and thoroughly discuss its mathematical properties. Then we apply the influence analysis to real ECoG data of a patient. The seizure onset zone is found as the area comprising the most influential electrodes.
The resulting seizure onset zone matches the result of the visual analysis performed by clinical experts.
The influence analysis is the main contribution of this thesis.

The second presented methodology is more practical in nature as it is based on the segmentation of the data. The methodology is also described in Graef et al.

(2012a).

The ECoG data are non-stationary, that means the data's properties change over time. We partition the data into segments where the data have the same properties. For this purpose we use a measure based on the physiological frequency bands of the human brain, this measure is called *band power measure*. After this first step we classify each segment with respect to its epileptic character. Segments showing rhythmic ϑ-activity (which is characteristic for temporal lobe seizures) are classified as epileptic. Combining the segmentation and the classification we are able to derive the start of the epileptic activity per channel as the begin of the earliest epileptic segment. The seizure onset zone is found as the area comprising the channels showing the first epileptic activity.

The application of this methodology to the aforementioned ECoG data also yields a result that correlates very well with the visual inspection of the clinicians.

The band power measure is the second major contribution of this thesis.

Summing up, both presented methods show promising first results as they both detect the seizure onset zone matching the visual analysis of the clinicians.

Statutory Declaration

I hereby certify that all work presented in this PhD thesis is my own, no other than the sources and aids referred to were used. All parts which have been adopted either literally or in a general manner from other sources have been indicated accordingly.

The main contribution of this PhD thesis is the influence analysis, which is my own work.
The second contribution is the band power measure. The main idea, the conception and the application is my work. Andreas Graef suggested to use quadratic distances instead of the absolute distances in the construction of the band power measure. Furthermore he conceived the second rule for segment classification. Apart from these two points the band power measure is my own work.

Vienna, September 2012

Christoph Flamm

Acknowledgment

Without any doubt this PhD thesis was the hardest work in my live. Challenges in multiple forms arose during this time of my live. Therefore, I want to thank all persons who supported me during this period.

My greatest thanks go to my family. Without their caring love this work and my whole studies would not have been possible.

Of course I want to thank my advisor Prof. Manfred Deistler. He challenged me in ways I would never have imagined.

Another person I want to thank is Prof. Michael Eichler. He encouraged me when I needed it the most.

Furthermore I want to thank Dr. Susanne Piker for teaching me EEG and Prof. Christoph Baumgartner for his support at the hospital.

Finally I want to thank all my friends who supported me. I especially want to mention Andreas Graef and Alexander Filler, who are not only my colleagues but dear friends. Besides countless hours of work we shared countless jokes and hours of laughter.

We advise reading chapter 3 before this joke.
It will help, maybe.
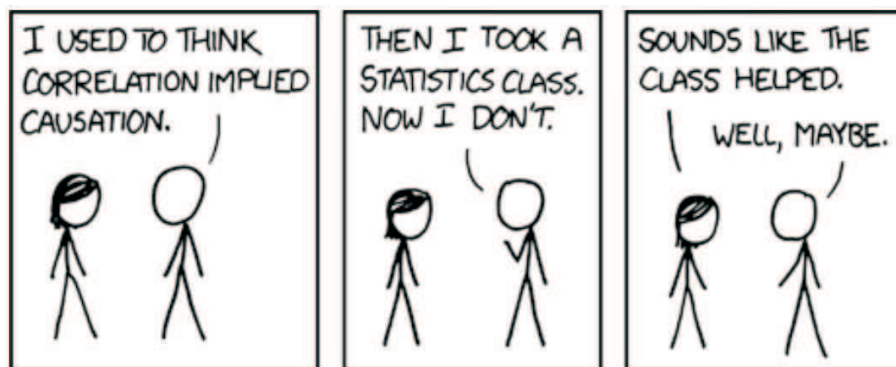
Figure 1: Webcomic entitled *correlation*. Source: http://xkcd.com/552/.

Contents

Chapter 1

Introduction

This thesis is an interdisciplinary contribution in the fields of statistics and epilepsy and should be regarded as such. The clear focus of this work is on statisticss as we will use it to analyze and model electrical brain activity. In this thesis we present two statistical methods for the detection of the epileptic seizure onset zone. The necessary medical and mathematical foundations will be provided for the reader.

Epilepsy is a common neurological disease which is characterized by recurring unprovoked seizures, see Baumgartner (2001). Seizures are characterized by abnormal synchronized brain activity in both hemispheres (generalized seizures) or in a circumscribed area (focal seizures). In this work we focus on temporal lobe epilepsy, where the seizures emanate from a circumscribed in the temporal lobe, the so called *seizure onset zone* or the seizure *focus*. If an anti-epileptic drug therapy cannot suppress the seizures, a resective epilepsy surgery removing the seizure onset zone is a valuable treatment option, see Schuele and Lüders (2008) and Wiebe et al. (2001). For this surgery the knowledge of the exact area of the seizure onset zone is essential.

The focus is normally determined by a visual inspection of the EEG data by clinical experts. If the scalp EEG does not provide sufficient information, subdural strip electrodes (directly placed on the cortex) are implanted. This method is called *electrocorticography(ECoG)*, see Pondal-Sordo et al. (2007). Like in the EEG case the seizure onset zone is determined by a visual inspection. This visual inspection is currently regarded as the gold standard of ECoG analysis, see Götz-Trabert et al. (2008) and Jenssen et al. (2011).

The aim of this thesis is to provide methods to aid the clinicians in the difficult visual analysis. In other words we present statistical methods for the detection of the seizure onset zone. These methods are meant as an objectivation and help for the clinicians.

In this thesis we present methods for seizure onset zone detection as well as their application to real ECoG data of a patient.

The first method will be based on the causal analysis of factor models, called influence analysis, see chapter 6. The second method will be based on the segmentation of the data and subsequent classification of the segments, see chapter 7.

This work is structured as follows:
In the remainder of this chapter we present the mathematical foundations.
In chapter 2 we introduce the reader to epilepsy. Especially we focus on temporal lobe epilepsy and its characteristics. We also present the ECoG data of a patient, which are analyzed in the later chapters.
The basics of causality are presented in chapter 3. We thoroughly discuss the concept of Granger causality in the univariate and the multivariate case. Furthermore, the most important dependence measures are discussed.
Graphical models are closely related to causality analysis and are presented in chapter 4. They are an easy way to analyze the inner structure of a stochastic process.
In chapter 5 we present factor models. The analysis of high-dimensional time series is often problematic due to the curse of dimensionality. Factor models are a useful tool for the analysis of such high dimensional systems as they reduce the dimension of the parameter spaces of the models used to describe these data. We also show how to use principal component analysis (PCA) for our purposes.
We present our first method for the detection of the epileptic seizure onset zone in chapter 6. The method is based on the causal analysis of factor models. All preliminaries can be found in the previous chapters. This chapter signifies the main contribution of this thesis as it merges the causal analysis, factor models and the neurophysiological aspects.
In chapter 7 we present our second method, which is based on the segmentation of the data. In this approach the data are segmented and subsequently classified. The combination of these two steps yields the seizure onset zone. This chapter is the second major contribution of this work.
This work is concluded by a thorough discussion and suggestions for the next steps.

1.1   Mathematical introduction

In chapter 2 we discuss the properties of EEG and ECoG data. These ECoG data are modeled as stochastic processes in this thesis.
In this thesis we distinguish between two different types of processes. In the classic Granger causal analysis (chapter 3), the discussion of causality measures (also chapter 3) and graphical modeling (chapter 4) we investigate n-dimensional stochastic processes $(y(t))_{t\in\mathbb{Z}}$ generated by n components. The mathematical properties of this kind of processes are discussed in this section.
In the factor model case we analyze n-dimensional stochastic processes $(x(t))_{t\in\mathbb{Z}}$, whose latent variable process $\chi(t)$ is generated by a small number $q < n$ of components, as is discussed in chapter 5 and 6.

We discuss the main differences of these two types of processes and the resulting influence on the causal analysis in section 6.1.

For notational purposes we simply write y when referring to the whole stochastic process $(y(t))_{t\in\mathbb{Z}}$, this also applies for all other processes. For a single realization/observation we will write $y(t)$.

For the classic Granger causal analysis, we consider an n-dimensional stochastic process $(y(t))_{t\in\mathbb{Z}}$, $y(t) : \Omega \to \mathbb{R}^n$, which is weakly stationary with mean zero. We refer to Hannan and Deistler (2012) and Brockwell and Davis (1991) for a mathematical introduction to time series analysis.

The covariance function of y is given by $\gamma(s) = \mathbb{E}\,y(t+s)y(t)'$. Although the covariance function in general does not contain the full information about the underlying stochastic process, the analysis presented here is based on the covariance only.

As is well known, see Rozanov (1967) and Hannan (1970), a stationary process has a representation of the form

$$y(t) = \int_{-\pi}^{\pi} e^{it\lambda}\,\mathrm{d}z(\lambda) \tag{1.1}$$

where $(z(\lambda)|\lambda \in [-\pi, \pi])$, $z(\lambda) : [-\pi, \pi] \to \mathbb{C}^n$ is a random process with orthogonal increments, which is uniquely defined by y.

The *spectral distribution function* $F(\lambda)$ of y is defined by $F(\lambda) = \mathbb{E}\,z(\lambda)z(\lambda)^*$, where $.^*$ denotes the conjugate transpose. For convenience we will use the notation $\mathrm{d}F(\lambda) = \mathbb{E}\,dz(\lambda)dz(\lambda)^*$. Note that $\mathrm{d}F(\lambda)$ describes the importance of a frequency band in terms of its contribution to the overall variance.

Under the assumption $\sum_{s=-\infty}^{\infty} \|\gamma(s)\| < \infty$ the spectral distribution function is absolutely continuous, and the *spectral density function* is defined as $f_{yy}(\lambda) = \mathrm{d}F(\lambda)/\mathrm{d}\lambda$ in the Radon Nykodym sense. In this case, there is a one-to-one relation between the covariance function and the spectral density:

$$\gamma(s) = \int_{-\pi}^{\pi} f_{yy}(\lambda)e^{i\lambda s}\mathrm{d}\lambda \tag{1.2}$$

$$f_{yy}(\lambda) = \frac{1}{2\pi}\sum_{s=-\infty}^{\infty} \gamma(s)e^{-i\lambda s}. \tag{1.3}$$

In this thesis we only consider linearly regular processes, see Rozanov (1967) and Hannan (1970), i.e. processes where the best linear least squares forecasts tend to zero if the forecast horizon tends to infinity. Linearly regular processes admit a Wold representation

$$y(t) = \sum_{m=0}^{\infty} K(m)\varepsilon(t - m) \tag{1.4}$$

where $\varepsilon(t)$ is n-dimensional white noise process, i.e. $\mathbb{E}\,\varepsilon(t) = 0$, $\mathbb{E}\,\varepsilon(s)\varepsilon(t)^* = \delta_{st}\Sigma$ and $K(m) \in \mathbb{R}^{n\times n}$, $\sum_{m=0}^{\infty} \|K(m)\|^2 < \infty$. Furthermore ε are the innovations of y, i.e. the one step ahead prediction errors of the best linear least squares

forecast of $y(t)$ given its past $\overline{y(t-1)}$, where (for stationary processes y) $\overline{y(t)} = closure(span(\{y(s)|s \le t\}))$ denotes the space spanned by the past and present of $y(t)$ in the Hilbert space of all square integrable random variables. Time t represents the present unless noted otherwise.

In addition we assume that the covariance matrix of the white noise Σ is non-singular.

An important special case of linearly regular processes are AR(∞) processes (AR stands for autoregressive).

For the remainder of the thesis we assume, that the spectral density of y is bounded uniformly above and below, i.e. there exists a real constant c such that

$$c^{-1}I_n \le f_{yy}(\lambda) \le cI_n{}^1 \text{ for all } \lambda \in [-\pi, \pi] \tag{1.5}$$

holds. According to Wiener and Masani (1957), this assumption yields that y has an *AR(∞) representation*

$$\sum_{m=0}^{\infty} A(m)y(t-m) = \varepsilon(t) \tag{1.6}$$

where $A(m) \in \mathbb{R}^{n \times n}$, $\sum_{m=0}^{\infty} \|A(m)\|^2 < \infty$ and $A(0) = I_n$ holds. The right-hand side ε is the same white noise process as in equation (1.4), see e.g. Geweke (1984). Additionally we assume that even $\sum_{m=0}^{\infty} \|A(m)\| < \infty$ in equation (1.6) holds, therefore we we also have $\sum_{m=0}^{\infty} \|K(m)\| < \infty$ in equation (1.4), see e.g. Brillinger (1981).

The interested reader may note, that assumption (1.5) ensures that Σ is non-singular.

We use z to denote the backshift operator on $\mathbb{Z}$: $z(y(t)|t \in \mathbb{Z}) = (y(t-1)|t \in \mathbb{Z})$, as well as a complex variable. Using this notation we may rewrite equation (1.6) in a shorter fashion

$$a(z)y(t) = \varepsilon(t), \tag{1.7}$$

where $a(z) = \sum_{m=0}^{\infty} A(m)z^m$ exists inside and on the unit circle. Because $A(0) = I_n$ and by using $\tilde{a}(z) = -\sum_{j=1}^{\infty} A(m)z^m$ we rewrite equation (1.7) as

$$y(t) = \tilde{a}(z)y(t) + \varepsilon(t). \tag{1.8}$$

Furthermore, we assume that the *stability condition* $\det a(z) \ne 0$ for $|z| \le 1$ holds.

---

[1]In this context $A < B$ means $B - A$ is a positive definite matrix.

There are two important points we want to highlight here:

First, representation (1.4) is a unique weakly stationary solution of (1.6), and it is called an AR(∞) process. For the sake of a simple notation we skip the (∞) sign henceforth.

Second, if assumption (1.5) holds for the whole process y, it also holds for all sub-processes, and therefore all sub-processes of y also have an AR representation.

Chapter 2

Epilepsy

Throughout this thesis we are interested in the temporal evolution of epileptic seizures. For the reader we provide a step by step introduction to all important terms and definitions of epilepsy.

2.1 Definition

According to Baumgartner (2001) *epilepsy* is a collective term for various diseases, which are characterized by recurring unprovoked seizures. It is important that the cause for these seizures remain between the seizures.

An *epileptic seizure* is the clinical manifestation of excessive hyper-synchronous activity of nerve cells. Depending on the involved brain areas and the seizure type, the clinical symptoms can range from impairment of consciousness to generalized convulsions.

Epilepsy is very common and has a prevalence of 0.7%, that means 0.7% of the general population suffers from epilepsy. The incidence, i.e. the number of new afflictions, is about 4-5 per 10.000 persons per year in industrial countries, see Hirtz et al. (2007) for further informations.

2.2 Electroencephalography

A commonly used and important tool in epilepsy research is electroencephalography. The combination of electro-(referring to electrical brain activity) encephalo-(referring to signals from the brain) and gram (or graphy, which means drawing) to the term *electroencephalography(EEG)* describes electrical neural activity of the brain. In other words, the EEG measures the electrical potential difference between two points on the scalp.

Richard Caton (1811-1926), an english scientist from Liverpool, used a galvanometer with two scalp electrodes to successfully record brain activity in the form of electrical signals in 1875.
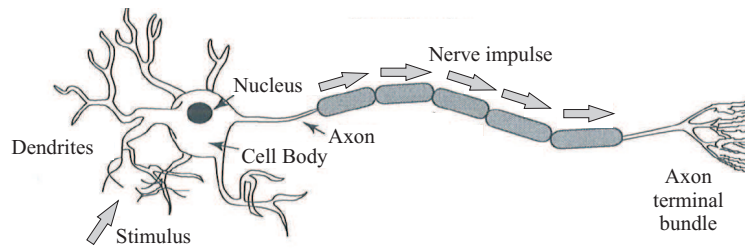
Figure 2.1: Schematic picture of a neuron, adopted from Atwood and MacKay (1989).

The discoverer of human EEG signals was Hans Berger (1873-1941), a german psychiatrist. Berger found the first proof for cortical generated potentials on July 24th, 1924 during a surgery. Berger (1929) is a fundamental work regarding EEG.

The presented historical facts are taken from Sanei and Chambers (2009) and Schneble (2003).

2.3   From cells to seizures

The central nervous system consists of a high number of nerve cells. Each of these nerve cell consists of a cell body, axons and dendrites, compare figure 2.1 for the scheme of a neuron. The axon is a long cylinder, which transmits electrical impulses. The dendrites are linked with dendrites or axons of other cells and transmit or receive impulses. The main activity of the central nervous system is related with the impulses traveling over the junctions (also called synapses) between dendrites and-or axons.

The information between two nerve cells, or in other words the temporal change in the membrane potential traveling along the axon, is called *action potential*. A potential of 60-70mV may be recorded under the cell membrane, and this potential changes with variations in the synaptic activities. The changes in these potentials are recorded and displayed by the EEG.

The presented biological facts are taken from Sanei and Chambers (2009).

As we mentioned before an epileptic seizure is the clinical manifestation of excessive hyper-synchronous activity of a continuum of nerve cells. This pathological synchronous activity starts at a small localized brain area and spreads to its surroundings, recruiting more and more cells in the process. This synchronization can affect both hemispheres of the brain (generalized seizure) or a circumscribed area in one hemisphere (focal seizure), see Baumgartner (2001).

2.4 Epilepsy surgery

About one third of the epilepsy patients suffer from therapy-resistant epilepsy, i.e. their epileptic seizures cannot be controlled by anti-epileptic drugs, see Engel (1996). A valuable treatment option for these patients is epilepsy surgery.

According to Baumgartner (2001) *epilepsy surgery* is a neurosurgical intervention in order to cure therapy-resistant epilepsy. The aim of this intervention is the removal of the epileptic tissue and the elimination of the seizures' cause.

The idea behind epileptic surgery is, that the seizures start from a localized area in the brain, also called the *seizure onset zone* or the *focus*, and spread from this area. The surgical extraction of this brain area should abolish the seizures.

Of course it is of highest importance to localize the seizure onset zone as exact as possible before the surgery. On one hand the seizures shall be stopped by the surgery, but on the other hand no neurological deficits shall be caused. Therefore, a thorough presurgical examination has to determine the extent of the epileptogenic tissue and the area of essential brain regions, like e.g. the motor cortex. We distinguish two types of examination methods: non-invasive (phase 1) and invasive (phase 2) ones, see Lüders (1992) and Engel (1996).

Non-invasive examination methods (phase 1) include:

- prolonged video-EEG-monitoring,

- structural imaging, e.g. magnetic resonance tomography (MRT),

- functional imaging, e.g. positron emission tomography (PET) and

- neuro-psychological tests, e.g. the Wada test.

If the non-invasive examination methods yield non-conclusive or inconsistent findings, invasive methods have to be used for a better localization of the seizure onset zone. These invasive methods are named after the types of intracranial electrodes used for the examination, the most common invasive examination methods (phase 2) are:

- epidural peg electrodes,

- subdural strip electrodes,

- subdural grid electrodes and

- implanted depth electrodes.

The examination with subdural strip electrodes is also called *electrocorticography(ECoG)*. It yields similar results as the EEG, but the electrodes are directly placed on the cortex. Therefore, the results from ECoG provide a better spatial resolution, see Behrens et al. (1994) or Zumsteg and Wieser (2000).

If phase 1 methods do not yield good results, ECoG is a good technique to localize the seizure onset zone. This localization is done by a visual inspection of the raw

ECoG data performed by the clinicians. Of course, this analysis is time-consuming and has to be conducted by experienced clinicians, but it is still regarded as gold standard, see Götz-Trabert et al. (2008) and Jenssen et al. (2011).

The first recorded epilepsy surgery was conducted by Wilder Penfield (1891-1976) in Montreal, see Schneble (2003). He removed a tumor from an adolescent's brain and thereby cured the patient from his daily seizures. Penfield closely worked together with Herbert Jasper (1906-1999), who was a specialist in the use of the EEG developed by Berger in order to localize the seizure onset zone.

2.5   Mathematical ECoG analysis

The aim of this thesis is an objectivation of the visual seizure onset zone localization as described in the previous section by mathematical methods. We want to aid the clinicians in the difficult localization task. There has recently been an increasing interest in the mathematical analysis of ECoG and EEG data, aiming at the quantification of the aforementioned synchronous activity in order to detect the seizure onset zone, see Kim et al. (2010) and Wilke et al. (2008).
We will present two mathematical methods for the localization of the seizure onset zone based on ECoG data. First, we present a methodology for the causal analysis of factor models, see chapter 6. Second, we propose a novel method based on the temporal delays of rhythmic ϑ-activity to detect the seizure onset zone, see chapter 7.

The mathematical analysis of EEG started early, when Berger and Dietsch applied Fourier analysis to the EEG data, see Dietsch (1932). The invention of the Fast Fourier Transformation algorithm promoted this field of research significantly.

2.6   Temporal lobe epilepsy

As the name indicates, *temporal lobe epilepsy (TLE)* refers to an epilepsy syndrome emanating from the temporal lobe, see figure 2.2. Depending on the exact area, we distinguish between *mesial* and *neocortical* temporal lobe epilepsies.
Mesial temporal lobe epilepsy (mTLE) is the most common epilepsy syndrome among therapy-resistant patients. Therefore, it is important to improve mathematical methods for the localization of the seizure onset zone for TLE patients. It is important to note that epilepsy surgery renders about 70-80% of the patients seizure free in the TLE case, see Baumgartner (2001).

Different epilepsy syndromes have different characteristic EEG correlates. These correlates are used for the localization of the epileptogenic zone. Temporal lobe epilepsy is typically distinguished by a rhythmic ϑ-activity (3-8 Hz) at the ictal (*ictal* means during a seizure) onset zone, see Foldvary et al. (2001). This means
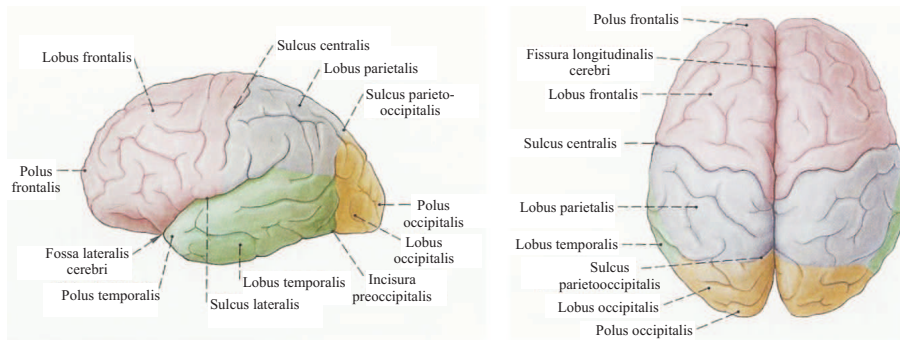
Figure 2.2: Schematic picture of the brain, adopted from Paulsen and Waschke (2010).

that ϑ- waves at the beginning of the seizure are a strong indicator for the seizure onset zone. Other indicators might also be possible.

For the remainder of this thesis we focus on ϑ-oscillations, which are especially important for TLE patients. The localization methods we will present in chapter 6 and 7 will exploit this rhythmic ϑ-activity.

2.7 Presentation of the data

In this thesis we will apply two proposed methods (see chapter 6 and 7) to ECoG data taken from a patient (male, 43 years) suffering from focal epilepsy. During the seizures he shows characteristic ϑ-waves, which are normally associated with temporal lobe epilepsy. We want the reader to get a grasp at the data and therefore present them here. [2]

The patient underwent a presurgical long-term video EEG monitoring at the Hospital Hietzing with Neurological Center Rosenhügel. Three subdural strip electrodes with a total of 25 channels were implanted. The electrode B1, which is far away from the (conjectured) seizure onset zone, was chosen as reference. Compare the *magnet resonance image (MRI)* in figure 2.3 for details. Because B1 is the reference electrode we have measurements of 24 channels available.

A Micromed® system with a sampling frequency of 1024Hz was used for the recording. Afterwards the data were preprocessed in Matlab®. First, the line interference was removed using a notch filer at 50Hz. Second, a high-pass filter at 1Hz got rid of physiologically irrelevant low-frequency contributions. Then the signals were downsampled to 128Hz after the application of a low-pass filter at 64Hz to

---

[2]Up to the date of this thesis the exact lay of the seizure onset zone of the patient is not clear. The clinicians initially conjectured the patient to have temporal lobe epilepsy due to the rhythmic ϑ-activity, but the visual analysis of the ECoG data suggests that his onset zone is rather temporo-occipital or occipital. The results of our analysis confirm this visual finding.
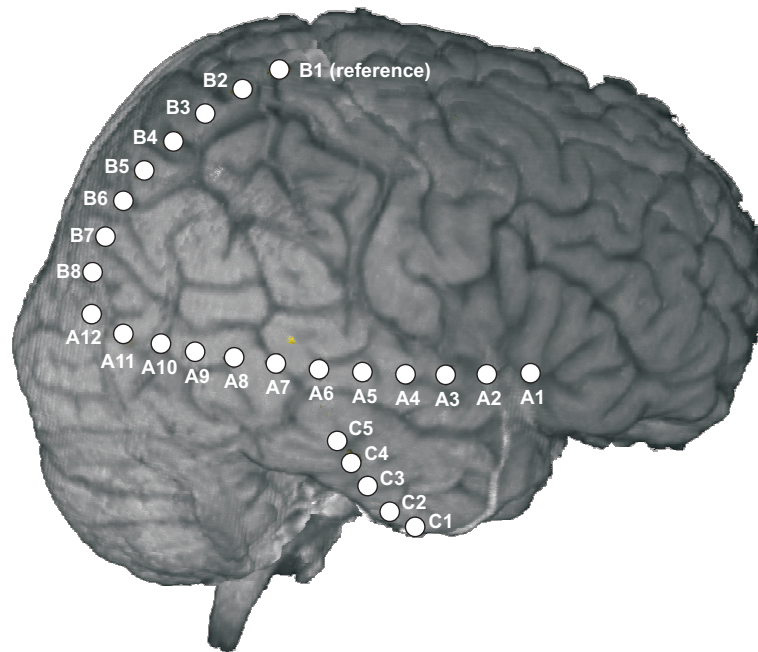
Figure 2.3: MRI picture of the brain including the electrode positions.

avoid aliasing.

During the examination time the patient had four epileptic seizures within 2 hours. We analyze the first three seizures because the data quality of the fourth seizure is bad. For a better understanding of the data, we will present the first 15 seconds of each seizure here. As we have already mentioned in section 2.6, we analyze the ϑ-activity in the initial seizure phase in order to draw conclusions regarding the seizure onset zone.

Seizure 1 is shown in figure 2.4, seizure 2 in figure 2.5 and seizure 3 in figure 2.6. The most important time points of the seizures, like the exact starting and ending times according to the visual analysis of the clinicians, are summarized in table 2.1.

Considering the visualization of the three seizures we see a common behavior. The initial stage of each seizure contains the following phases:

- First signs: The occurrence of *high-frequency oscillations(HFOs)* (75 Hz) signifies the beginning of the epileptic seizure. These HFOs have a very small amplitude and a short duration, which makes their visual detection difficult. Recently high-frequency oscillations have seen a lot of interest, because they also are used for the localization of the seizure onset zone.

- Paroxysmal activity: The blocks of high amplitude 30 Hz activity affect only some channels. Their influence on the epileptic seizures for this patient is
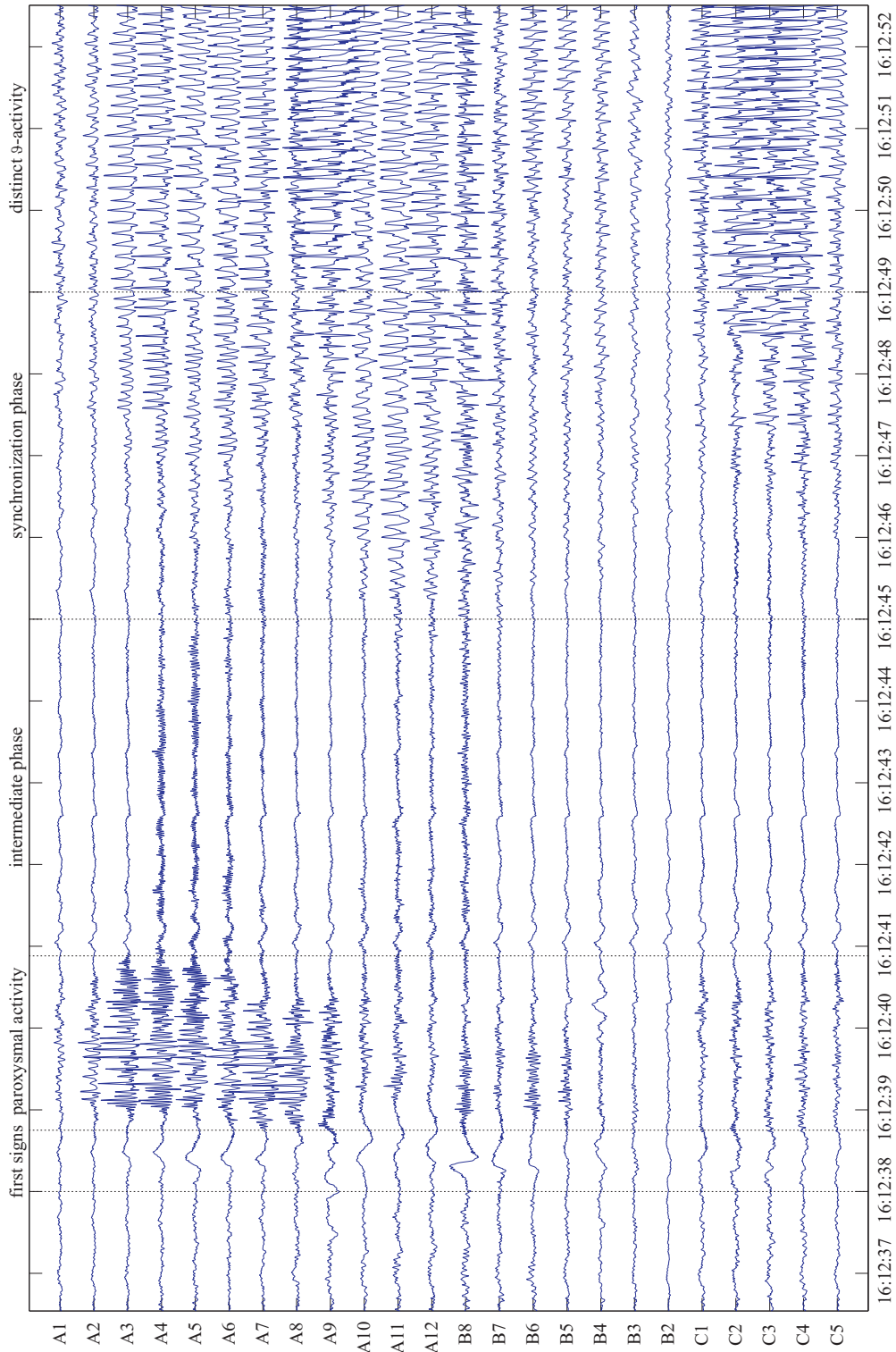
Figure 2.4: ECoG data of the patient, seizure 1, the first 15 seconds of the seizure are displayed.
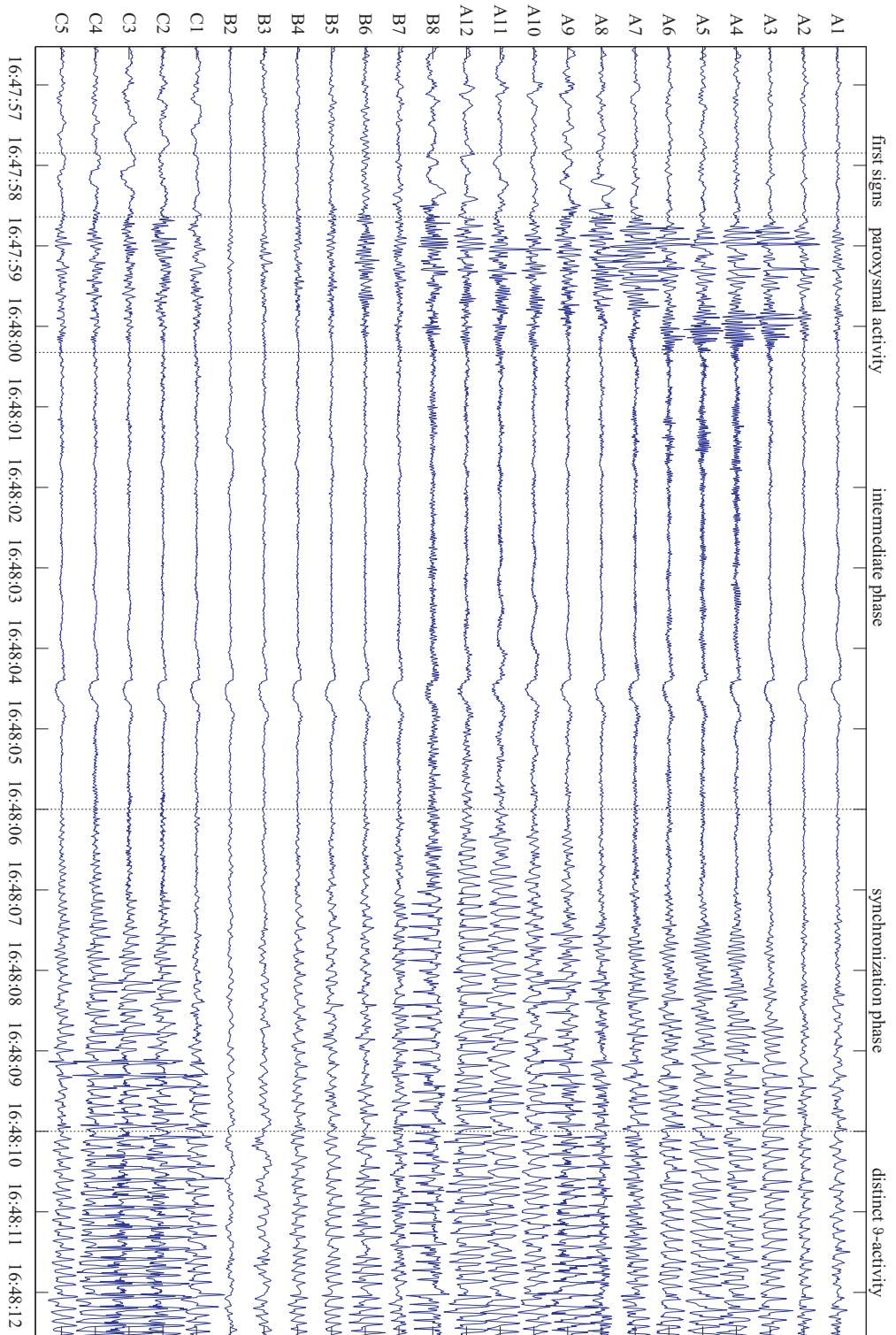
Figure 2.5: ECoG data of the patient, seizure 2, the first 15 seconds of the seizure are displayed.
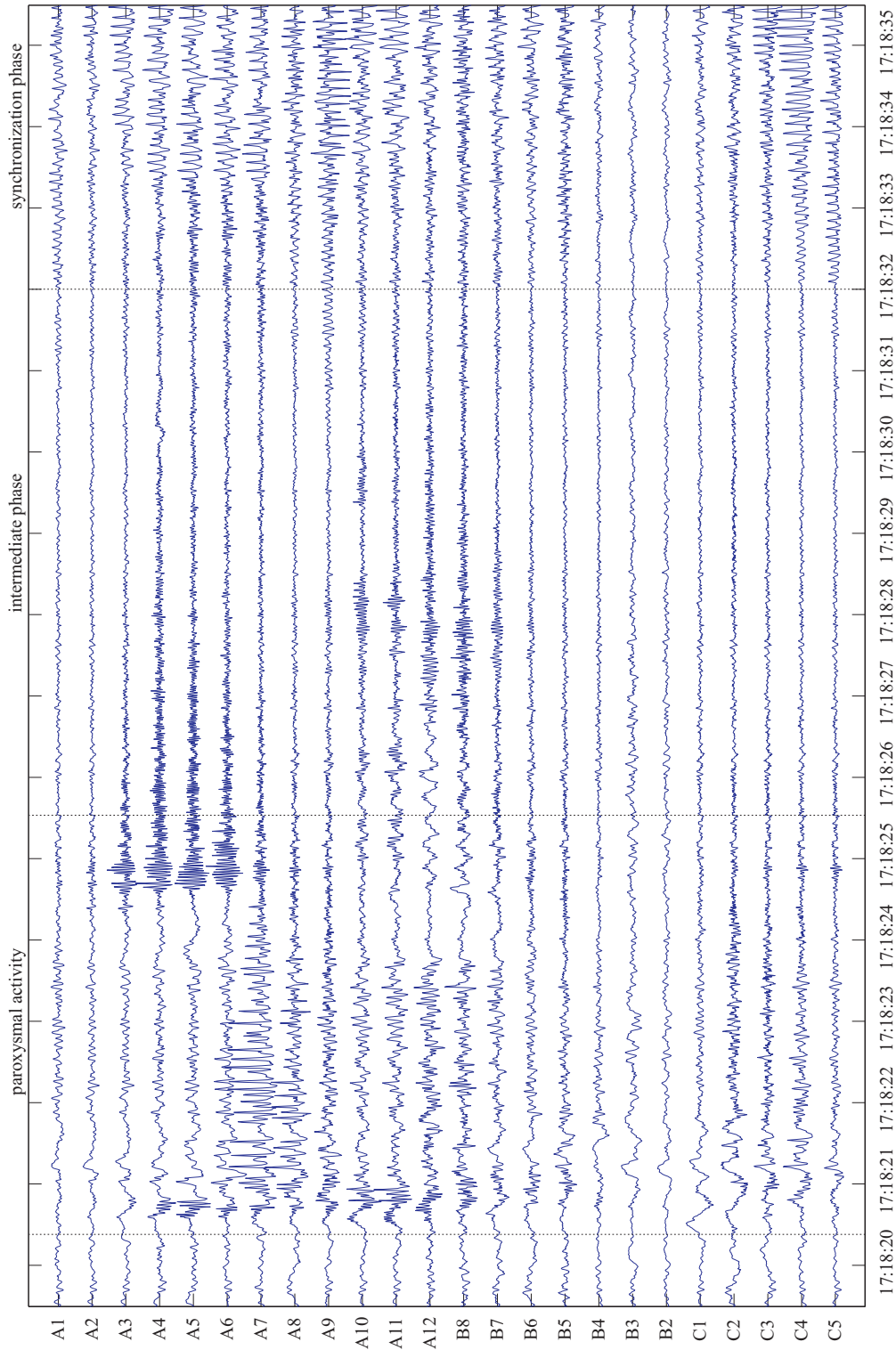
Figure 2.6: ECoG data of the patient, seizure 3, the first 15 seconds of the seizure are displayed.

not clear as they also occur during the seizures.

- Intermediate phase: This is the time period between the paroxysmal activity and the synchronization phase.

- Synchronization Phase: The beginning of the rhythmic ϑ-activity (3-8 Hz) signifies the beginning of this phase. The channels of the seizure onset zone are the first to show the distinct ϑ oscillations. As time progresses the ϑ-activity spreads to the other uninvolved channels, this behavior is called *seizure propagation*. This characteristic behavior leads to a synchronization of almost all observed channels.

  This is the most important phase for our analysis, because the first channels showing the distinct ϑ-rhythm are said to be the seizure onset zone. We will return to this point later in chapter 7.

- Distinct ϑ-activity: In this phase nearly all observed channels show a characteristic ϑ-activity. This phase has a very long duration compared to the other ones. The clinical signs of the epileptic seizures shortly start after all channels show the described distinct ϑ-rhythm.

Seizures 1 and 2 exactly show the described behavior, whereas the high frequency oscillations are not present in seizure 3.

The focus of this thesis lies on the analysis of the ϑ-activity, therefore the HFOs and the paroxysmal activity will not be considered in this work. For the interested reader we mention that HFOs also have a good localizing value, see e.g. Engel and da Silva (2012) and Zijlmans et al. (2011).

In chapter 7 we will analyze all three presented seizures and in chapter 6 we will investigate the first seizure in detail.
In figure 7.4 the initial stage of the first seizure is shown, this figure also includes the beginning of the epileptic activity per channel marked by three clinical experts. The results of this figure are summed up in table 6.1.

2.8   Problems in ECoG data analysis

The ECoG data presented in section 2.7 are biological real world data. Therefore, a lot of problems or rather challenges may occur in the data analysis.
In the works Graef (2008), Graef et al. (2008) and Schuster and Kalliauer (2009) seizure onset zone localization was done based on ECoG data analysis. Three main problem fields have been identified in these works:

- Stationarity: As aforementioned the ECoG data are biological data. We want to apply stationary methods, but the data are non-stationary. The main ideas

seizure	time	event
1	16:12:38	beginning
1	16:12:45	start of rhythmic ϑ-activity
1	16:12:50	clinical start
1	16:14:18	clinical ending
1	16:15:07	ending
2	16:47:58	beginning
2	16:48:06	start of rhythmic ϑ-activity
2	16:48:22	clinical start
2	16:49:46	clinical ending
2	16:50:33	ending
3	17:18:20	beginning
3	17:18:32	start of rhythmic ϑ-activity
3	17:18:55	clinical start
3	17:20:28	clinical ending
3	17:21:20	ending

Table 2.1: Clinical findings of the three seizures, by a visual inspection of the clinical experts.

to solve this problem include: Segmentation of the data into stationary segments, see e.g. Ombao et al. (2005) and Inouye et al. (1995), the use of a sliding window, see e.g. Bodenstein and Praetorius (1977) or Gath et al. (1992), or adaptive estimation, see e.g. An and Gu (1989).

- Dimension reduction: Due to the high correlation (caused by the spatial proximity of the electrodes and the synchronization of the ϑ-activity in ictal periods) there occur problems in the estimation of the whole 24 channel system. The main idea to solve this problem focuses on the reduction of the information, i.e. to extract the important information out of the data. Ideas to cope with this problem include: the use of factor models, see e.g. Deistler et al. (2010), or the selection of important channels for the analysis, see e.g. Graef et al. (2012b).

- Dependence (or causality) measures: The main idea behind the application of dependence measures in seizure onset zone localization is the following: Because the epileptic activity (i.e. the ϑ-rhythm) spreads from the seizure onset zone to the other channels, one could imagine that the seizure onset zone *influences* or *causes* the other channels. The literature offers a wide variety of such measures, see e.g. Baccala and Sameshima (2001) or Kaminski and Blinowska (1991), but the question is, which measure is best fit for seizure onset zone localization?

For more information on dependence measures see chapter 3.

As aforementioned we will present two methods for the detection of the seizure onset zone later on.

In chapter 6 we propose a methodology which is based on dimension reduction and and dependence measures. In particular we conduct a causal analysis of factor models.

A more practical approach is presented in chapter 7. The second proposed methodology is based on a segmentation method. For this segmentation we use a novel measure based on the physiological frequency bands of the brain.

Chapter 3

Causality

In this thesis we are interested in the detection of the epileptic seizure onset zone based on ECoG data. One possible approach for this detection is the use of causal analysis in the context of time series. Following this idea we analyze the dependencies between component processes of a multivariate stationary process.

It is important to distinguish between directed and undirected dependencies. Most of the undirected dependence relations are symmetric in nature, but for our purposes the directed relations are more important.
Furthermore, we distinguish between direct influences (between two component processes) and indirect influences (mediated by other components).

A common and renown concept of causality in the context of time series is Granger causality, which yields directed dependence relations. The idea behind Granger causality is based on the predictability of a stochastic process.
The concept of Granger causality as well as its implementation by regular AR-systems have been introduced in Granger (1969). The original definition focused on the relation between univariate processes, but multivariate extensions have also been proposed. For our purposes we use the multivariate definition of Granger causality introduced in Eichler (2007).

Besides the dependence structure of a stationary process itself, the strength of each individual dependence relation is often important. There exist various ideas for dependence measures in the literature, and we present the most important ones in this chapter.

We give the exact definitions of the dependence measures. These definitions are based on the (second order) population statistics. For the application to actual data statistical testing is necessary, but this is not in the focus of this thesis and will only be discussed briefly.

19

This chapter signifies the basis for the following parts of this thesis, where we will extensively draw information from here. All definitions in this chapter are based on the mathematical introduction in section 1.1.

We start with an historical introduction to causality in section 3.1. We present the concept of Granger causality in section 3.3 and its multivariate extension in section 3.4. Undirected dependence measures are presented in section 3.2 as well as directed ones in section 3.5. We briefly discuss the application of Granger causality to actual data in section 3.6.

3.1   History

As causality is one of the main concerns of this thesis, we present a brief historical summary concerning the evolution of the notion causality. The following informations are taken from the epilogue of Pearl (2000).

Every human has a basic thought what causality is or what the cause for an effect is. But despite this fact the notion of causality is shrouded in mystery and controversy when it comes to its definition for scientists and philosophers.
The urge to ask *why* is a very old habit of humankind. In the bible God asks Adam if he ate from the tree of knowledge. God wanted to know the facts and Adam brought forth an explanation of his deeds. The statement of this anecdote is clear: causal explanation is a man-made concept.

In ancient times only gods, humans and animals could cause things to happen, objects or events were not considered causes. This view of the world changed with the rise of engineering. With a large enough lever Archimedes would have moved the earth itself.
Archimedes considered the purpose of an event or object as its complete explanation. He even called it the final cause.
In the renaissance God's role as the final cause was taken over by human knowledge, and the notion of causality changed.

A drastic change in the way causality was considered and moreover a drastic change for science itself happened, when Galileo published his book *Discorsi* in 1638. He proposed two rules for proper scientific investigations:

1. Description first, explanation second. This means that the *how* precedes the *why*.

2. The description has to be done in mathematical language, namely equations.

This idea caused a lot of uproar among scientists and philosophers of this era, because why should nature be describable by mathematical equations. But we all

know that mathematical equations prevailed as a good description tool for natural observations and we all know what happened to Galileo.

Physicists embraced Galileo's first maxim and it led to a lot of empirical laws in physics, like Hooke's law or Joule's law. Philosophers discussed Galileo's rules thoroughly.

Hundred years after the Discorsi the philosopher David Hume argued that the *why* is not second to the *how*, but that the *why* is superfluous as it is subsumed by the *how*. He also writes about the flame and the heat as cause and effect, and infers the existence of the effect from that of the cause. Thus causal connections are solely a product of human observation.

He was not aware that his definition was flawed, because as we today know: *correlation does not imply causation*.

Bertrand Russell wrote in 1913 that causality is a relic of bygone ages. He complained about physics where the laws were bi-directional but the often discussed causality was not. The physicists did not pay attention to this discrepancy and had great success in splitting the atom.

In another discipline the need for the distinction between causal relations and other relationships rose, namely statistics. In 1888 Francis Galton measured the length of forearms and the size of people's heads. He related these two quantities strictly based on data and conjectured that the occurring co-relation was due to a common cause.

Karl Pearson, one of Galton's students, discarded the concept of cause and effect in 1911. He introduced so called contingency tables as the ultimate scientific statement. Thus Pearson denied the need for an independent concept of causality beyond correlation, he exterminated the notion causality from statistics before it had a chance to take root.

In the year 1936 Sir Ronald Fisher formulated the randomized experiment. This is the only scientifically proven method of testing causal relations from data and to this day, the one and only causal concept permitted in mainstream statistics.

Today scientists and even statisticians have an ambivalent relation to causality. We want to underline this by a quote of Terry Speed:

> *Considerations about causality should be treated as they have always been treated in statistics: preferably not at all but, if necessary, then with very great care.*

In Pearl (2000) a novel concept for the causal analysis of temporally non-ordered data was introduced, based on the idea of interventions.

In this thesis we are interested in the causal analysis of temporally ordered data, namely time series. In a certain sense this is easier than the non-ordered case. We use the causal concept proposed in Granger (1969) based on an idea in Wiener (1956). The two basic principles of this concept are that the cause precedes the effect, and that cause and effect are correlated. In other words, the knowledge of the cause helps the prediction of the effect.

We hope to have given the reader a small glimpse at causality's history and its problems.
Considering all the aforementioned problems the causal analysis of a multivariate system is not an easy task, especially not the causal analysis of brain structures during an epileptic seizure.

3.2   Undirected dependence measures

In this section we present the most prominent undirected measures for the dependence between component processes of an n-dimensional stationary processes.

3.2.1   Coherence

From the spectral representation of a stationary process (1.1) we obtain a measure of the strength of linear dependence in the frequency domain. Let y_i and y_j be univariate sub-processes of y, with the corresponding orthogonal increment processes $z_i(\lambda)$ and $z_j(\lambda)$ respectively.
The idea of the *coherence* is to measure the squared coefficient of correlation between $dz_i(\lambda)$ and $dz_j(\lambda)$

$$C_{ij}^2(\lambda) = \frac{|\mathbb{E}\{dz_i(\lambda)\overline{dz_j(\lambda)}\}|^2}{\mathbb{E}|dz_i(\lambda)|^2\mathbb{E}|dz_j(\lambda)|^2} = \frac{|f_{ij}(\lambda)|^2}{f_{ii}(\lambda)f_{jj}(\lambda)} \tag{3.1}$$

where f_{ij} is the (i, j)-element of the spectral density f_{yy} from equation (1.3). Thus, the coherence is a frequency specific measure for the dependence between y_i and y_j, which is bounded between 0 and 1. By construction it is a measure of the strength of dependence between the frequency weights $dz_i(\lambda)$ and $dz_j(\lambda)$.

The coherence is a very simple measure, and since $C_{ij}^2(\lambda)$ is obviously symmetric, it is not possible to detect a direction of influence from $C_{ij}^2(\lambda)$.
The second disadvantage is its incapability to distinguish between direct and indirect influences. This means if $C_{ij}^2(\lambda)$ indicates linear dependence between y_i and y_j, there might be a direct influence between the two channels or the dependence could be mediated via a third channel y_k (or via a path of other channels).

3.2.2 Partial spectral coherence (PSC)

As the ordinary coherence is not capable to distinguish between direct and indirect dependencies, we are searching for a measure which may do so. This leads us to the partial spectral coherence (PSC). The idea of the PSC is simple, we remove the influence of all other channels before considering the dependence between y_i and y_j.

The partial spectral coherence is constructed in the following way: In order to measure the dependence between y_i and y_j ($i \neq j$) after removing the influence of all other variables, $y_{V \setminus \{i,j\}} = (y_k | k \neq i, j)'$ say, we project y_i as well as y_j onto the Hilbertspace spanned by past, present and future $y_{V \setminus \{i,j\}}$ in the L^2 over the underlying probability space. This projection leads to the residuals η_i and η_j

$$\eta_i(t) := \eta_{i|V \setminus \{i,j\}}(t) = y_i(t) - \sum_{k=-\infty}^{\infty} D_i(k) y_{V \setminus \{i,j\}}(t-k) = y_i(t) - d_i(z) y_{V \setminus \{i,j\}}(t) \quad (3.2)$$

$$\eta_j(t) := \eta_{j|V \setminus \{i,j\}}(t) = y_j(t) - \sum_{k=-\infty}^{\infty} D_j(k) y_{V \setminus \{i,j\}}(t-k) = y_j(t) - d_j(z) y_{V \setminus \{i,j\}}(t)$$

where the filters $d_i(z)$ and $d_j(z)$ minimize the variance of the residuals. As a side note this definition can simply be extended from component processes to multivariate sub-processes of y as follows, for disjoint sub-sets $A, B, C \subset V$ we have

$$\eta_{A|C}(t) = y_A(t) - \sum_{k=-\infty}^{\infty} D_A(k) y_C(t-k) = y_A(t) - d_A(z) y_C(t) \quad (3.3)$$

$$\eta_{B|C}(t) = y_B(t) - \sum_{k=-\infty}^{\infty} D_B(k) y_C(t-k) = y_A(t) - d_B(z) y_C(t)$$

where the filters $d_A(z)$ and $d_B(z)$ minimize the variance of the residuals.
We resume our analysis and now look at the spectrum of the process $(\eta_i, \eta_j)'$ and let $f_{\eta_i \eta_j}$ denote the corresponding cross-spectrum. This cross-spectrum is a frequency specific measure for the dependence between between y_i and y_j given all $y_{V \setminus \{i,j\}}$. Scaling leads to the definition of the *partial spectral coherence* (PSC)

$$R^2_{ij|V \setminus \{i,j\}}(\lambda) = \frac{|f_{\eta_i \eta_j}(\lambda)|^2}{f_{\eta_i \eta_i}(\lambda) f_{\eta_j \eta_j}(\lambda)}. \quad (3.4)$$

As has been shown, see e.g. Dahlhaus (2000), there exists a more convenient way to compute the partial spectral coherence using the inverse of the spectral density $f_{yy}^{-1}(\lambda)$ of the original process y:

$$R^2_{ij|V \setminus \{i,j\}}(\lambda) = \frac{\left|(f_{yy}^{-1}(\lambda))_{ij}\right|^2}{(f_{yy}^{-1}(\lambda))_{ii}(f_{yy}^{-1}(\lambda))_{jj}} \quad (i \neq j) \quad (3.5)$$

where $(f_{yy}^{-1}(\lambda))_{ij}$ is the (i, j)-element of $f_{yy}^{-1}(\lambda)$.

The partial spectral coherence is bounded between 0 and 1. Obviously it is a symmetric measure, so no dependence direction may be inferred by its use. The clear advantage of the PSC is, that it measures the direct influence from y_i to y_j, because the indirect influences via other channels are filtered out. This property is important and useful.

The PSC also has as another advantage: Given actual data, instead of fitting a finite AR model, the partial spectral coherence can be estimated based on a non-parametric spectral estimator using equation (3.5). For this procedure a test with the null hypothesis $\mathcal{H}_0 : R^2_{ij|V\setminus\{ij\}}(\lambda) = 0$ has been described in Dahlhaus et al. (1997). The interested reader may be referred to Brillinger (1981) for a thorough theoretical discussion of the PSC.

3.3   Granger causality

There have been long and thorough discussions about causality throughout the last decades, see section 3.1, and there exist various ideas how to define causality.

The causality concept we will mainly use in this thesis is Granger causality, as introduced in Granger (1969). This section will heavily draw from this paper.

As already stated Granger causality is a *concept* of causality and we also present the framework for its application. The theoretical considerations are based on the second population moments. For the application of Granger causality to actual data statistical testing is necessary as will be discussed in section 3.6.

Although the application of the Granger causality concept only yields a statement whether one stationary process causes another or not, it is easy to derive a causality measure based on Granger causality as we will see in subsection 3.5.1.

The basic idea behind Granger causality was introduced in Wiener (1956): A variable (or time series) is called *causal* for another variable if the prediction of the second is improved by incorporating information about the first in the analysis. However, Wiener lacked the framework for the application of the concept in his work.

Before moving further we introduce some helpful notations in order to present the original definition of Granger causality. For a stationary process z, let $\overline{z(t)} = closure(span(\{z(s)|s \leq t\}))$ denote the space spanned by the past and present of z. Time t represents the present unless otherwise noted. Let y_1 and y_2 be two univariate stationary processes. Denote the optimal predictor of $y_2(t + 1)$ using the set of values $\overline{y_1(t)}$ by $P(t + 1; y_2|\overline{y_1})$. For instance, $P(t + 1; y_2|\overline{y_2})$ will denote the best linear predictor of $y_2(t + 1)$ using past and present of y_2. Let $\sigma^2(y_2|\overline{y_1})$ denote the

variance of the prediction error $\varepsilon(t + 1) = y_2(t + 1) - P(t + 1; y_2|\overline{y_1})$. Furthermore, we use $u(t)$ to denote all the information in the universe up to time t. The notation $u\backslash y_1$ refers to all information in the universe apart from the information of process y_1. Using these notations we may state the following definition.

Definition 3.1 (Original definition of causality according to Granger)**.** We say that y_1 is causing y_2, if

$$\sigma^2(y_2|\overline{u}) < \sigma^2(y_2|\overline{u\backslash y_1}). \tag{3.6}$$

Thus y_1 is causal for y_2 if the prediction from all information is better than the prediction from all information apart from y_1.

As a side note we want to mention the implicit assumption that the information represented in y_1 is not included in other parts of u. We will come back to this point later.

Granger also gave the definition of *feedback* and *causality lag* in Granger (1969), but they are of no importance to us. There is only one additional definition we want to list here.

Definition 3.2 (Definition of instantaneous causality according to Granger)**.** We say that y_1 is instantaneously causing y_2, if

$$\sigma^2(y_2|\overline{u}, y_1(t + 1)) < \sigma^2(y_2|\overline{u}) \tag{3.7}$$

In other words $y_2(t + 1)$ is better predicted when additionally using $y_1(t + 1)$ in the prediction.

Granger made some restrictions in order to reach a testable form for his original definition of causality.

First, the knowledge of all information in the universe u is unlikely, so we rather use a stationary process representing the relevant information d, where y_2 and y_1 are component processes of d. Thus the definition of causality is now relative to d.

Second, in practice it will usually not be possible to derive complete optimal (in the least squares sense) predictors, therefore we restrict ourselves to linear predictors only. In the Gaussian case the optimal predictors would be linear.

Third, it can be argued that the variance is not a proper criterion to measure the difference between the predictor $P(t + 1; y_2, .)$ and $y_2(t + 1)$. However, the variance seems to be a natural criterion, particularly in conjunction with linear predictors.

Taking all these restrictions into account, Granger called his definition of causality the *linear causality in mean with respect to d*.

Although Granger proposed a concept for the causality between two univariate processes with respect to a multivariate process, a lot of authors nowadays address Granger causality as the special bivariate case of his definition.

Definition 3.3 (Original definition of bivariate causality according to Granger)**.** We say that y_1 is causing y_2, denoted by $y_1 \to y_2$, if

$$\sigma^2(y_2|\overline{y_2}) > \sigma^2(y_2|\overline{y_2}, \overline{y_1}). \tag{3.8}$$

Thus y_1 is causal for y_2 if the knowledge of the present and past of y_1 improves the prediction of $y_2(t + 1)$, i.e. the variance of the prediction error is smaller when using the past and present of both y_2 and y_1 compared to using only the past and present of y_2 itself.

Criterion (3.8) can also be checked as follows. Let $y(t) = (y_1(t), y_2(t))'$ in the sense of section 1.1, we then analyze the causal influence from y_1 to y_2 by considering the joint AR representation (1.8) at time point $t + 1$, which exists according to the assumptions made,

$$\begin{pmatrix} y_1(t + 1) \\ y_2(t + 1) \end{pmatrix} = \underbrace{\tilde{a}(z) \begin{pmatrix} y_1(t + 1) \\ y_2(t + 1) \end{pmatrix}}_{} + \begin{pmatrix} \varepsilon_1(t + 1) \\ \varepsilon_2(t + 1) \end{pmatrix} \tag{3.9}$$

$$\begin{pmatrix} \sum_{m=1}^{\infty} A_{11}(m)y_1(t + 1 - m) + \sum_{m=1}^{\infty} A_{12}(m)y_2(t + 1 - m) \\ \sum_{m=1}^{\infty} A_{21}(m)y_1(t + 1 - m) + \sum_{m=1}^{\infty} A_{22}(m)y_2(t + 1 - m) \end{pmatrix}$$

where $\mathrm{Cov}((\varepsilon_1(t), \varepsilon_2(t))') = \Sigma$.

Because of the properties of AR representations, the lower line of equation (3.9) represents the orthogonal projection of $(y_1(t + 1), y_2(t + 1))'$ onto its past, which is its best linear predictor.

Therefore, we see that y_1 *is non-causal for* y_2 if $\tilde{a}_{21}(z) = 0$ (i.e. $A_{21}(m) = 0 \; \forall m$), which is equivalent to an unchanged prediction error when using $\overline{y_1(t)}$ additional to $\overline{y_2(t)}$. Otherwise y_1 is causal for y_2.

Note that this alternative version of criterion (3.8) is based on non-causality rather than causality. In the next section we will continue exploiting non-causality relations.

As aforementioned, the original bivariate definition of causality is often called Granger causality in modern literature. Granger's original definition is seldom used as there are no (explicit) requirements on u or respectively d.

Different authors proposed multivariate extensions to Granger's bivariate definition of causality, but these extensions are not necessarily consistent with the original definition. A non-exhaustive list of extensions and modifications is Granger (1980), Granger (1988), Hosoya (1977) and Florens and Mouchart (1985).

We will use the multivariate extension of Granger's classic definition introduced in Eichler (2007) as presented in the next section.

There have also been modifications to Granger's definition in order to analyze non-linear causality, see e.g. Freiwald et al. (1999) and Marinazzo et al. (2011).

We want to refer to Bressler and Seth (2011) for a good overview and additional information regarding Granger causality.

The interested reader may have noted, that there is a difference between the concept of Granger causality and the framework for its application. For the remainder of this thesis, we will address the application of the concept of Granger causality rather than the concept itself when we speak of Granger causality.

3.4 Multivariate Granger causality

In this section we present a multivariate extension of Granger's classic definition of causality which was proposed in Eichler (2007). We also present two additional equivalent criteria for this definition and discuss important facts regarding the multivariate extension.

For the remainder of this section let $(y(t))_{t \in \mathbb{Z}}$ be an n-dimensional stationary process satisfying the assumptions of section 1.1. Let $V = \{1, \dots, n\}$, we associate the elements of V with the component processes of y. We use the notation $y_V(t)$ for $y(t)$ to stress the fact that the elements of V correspond to the one-dimensional component processes of y. To refer to a sub-process corresponding to $S \subseteq V$ we write $y_S(t) = (y_s(t)|s \in S)'$, for a component process we simply write y_i.
Deducted from the assumptions in section 1.1 each sub-process y_S ($S \subseteq V$) of y_V satisfies the same assumptions as y_V itself, and therefore it also has an AR representation (corresponding to representation (1.7))

$$a^S(z)y_S(t) = \varepsilon_S(t), \tag{3.10}$$

where $a^S(z)$ is the respective AR power series and $\varepsilon_S(t)$ is the corresponding white noise process, $\text{Cov}(\varepsilon_S(t)) = \Sigma^S > 0$. Analogously to representation (1.8)) we get

$$y_S(t) = \tilde{a}^S(z)y_S(t) + \varepsilon_S(t). \tag{3.11}$$

Let A and B be disjoint subsets of $S \subseteq V$, then $a_{BA}^S(z) = (a_{ba}^S(z)|a \in A, b \in B)$ refers to the sub-matrices of $a^S(z)$ corresponding to the influence from y_A to y_B.
With these notations we may state the following definition.

Definition 3.4 (Definition of multivariate Granger non-causality according to Eichler). Let A and B be disjoint subsets of $S \subseteq V$. Then the process y_A is *Granger non-causal for y_B with respect to y_S* (denoted by $y_A \nrightarrow y_B|y_S$) if

$$\tilde{a}_{BA}^S(z) = 0, \tag{3.12}$$

($\Leftrightarrow A_{BA}^S(m) = 0 \ \forall m$) i.e. if the sub-matrices of the AR power series in representation (3.11) corresponding to the influence from y_A to y_B is zero. Of course, this is equivalent to $a_{BA}^S(z) = 0$.
In other words, the past and the present of $y_A(t)$ do not influence the linear prediction of $y_B(t + 1)$ in the AR representation of y_S.

Note that this definition is formulated in terms of non-causality, if y_A is *not* Granger non-causal for y_B with respect to y_S, we say that y_A *is causal for y_B with respect to* y_S, denoted by $y_A \rightarrow y_B | y_S$. If the relation (non-causality or causality) between y_A and y_B with respect to y_S has yet to be determined we will write $y_A \overset{?}{\rightarrow} y_B | y_S$.

This definition of causality is sometimes referred to as *conditional Granger causality*, because the causal effect of y_A on y_B conditioned on $y_{S \setminus \{A \cup B\}}$ is analyzed.

Furthermore, we want to note that the coefficients in $\tilde{a}^S$ represent the (orthogonal) projection coefficients of $y_S(t+1)$ onto $\overline{y_S(t)}$ in analogy to the bivariate case in equation (3.9).

In the following we present two equivalent criteria, which were also proposed in Eichler (2007).

Criterion (3.12) is based on the autoregressive coefficients and it is equivalent to

$$\det(\Sigma(y_B | \overline{y_S})) = \det(\Sigma(y_B | \overline{y_{S \setminus A}})), \tag{3.13}$$

where $\Sigma(y_B | \overline{y_S})$ denotes the covariance matrix of the prediction error when predicting $y_B(t+1)$ from $\overline{y_S(t)}$.

In other words y_A is Granger non-causal for y_B with respect to y_S if the determinant of the prediction error covariance matrix does not decrease when using y_A additional to $y_{S \setminus A}$ for the prediction of y_B, i.e. it is unchanged.

This criterion relates to Granger's original definition, which is also based on the prediction error.

In Barrett et al. (2010) the use of the trace instead of the determinant is discussed.

Before we move to a third equivalent criterion we introduce the following notation. For $A, B, C \subseteq V$ we remove the influence of $\overline{y_C(t)}$ from $y_A(t)$ and $y_B(t)$, compare equation (3.3),

$$\eta_{A | \overline{C}}(t) = y_A(t) - \sum_{k=0}^{\infty} D_A(k) y_C(t-k) = y_A(t) - d_A(z) y_C(t) \tag{3.14}$$

$$\eta_{B | \overline{C}}(t) = y_B(t) - \sum_{k=0}^{\infty} D_B(k) y_C(t-k) = y_A(t) - d_B(z) y_C(t)$$

where the filters $d_A(z)$ and $d_B(z)$ minimize the variance of the residuals. We use the following abbreviation if the residuals are uncorrelated

$$y_A(t) \perp y_B(t) | \overline{y_C(t)} :\Longleftrightarrow \mathbb{E}\, \eta_{A | \overline{C}}(t) \eta_{B | \overline{C}}(t)' = 0 \text{ for all } t \in \mathbb{Z}. \tag{3.15}$$

Furthermore, we use

$$y_A(t) \perp \overline{y_B(t)} | \overline{y_C(t)} :\Longleftrightarrow y_A(t) \perp z | \overline{y_C(t)} \text{ for all } z \in \overline{y_B(t)}, \text{ for all } t \in \mathbb{Z}. \tag{3.16}$$

A third equivalent criterion for Granger non-causality can be stated in terms of a non-correlation relation,

$$y_B(t+1) \perp \overline{y_A(t)} | \overline{y_{S \setminus A}(t)} \text{ for all } t \in \mathbb{Z}. \tag{3.17}$$

These non-correlation relations are often called conditional independence relations due to their origin from the iid graphical model case. In our setting this name is misleading because we only have non-correlation.

In criterion (3.17) we see why it is actually more convenient to work with non-causality relations, because they signify non-correlation relations or rather conditional independence relations. In chapter 4 we will further work with these conditional independence relations.

Of course, these three criteria are equivalent, but depending on the circumstances the effort for checking (in theoretical considerations or for statistical testing with actual data) may differ.

For the sake of completeness, we also present the following definition, which is the multivariate extension of Granger's instantaneous causality.

Definition 3.5 (Definition of contemporaneous non-correlation). The processes y_A and y_B are *contemporaneously uncorrelated with respect to y_S* (denoted by $y_A \not\sim y_B | y_S$) if

$$\Sigma_{AB}^S = 0,$$

i.e. if the sub-matrix of Σ^S corresponding to the sets A and B (in analogy to the definition of $a_{BA}^S(z)$) is zero. This criterion can also be formulated in terms of a conditional independence relation

$$y_A(t+1) \perp y_B(t+1) | \overline{y_S(t)} \text{ for all } t \in \mathbb{Z}. \tag{3.18}$$

In our considerations we will rather focus on the non-causality relations than the (contemporaneous) non-correlation relations.

In a certain sense the (contemporaneous) non-correlation relations are independence relations, where the temporal sampling resolution is not sufficient to detect the direction of the independence.

We conclude our thoughts concerning (contemporaneous) non-correlation relations by the following fact. In the context of Granger causality analysis the leading AR coefficient $A(0)$ is always fixed as the identity. Therefore the covariance matrix of the errors is also fixed, but there exist other causality definitions where the leading AR coefficient does not need to be the identity, see Faes and Nollo (2010).

There are some important points we want to highlight in connection to the definition of Granger non-causality at this point.

The causality relation between y_A to y_B with respect to y_S of course depends on S. We emphasize this, because the causality relations could change for changing

subsets. For example, if $y_A \rightarrow y_B|y_S$ holds, $y_A \nrightarrow y_B|y_{S'}$ could hold for $S' \subset S$. At first this might sound counter-intuitive, but at second glance it is clear. We will return to this problem in chapter 4. Granger also stressed the importance of the information series d in his original definition, because a change of d to d' could change the causal relations.

For a complete causal investigation of the stationary process y, one has to consider all independence relations for all arbitrary sub-processes of the form $y_A \overset{?}{\rightarrow} y_B|y_S$ (and all (contemporaneous) non-correlation relations). For a large number n this can be a huge effort, but a short and elegant way to gain more insights in the structure of the process is graphical modeling, which will be discussed in chapter 4.

However, we restrict our analysis to one-dimensional sets A and B, i.e. we only investigate the causality between two component processes given the information of the whole process. In other words, we concentrate on relations of the form $y_i \overset{?}{\rightarrow} y_j|y_V$. This is reasonable, because these relations are the basis for graphical modeling, and therefore they can be used to infer relations for arbitrary sub-processes, see chapter 4.

The analysis of the causality between two univariate component processes given the whole process is straight forward using the three aforementioned criteria. However, for a better understanding we want to state the causality definition for this restricted case. Using the notations we introduced above we have the following.

Definition 3.6 (Definition of Granger non-causality)**.** Let y_i and y_j ($i \neq j$) be two component processes of y_V. Then the process y_i *is Granger non-causal for y_j with respect to y_V* (denoted by $y_i \nrightarrow y_j|y_V$) if one of the three following equivalent criteria holds

$$\tilde{a}_{ji}(z) = 0, \tag{3.19}$$

$$\sigma^2(y_j|\overline{y_V}) = \sigma^2(y_j|\overline{y_{V\setminus\{i\}}}) \text{ or} \tag{3.20}$$

$$y_j(t+1) \perp \overline{y_i(t)|y_{V\setminus\{i\}}(t)} \text{ for all } t \in \mathbb{Z}. \tag{3.21}$$

The processes y_i and y_j are *contemporaneously uncorrelated with respect to y_V* (denoted by $y_i \nsim y_j|y_V$) if one the two following equivalent criteria holds

$$\Sigma_{ij} = 0 \tag{3.22}$$

$$y_i(t+1) \perp y_j(t+1)|\overline{y_V(t)} \text{ for all } t \in \mathbb{Z}. \tag{3.23}$$

Of course all aforementioned remarks concerning the general definition also apply to the restricted definition. Criterion (3.19) is based on the AR coefficients, criterion (3.20) focuses on the variance of the prediction error and criterion (3.21) signifies the respective conditional independence relation.

The interested reader surely has noted one important fact. The definition of non-causality according to criterion (3.20) equals non-causality in the sense of Granger's classic definition if the relevant information d equals y_V. In other words, if the important information comprises all observed processes the two definitions of causality are equal. Thus Eichler's multivariate definition of (non-)causality is a reasonable extension of Granger's classic definition.

The above definitions and criteria for multivariate Granger causality are based on regular AR systems. This means that the number of observations, i.e. the dimension of y, equals the number of driving components, i.e. the dimension of the white noise process ε (with a regular covariance matrix). Therefore, Granger causality analysis should be used for data satisfying this requirement. Typically observed data can be modelled by a regular AR system, but in practical applications this sometimes leads to ill-conditioned covariance matrices. This problem often occurs with high-dimensional co-moving data such as EEG (or ECoG) data. In chapter 5 and 6 we will extract the co-moving parts of such data for further investigation. These co-moving parts are normally generated by a far smaller number of components than the number of observations. A normal Granger causal analysis according to the definitions presented in this section would yield misleading results for such data.

For the remainder of this thesis, we will use the term Granger causality to refer to the definition of multivariate Granger causality based on criteria (3.12), (3.13) and (3.17). Of course the restricted case according to criteria (3.19), (3.20) and (3.21) is included in this notion.

3.5   Directed dependence measures

In the last section we encountered the definition of Granger causality, its application only yields a (non-)causality relation. Suppose y_i is causal for y_j ($y_i \rightarrow y_j|y_V$), but this relation does not give any information about the strength of the causality. We want to know how y_j changes for changes of y_i (in our linear setting). It is a naturally arising question how to measure the strength of the influence between two processes.
Therefore, we present the most common dependence measures from neuroscience literature in this section. In Blinowska (2011) a good overview over the most common dependence measures can be found.

The measures related to Granger causality could be called causality measures, but the more common term, also including the measures not based on Granger causality, is simply *dependence measures*. Note that other definitions of causality are viable too, so other measures could also be called causality measures. But in this

thesis we focus on Granger causality as our main concept of causality.

The measures gathered in this section are mainly from neuroscience literature, and sometimes we will refer to single component processes y_i of y_V as channels.
When we speak of the dependence (or influence) from y_i to y_j, y_i is called the *source channel* and y_j is called the *target channel*. In this context we speak of *all* target channels, when we refer to the channels y_m ($m \neq i$) with $a_{mi}(z) \neq 0$, i.e. we mean all channels influenced by y_i. Respectively we mean all channels y_m ($m \neq j$) with $a_{jm}(z) \neq 0$ when we speak of *all* source channels. This is a sloppy but intuitive formulation.

Our presentation of these measures is theoretical in nature. The application of these dependence measures to actual data requires statistical testing, but this is not in the focus of this section.

3.5.1 Measure of conditional linear dependence

The first measure we want to present is directly related to Granger causality. It was originally proposed in Geweke (1984), its predecessor, the measure of linear dependence, was proposed in Geweke (1982).

For the motivation of this measure we recall the definition of non-causality according to criterion (3.20), which reads $\sigma^2(y_j|\overline{y_V}) = \sigma^2(y_j|\overline{y_{V\setminus\{i\}}})$. We then simply construct the log ratio of these two terms.
The *measure of conditional linear dependence from y_i to y_j with respect to y_V* is defined as

$$\mathcal{F}_{i \to j|V\setminus\{i,j\}} = \ln \frac{\sigma^2(y_j|\overline{y_{V\setminus\{i\}}})}{\sigma^2(y_j|\overline{y_V})}, \tag{3.24}$$

where $\sigma^2(y_j|\overline{y_V})$ denotes the variance of the prediction error when predicting $y_j(t+1)$ from $\overline{y_V(t)}$.

Although it is called dependence measure, it actually measures Granger causality. The measure (3.24) equals zero in case of non-causality, and yields positive values in the case of causality. Further details regarding the properties of this measure can be found in Geweke (1984) and Barrett et al. (2010).
The application of the measure of conditional linear dependence to actual data and the corresponding statistical testing is briefly described in Geweke (1984).
The measure of conditional linear dependence is sometimes called *Granger causality index* (GCI), see e.g. Winterhalder et al. (2005).

3.5.2 Directed transfer function (DTF)

The second directed dependence measure we want to present is the directed transfer function as proposed in Kaminski and Blinowska (1991). Using the moving

average coefficients of equation (1.4) we define the *transfer function* as $k(z) = \sum_{m=0}^{\infty} K(m)z^m$ and note that $k(z) = a(z)^{-1}$ holds under the assumptions of section 1.1.

The transfer function itself measures the influence from the driving white noise process to the observed variables, the DTF is a normalized variant of this influence. The question is, how do the influences from the transfer function translate to the dependencies between component processes of y_V?

The *directed transfer function* (DTF) from y_i to y_j (with respect to y_V) is defined as

$$\gamma_{ji}^2(\lambda) = \frac{|k_{ji}(\lambda)|^2}{\sum_{m=1}^{n} |k_{jm}(\lambda)|^2} \tag{3.25}$$

where $k(\lambda)$ is a short notation for $k(e^{-i\lambda})$ (here we used $i = \sqrt{-1}$) and $k_{ji}(z)$ is the (j,i) element of the transfer function $k(z)$.

The directed transfer function is a directed dependence measure. Its denominator provides a normalization, and therefore the DTF is bounded by 0 and 1. This normalization is with respect to all source channels.

The nominator of the DTF measures the total information flow from y_i to y_j. This fact can be seen by expanding $k(z) = a(z)^{-1}$ as a geometric series using the notations introduced in section 1.1

$$k(z) = a(z)^{-1} = (I - \tilde{a}(z))^{-1} = \sum_{m=0}^{\infty} \tilde{a}(z)^m = I + \tilde{a}(z) + \tilde{a}(z)^2 + \dots .$$

Considering the off diagonal elements (i.e. $i \neq j$) we obtain

$$k(z)_{ji} = (a(z)^{-1})_{ji} = \tilde{a}(z)_{ji} + \sum_m \tilde{a}(z)_{jm}\tilde{a}(z)_{mi} + \sum_{m,\ell} \tilde{a}(z)_{jm}\tilde{a}(z)_{m\ell}\tilde{a}(z)_{\ell i} + \dots .$$

This shows that the nominator is the sum of the direct and all indirect information flows from y_i to y_j.

The DTF measures the total information flow between two components in a multivariate system including direct and indirect influences. Therefore, no conclusions may be drawn concerning the direct pathways of the information propagation. So the DTF is not useful in cases when we want to find the causal structure of a multivariate system.

Following this line of thought, it was shown in Eichler (2006c) that the directed transfer function does not detect Granger causal relations except in the bivariate case.

The properties of the directed transfer function are reviewed in the original paper Kaminski and Blinowska (1991) and a thorough discussion can be found in Kaminski et al. (2001). The DTF is an often used and common dependence measure in neuroscience literature.

3.5.3   Direct directed transfer function (dDTF)

In order to overcome the problem of indirect information flows that we encountered when discussing the directed transfer function, Korzeniewska et al. proposed a combination measure of the directed transfer function (DTF) and the partial spectral coherence (PSC).

The *direct directed transfer function* (dDTF) from y_i to y_j (with respect to y_V) was proposed in Korzeniewska et al. (2003) and is defined as

$$\delta^2_{ji}(\lambda) = \gamma^2_{ji}(\lambda) R^2_{ji|V\setminus\{i,j\}}(\lambda). \tag{3.26}$$

The DTF is used to identify the direction of the information flow, and the PSC is used to filter out the indirect flows, so the directed direct information flows are the only remaining ones.

As has been pointed out in Eichler (2006a) the statistical properties of the dDTF have not been investigated so far and an analysis of actual data based on the dDTF could detect wrong relationships.

3.5.4   Partial directed coherence (PDC)

Another simple and yet efficient idea is to construct a frequency specific dependence measure directly based on the AR coefficients. The *partial directed coherence* (PDC) from y_i to y_j (with respect to y_V) was introduced in Baccala and Sameshima (2001) and is defined as

$$\pi^2_{ji}(\lambda) = \frac{|\tilde{a}_{ji}(\lambda)|^2}{\sum_{m=1}^n |\tilde{a}_{mi}(\lambda)|^2}, \tag{3.27}$$

where $\tilde{a}(\lambda)$ is a short notation for $\tilde{a}(e^{-i\lambda})$ (here we used $i = \sqrt{-1}$) and $\tilde{a}_{ji}(z)$ is the (j, i) element of the AR power series $\tilde{a}(z)$.

The PDC can be seen as the direct information flow from y_i to y_j normalized with respect to all target channels. Due to the normalization the partial directed coherence is bounded between 0 and 1.

The careful reader may have noted, that this normalization is different than the one of the DTF. The DTF is normalized with respect to all source channels for one target channel. This is reasonable, because the original meaning of k_{ji} is the influence from ε_i to the target channel y_j. In contrast, the PDC is normalized with respect to all target channels for one source channel, so the partial directed coherence will mark the strongest influences departing from a single channel. Of course other normalizations would be conceivable too.

The *partial* as part of the PDC's name stems from the derivation of the measure, where Baccala and Sameshima factorized the partial spectral coherence and skipped some components. The word partial normally refers to the removal of the

influences of all other components when considering the influence from y_i to y_j, or in other words the direct influence from y_i to y_j. As the AR coefficients themselves signify these direct influences and the PDC is based on these coefficients, the partial directed coherence only shows direct effects.

Obviously the PDC is a directed measure. The partial directed coherence is widely used in neuroscience literature and a renown dependence measure. Its properties are discussed in the original paper Baccala and Sameshima (2001), and a more thorough discussion can be found in Schelter et al. (2005). There exist a lot of extensions of the classic PDC, see e.g. Baccala et al. (2007) and Faes and Nollo (2010).

The advantage of the partial directed coherence is the clear interpretation as the normalized direct information flow. Furthermore, it has a clear connection to Granger causality. Non-influence in terms of PDC is equivalent to Granger non-causality as can be easily seen,

$$\pi_{ji}^2(\lambda) = 0 \ \forall \lambda \in [0, 2\pi] \iff \tilde{a}_{ji}(\lambda) = 0 \ \forall \lambda \in [0, 2\pi] \iff$$
$$\iff A_{ji}(m) = 0 \ \forall m \iff \tilde{a}_{ji}(z) = 0.$$

Thus dependence in the PDC sense is equivalent to Granger causality. Of course, this is a theoretical connection.

For the practical implementation of the dependence analysis using the PDC statistical testing is necessary. For further information on this topic we refer to Schelter et al. (2005).

3.5.5 Generalized partial directed coherence (gPDC)

A disadvantage of the PDC is, that it is not scale invariant, meaning that it is not invariant under different choices of the unit of measurement. To overcome this problem an extension of the partial directed coherence was introduced in Baccala et al. (2007).
The *generalized partial directed coherence* (gPDC) from y_i to y_j (with respect to y_V) is defined as

$$\tilde{\pi}_{ji}^2(\lambda) = \frac{\Sigma_{jj}^{-1} |\tilde{a}_{ji}(\lambda)|^2}{\sum_{m=1}^{n} \Sigma_{mm}^{-1} |\tilde{a}_{mi}(\lambda)|^2} \tag{3.28}$$

where Σ_{jj} is the (j, j)-component of the error covariance matrix.
This modification turns out to be more robust than the PDC when processing actual data.

The PDC is a widely used dependence measure and there exist a lot of modifications in literature, the gPDC is just one of its various extensions. There is only one

extension beside the gPDC we want to mention here, it is the so called renormal-ized partial directed coherence as introduced in Schelter et al. (2009).

3.5.6 Extrinsic-to-intrinsic-power ratio (EIPR)

In the discussion of the partial directed coherence we mentioned, that in the cal-culation of the PDC from y_i to y_j also the other channels $y_{V \setminus \{i,j\}}$ are involved via the normalization. In Hartmann et al. (2008) a novel dependence measure, called extrinsic-to-intrinsic-power ratio (EIPR), was proposed, where only the source and the target channel are involved in the calculation of the measure.

For the derivation of the EIPR we rewrite the line of equation (1.8) corresponding to y_j in the following way

$$
\begin{aligned}
y_j(t) &= \sum_{i=1}^{n} \tilde{a}_{ji}(z) y_i(t) + \varepsilon_j(t) = \underbrace{\tilde{a}_{jj}(z) y_j(t)}_{\mu_{jj}(t)} + \sum_{i \in V \setminus \{j\}} \underbrace{\tilde{a}_{ji}(z) y_i(t)}_{\mu_{ji}(t)} + \varepsilon_j(t) = \\
&= \mu_{jj}(t) + \sum_{i \in V \setminus \{j\}} \mu_{ji}(t) + \varepsilon_j(t),
\end{aligned}
$$

where we call $\mu_{jj}(t)$ the *intrinsic contribution term* and $\mu_{ji}(t)$ the *partial extrinsic contribution term*.
The *extrinsic-to-intrinsic-power ratio* (EIPR) from y_i to y_j (with respect to y_V) is defined as

$$
\eta_{ji}^2 = \frac{\text{Var}(\mu_{ji}(t))}{\text{Var}(\mu_{jj}(t))}. \tag{3.29}
$$

The EIPR is physiologically motivated as it shows similarities to a signal-to-noise ratio.
A thorough discussion of the EIPR's properties can be found in Graef et al. (2012b). In this article the relation between the EIPR and the PDC is described as well as the connection between the EIPR and Granger causality.

3.5.7 Other dependence measures

In this section we gave a brief overview over the most common dependence mea-sures in neuroscience literature. Of course there exists a very broad range of mea-sures between two signals in the literature based on different ideas and models. These measures differ in the kind of influence they observe, i.e. direct or indi-rect influence, as well as the considered directionality, i.e. directed or symmetric measures.
We presented only models based on the linear framework of section 1.1, but there are also non-linear approaches like e.g Freiwald et al. (1999) and Marinazzo et al. (2011).
A lot of measures can also be found in information theory literature, see e.g. Gabor (1946) and Schreiber (2000).

3.6 Statistical Inference

In this section we want to answer the question, how to actually compute Granger non-causality relations.

In practice we start by fitting an AR(p)-model to the observed data $y(t), T = 1, \ldots T$ (instead of the theoretical AR(∞)-model). The order p of the model is determined by means of an information criterion like AIC or BIC. For further information on the estimation we refer to Hannan and Deistler (2012) or Lütkepohl (2007). This procedure yields an estimation of the AR polynomial $\hat{a}(z)$.

There are several ideas how to construct a statistical test for Granger non-causality, and we briefly present two of them.

Our first statistical test is directly based on the AR coefficients and is described in Eichler (2005). According to criterion (3.19) y_i is Granger non-causal for y_j (with respect to y_V) if $A_{ji}(m) = 0$ for $m = 1, \ldots, p$. In practice the estimations $\hat{A}_{ji}(m)$ will not be exactly zero, so we have to apply a statistical test. If the observations $y_V(t)$ are normally distributed, the AR estimates $\hat{A}_{ji}(m)$ (computed by the Yule Walker equations) are asymptotically jointly normally distributed with mean $A_{ji}(m)$. Under the null hypothesis that y_i is Granger non-causal for y_j (with respect to y_V) ($\mathcal{H}_0 : y_i \nrightarrow y_j | y_V$), a simple test statistics S_{ji} can be constructed based on the estimations of the AR coefficients. The statistics S_{ji} is then asymptotically χ^2 distributed with p degrees of freedom.

We have seen in subsection 3.5.4 that non-dependence in the PDC sense for all frequencies is equivalent to Granger non-causality. Therefore, we can statistically test if the PDC is zero for all frequencies in order to yield a Granger non-causality relation. In Schelter et al. (2005) this procedure and the statistical properties of its test are discussed thoroughly.

As we will see in chapter 6 an F-test on the coefficients $\hat{a}(z)$ is also possible to determine Granger causal relations.

3.7 Partial Granger causality

The identification of causal structures in biological systems can be confounded by exogenous inputs. This means that the recorded variables are likely to be influenced by unrecorded variables. A novel variant of the measure of conditional linear dependence from subsection 3.5.1 was introduced in Guo et al. (2008) in order to cope with these influences.

The idea of this novel measure called partial Granger causality is to first calculate the prediction errors and then to partialize out the instantaneous effects of y_i on y_j. The log ratio of the partialized prediction errors is the considered measure of causality.

Before giving the definition we repeat the definition of the measure of conditional linear dependence from equation (3.24) for a better understanding

$$\mathcal{F}_{i \to j | V \backslash \{i,j\}} = \ln \frac{\sigma^2(y_j | \overline{y_{V \backslash \{i\}}})}{\sigma^2(y_j | \overline{y_V})} = \ln \frac{\mathrm{Var}(y_j | \overline{y_{V \backslash \{i\}}})}{\mathrm{Var}(y_j | \overline{y_V})}.$$

The causality measure *partial Granger causality* from y_i to y_j (with respect to y_V) is defined as

$$\mathcal{F}^P_{i \to j | V \backslash \{i,j\}} =$$
$$= \ln \frac{\mathrm{Var}(y_j | \overline{y_{V \backslash \{i\}}}) - \mathrm{Cov}(y_j y_{V \backslash \{i,j\}} | \overline{y_{V \backslash \{i\}}}) \mathrm{Var}(y_{V \backslash \{i,j\}} | \overline{y_{V \backslash \{i\}}})^{-1} \mathrm{Cov}(y_{V \backslash \{i,j\}} y_j | \overline{y_{V \backslash \{i\}}})}{\mathrm{Var}(y_j | \overline{y_V}) - \mathrm{Cov}(y_j y_{V \backslash \{i,j\}} | \overline{y_V}) \mathrm{Var}(y_{V \backslash \{i,j\}} | \overline{y_V})^{-1} \mathrm{Cov}(y_{V \backslash \{i,j\}} y_j | \overline{y_V})},$$

where $\mathrm{Var}(y_j | \overline{y_V})$ denotes the variance of the prediction error when predicting $y_j(t+1)$ from $\overline{y_V(t)}$ and $\mathrm{Cov}(y_j y_{V \backslash \{i,j\}} | \overline{y_V})$ denotes the covariance of the prediction errors when predicting $y_j(t+1)$ (or respectively $y_{V \backslash \{i,j\}}(t+1)$) from $\overline{y_V(t)}$.
In this definition we see that the instantaneous effects of y_i on y_j are partialized out. We already encountered a partialization in subsection 3.2.2 in a dynamic context, but for the partial Granger causality it is just a static partialization.

Due to the removal of these instantaneous effects, the measure shows positive features when applied to data with exogenous inputs. The properties of this measure are discussed in detail in the original paper Guo et al. (2008) and are revisited in Barrett et al. (2010).

The interested reader may have noted, that the nomination partial Granger causality is sloppy for a measure of causality, because it is misleading in regard to the definitions of sections 3.3 and 3.4. It should rather refer to a type of (non-)causality where partialized prediction errors are compared following the idea of criterion (3.13).

If there is no correlation between the components of the prediction errors, the partial Granger causality equals the measure of conditional linear dependence, in this case it would measure (Granger) causality. In case of correlations the connection to Granger causality is more difficult, but as mentioned in the last paragraph partial Granger causality could be the basis for the definition of a novel type of (non-)causality.

3.8   Additional information

In this chapter we presented the concept of Granger causality as well as undirected and directed dependence measures. There are further important topics in this context we want to mention.

On one hand we introduce ideas, which we will revisit in later chapters, and on the other hand we present additional information for the interested reader.

3.8.1 Granger causality

The framework for the application of Granger causality is based on regular AR systems, i.e. the covariance matrix Σ of the errors has full rank. In this case the three criteria (3.12), (3.13) and (3.17) are equivalent, or respectively criteria (3.19), (3.20) and (3.21). Especially the projection (from $y(t+1)$ on $\overline{y_V(t+1)}$) coefficients in criterion (3.19), which are displayed in $a(z)$, are unique. A zero projection coefficient signifies a Granger non-causality relation.

In applications with high-dimensional co-moving data, such as e.g. EEG or ECoG data, the error covariance Σ will be almost singular. If Σ is singular the projection coefficients are not necessarily unique (although the projection itself is). If the projection coefficients are not unique, the contribution of e.g. $\overline{y_i(t)}$ to $y_j(t+1)$ is no longer well defined and therefore the notion of Granger causality itself is no longer well defined in this case.

In chapter 6 we present a methodology for the causal analysis of a singular AR system, i.e. the covariance matrix Σ of the errors is singular. The idea will be to rewrite the singular AR representation $a(z)y(t) = \varepsilon(t)$ into an MA representation $y(t) = w(z)\varepsilon(t)$, where $w(z)$ is a tall transfer function. By analyzing selections of $w(z)$ we will yield influence statements for the whole process.

To our knowledge the first Granger causal analysis of a process via an MA representation was proposed in Sims (1972).

As we mentioned in the discussion of the multivariate Granger causality, it is very important to specify the conditioning set for the causal relations. This means that the relation $y_i \overset{?}{\to} y_j | y_V$ could change when we consider a sub-system S of V ($S \subset V$). This can be extended to a second sub-system $R \subset S$ where the relation could change again, and so on. These changing causality relations are called spurious causalities, they were first mentioned in the original paper Granger (1969) and the trivariate case was analyzed in Hsiao (1982). If we know the original system, we will see in chapter 4 how to retrieve non-causality relations for sub-systems.

The other direction is much more sophisticated. If we have only observed a sub-system, is it possible to state causal relations for the original system? But the more pressing question is, how do we notice that we have only observed a sub-system? Especially when working with real data we can never be sure to have included all important variables in our investigation. But at least there are some indicators for the presence of a larger original system.

In all our considerations (apart from chapter 4) we assume that all important variables are included in the analysis. This implicates that the relations $y_i \overset{?}{\to} y_j | y_V$ really indicate direct relations. If V would be the sub-set of a larger original system U, these relations could include direct as well as indirect effects. If not all

important variables are included in the analysis, we refer to Eichler (2005) and Eichler (2009).

When speaking of causality in neuroscience literature there are two prevalent concepts. We thoroughly discussed Granger causality, which is based on a linear time-discrete model. The second renown causality concept is called *dynamic causal modeling* and was proposed in Friston et al. (2003). Its framework is based on a non-linear time-continuous model.

3.8.2 Dependence measures

We presented spectral and non-spectral dependence measures, where the spectral measures are dependent on the frequency λ. Both types of measures have advantages and disadvantages.
Spectral dependence measures can be evaluated for each frequency. On one hand this is an advantage if you are interested in the dependence of specific frequency bands, e.g. the dependence between channels in the ϑ-band in EEG data. But on the other hand a rule has to be defined for the calculation of an overall dependence statement (summing up the information for all frequencies), and this rule could be an error source.
Non-spectral dependence measures normally yield only a single number to quantify the strength of the dependence.

As aforementioned there exist a lot of different dependence measures in neuroscience literature. A good comparison of the most important measures can be found in Florin (2010). In this work the application of the measures to actual data via statistical testing is described thoroughly.

Up to this point we only talked about the stationary case. This means that the dependence between y_i and y_j does not change over time. Of course, this assumption is not reasonable for real world data, especially for EEG and ECoG data. A simple idea to cope with instationarities is the use of a sliding window, where the data are assumed to be stationary within a single window, see e.g. Bodenstein and Praetorius (1977). Another interesting idea is the use of state space modeling as described in Sommerlade et al. (2012).

The main goal of this thesis is the analysis of the dependence structure of EEG/ ECoG data. As we mentioned before a lot of dependence measures have been invented in the neuroscience context. For the interested reader we present a list of selected publications regarding the application of dependence measures to EEG/ ECoG data: Chavez et al. (2003), Matysiak et al. (2005), Osterhage et al. (2007) and Florin et al. (2011).

Chapter 4

Graphical modeling

The goal of our considerations is the analysis of the *inner structure* of an n-dimensional stationary process. As we have discussed in chapter 3, the analysis of the whole process as well as of all sub-processes is necessary in order to fully understand the underlying structure. Of course this kind of analysis would be computationally intensive. In this chapter we present an easy and elegant way for gaining more insights in the structure of a process. This approach is called graphical modeling.

In general graphical modeling refers to the use of graphs and graph theory in order to analyze the causal structure of a multivariate stochastic variable. The basic idea of graphical modeling is to find a connection between the conditional independence relations of the variable, i.e. orthogonality relations (in the corresponding spaces), and the separation properties of a graph corresponding to the structure of the variable.

In the last decades there has been a substantial interest in graphical modeling and a lot of research has been conducted. Most of this research has been focused on the dependence structure in the iid case. A not exhaustive list of surveys on this topic is Lauritzen (1996), Whittaker (2000), Edwards (2000), Pearl (2000) and Cowell et al. (2007).

In this thesis we are concerned with the causal analysis of ECoG data, i.e. temporally ordered data. Therefore, we focus on graphical modeling for time series. To the best of our knowledge this kind of analysis was introduced in Brillinger (1996) and Dahlhaus (2000). A good overview concerning the topic of graphical modeling for time series is Eichler (2006a).

Although other types of graphical models are conceivable, we limit ourselves to graphs where the vertices correspond to the one-dimensional component processes of a multivariate stationary process, see Dahlhaus and Eichler (2003). In this context the edges will signify dependence between the components.

It is important to note that graphical models are more than mere graphs where the

dependencies between the components are depicted. Based on the properties of
the graphical model and the underlying stationary process, conclusions regarding
sub-processes can be drawn in an easy way.
This is the main reason we are presenting graphical models in this thesis, they are
an easy way to analyze the inner structure of a regular AR model.

In this chapter we present two types of graphical models for time series and their
use. First, the partial correlation graphs, which are undirected graphical models,
and second Granger causality graphs, which are mixed graphs, because they con-
tain directed and undirected edges.
We thoroughly discuss Granger causality graphs and their properties. Furthermore,
we briefly consider statistical inference and how to address systems with unob-
served variables.

Throughout this chapter we analyze n-dimensional stationary processes satisfying
the assumptions of section 1.1 and use the definitions from chapter 3.
We introduce some notations from graph theory in section 4.1. The partial correla-
tion graphs are discussed in section 4.2 and the Granger causality graphs in section
4.3. The practical application of graphical models is discussed in section 4.5. In
section 4.4 we talk about the treatment of unobserved variables.

4.1   Graph theory

In this section we compile useful definitions and notations from graph theory.
A graph $G = (V, E)$ consists of a *vertex set $V = \{1, \dots, n\}$* and an *edge set E*. Ac-
cording to the types of edges in E we distinguish between undirected, directed and
mixed graphs. Undirected graphs contain only undirected edges, directed graphs
contain only directed edges and mixed graphs contain both types. With the intro-
duction of the vertex set V we see where the notation y_V for y stems from.
We can easily draw a graphical representation of a graph $G = (V, E)$ by display-
ing the vertices with circles and drawing the edges of E. For undirected graphs we
draw an edge $a \,\relbar\!\relbar\, b$ in the graphical representation if $a \relbar\!\relbar b \in E$ (the second edge
for $b \relbar\!\relbar a \in E$ is omitted). For directed graphs we draw an edge $a \longrightarrow b$ in the
graphical representation if $a \longrightarrow b \in E$. Figure 4.1 shows an exemplary undirected
and a mixed graph. Mixed graphs show a combination of directed and undirected
edges.

A *path* in a graph $G = (V, E)$ is a sequence $\pi = (e_1, \dots, e_k)$ of edges $e_i \in E$ and a
corresponding sequence of vertices $(v_0, \dots, v_k)$ where e_i connects v_{i-1} and v_i. The
vertices v_0 and v_k are the *endpoints* of the path, the remaining vertices $v_1, \dots, v_{k-1}$
are called *intermediate vertices*.
In undirected graphs, for sets $A, B, S \subset V$ (pairwise disjoint) we say that C *sepa-
rates A and B*, if every path from an element of A to an element of B contains at

Figure 4.1: (a) Example of an undirected graph. (b) Example of a mixed graph containing a direct and an undirected edge.

least one element of the separation set S. This is an intuitive definition.

For mixed graphs, which contain directed $\longrightarrow$ and undirected edges $\text{-}\,\text{-}\,\text{-}$, we have to go into more detail for the definition of separation. There exist different types of separation definitions for mixed graphs. For our purposes we will use the notation of m-separation as introduced in Richardson (2003), which we explain in the following.

Given a path π, an intermediate vertex c of the path is called a *collider* of the path if the edges preceding and succeeding c on the path both have an arrowhead or a dashed tail at c, i.e. $\longrightarrow c \longleftarrow$, $\longrightarrow c \text{-}\,\text{-}\,\text{-}$, $\text{-}\,\text{-}\,\text{-} c \longleftarrow$ or $\text{-}\,\text{-}\,\text{-} c \text{-}\,\text{-}\,\text{-}$, otherwise c is a *non-collider* of the path. Let S be a subset of V and let i and j be two vertices not in S. Then a path π between the vertices i and j is called *m-connecting* given S if

(i) every non-collider of the path is not in S and

(ii) every collider of the path is in S,

otherwise the path is *m-blocked* given S. If all paths between i and j are m-blocked given S, then i an j are called *m-separated* given S. In analogy two disjoint subsets I and J are called *m-separated* given S if for every pair $i \in I$ and $j \in J$ the vertices i and j are m-separated given S.

At first glance this definition might seem complicated, but actually it is very intuitive. To check the separation between two sets we have to consider all paths between the components of these sets. All of these paths have to be m-blocked in order that the sets are m-separated, only one m-connecting path between two components of the sets will unseparate them. For each path we have to check the two aforementioned criteria, or equivalently we could look for a non-collider in S or a collider not in S in order for the path to be m-blocked.

In non-mathematical terms a non-collider lets information pass through, a collider blocks the information on a path and the set S switches colliders and non-colliders. With this easy interpretation you simply have to check if any path between two sets lets the information flow unblocked.

For clarification purposes we give a short example of m-separation. Consider the mixed graph in figure 4.2. We want to analyze the separation relations between the vertices 1 and 4.
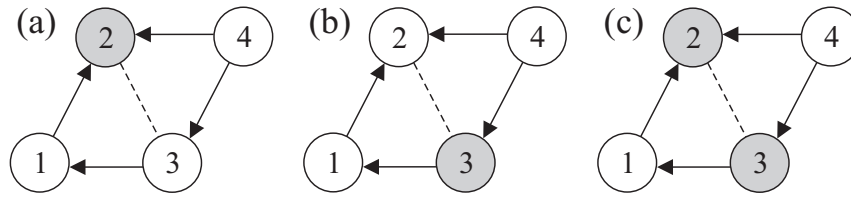
Figure 4.2: Example of a mixed graph for the illustration of m-separation. (a) $S = \{2\}$, vertices 1 and 4 are connected. (b) $S = \{3\}$, vertices 1 and 4 are m-separated . (c) $S = \{2, 3\}$, vertices 1 and 4 are connected .

If S is the empty set the path $1 \longleftarrow 3 \longleftarrow 4$ is m-connecting, therefore 1 and 4 are not separated. In the case that $S = \{2\}$, as seen in figure 4.2 (a), the path $1 \longleftarrow 3 \longleftarrow 4$ is m-connecting and the vertices 1 and 4 are not separated. In both cases we simply found an m-connecting path which yielded the separation relation.

The analysis is more interesting if $S = \{3\}$, as seen in in figure 4.2 (b). In order to check the separation between vertex 1 and 4 we have to consider all paths between them. The most obvious paths are $1 \longrightarrow 2 \longleftarrow 4$, $1 \longleftarrow 3 \longleftarrow 4$, $1 \longrightarrow 2 \dashdash 3 \longleftarrow 4$ and $1 \longleftarrow 3 \dashdash 2 \longleftarrow 4$. Vertex 2 is a collider not in S of the path $1 \longrightarrow 2 \longleftarrow 4$, therefore the path is m-blocked. Vertex 3 is a non-collider in S of the path $1 \longleftarrow 3 \longleftarrow 4$, hence the path is m-blocked. In both paths $1 \longrightarrow 2 \dashdash 3 \longleftarrow 4$ and $1 \longleftarrow 3 \dashdash 2 \longleftarrow 4$ the vertex 2 is a collider not in S, therefore both paths are m-blocked. Looking at these 4 paths is not enough since for the analysis we have to check all possible paths also including self-intersecting paths. We choose another approach for the reasoning. Every path passing through vertex 2 contains it as an collider, for this reason all these paths are m-blocked since 2 is not in S. The only path between vertex 1 and 4 not containing vertex 2 is $1 \longleftarrow 3 \longleftarrow 4$, which is m-blocked since vertex 3 is a non-collider in S. It follows that all paths between vertices 1 and 4 are m-blocked given $S = \{3\}$ and therefore vertices 1 and 4 are m-separated given $S = \{3\}$.

In the case $S = \{2, 3\}$, as seen in figure 4.2 (c), the path $1 \longrightarrow 2 \longleftarrow 4$ is m-connecting vertices 1 and 4 because the vertex 2 is a collider in S, therefore the vertices 1 and 4 are not m-separated given $S = \{2, 3\}$.

For our purposes we need one last definition. A path π between vertices i and j is called *j-pointing* if it has an arrowhead at the endpoint j. More generally, a path between I and J is *J-pointing* if it is j-pointing for at least one $j \in J$.

A path π between vertices i and j is called *bi-pointing* if it has an arrowhead at both endpoints. A path between I and J is *bi-pointing* if it is bi-pointing for at least one pair $(i, j) \in I \times J$.

4.2 Partial correlation graph

As the name suggests, the partial correlation graph is based on the partial spectral coherence (PSC), see subsection 3.2.2 . It is an undirected graph and was introduced in Dahlhaus (2000).

The vertices in the partial correlation graph signify the components of y_V and the edges are characterized by the PSC. The idea behind the partial correlation graph is, that an edge $i \text{---} j$ is missing, if the components y_i and y_j are uncorrelated conditional on the other components of the process.

We want to establish a connection between the separation relations of the partial correlation graph and conditional independence relations of the process. For this reason we introduce some notations. We recall the definition of the residual η_i (and η_j, see subsection 3.2.2), which signifies y_i removed by the influences of all other (past, present and future) components apart from y_i and y_j, see equation (3.2). Then we define the following relation

$$y_i \perp y_j | y_{V \setminus \{i,j\}} :\Longleftrightarrow Cov(\eta_i(t)\eta_j(t+s)) = 0 \; \forall s. \tag{4.1}$$

Of course, this relation is also equivalent to the PSC being zero, $R^2_{ij|V \setminus \{i,j\}}(\lambda) = 0 \; \forall \lambda$. The definition of the relation $\perp$ can be extended for arbitrary multivariate subprocesses of y_V using the multivariate extension of the partial spectral coherence. We have

$$y_A \perp y_B | y_C \Longleftrightarrow Cov(\eta_{A|C}(t)\eta_{B|C}(t+s)) = 0 \; \forall s \Longleftrightarrow R^2_{AB|C}(\lambda) = 0 \; \forall \lambda. \tag{4.2}$$

In general relations of the form $y_A \perp y_B | y_C$ are called conditional independence relations. They are interpreted as orthogonality relations in the corresponding probability spaces.

Using relation (4.1) we may define the partial correlation graph.

Definition 4.1. Let y_V be a multivariate stationary process satisfying the assumptions of section 1.1. The *partial correlation graph* associated with y_V is a graph $G_{PC} = (V, E)$ with vertex set $V = \{1, \ldots, n\}$ and edge set E such that for $i, j \in V$ ($i \neq j$) the edge $i \text{---} j \notin E$ if and only if $y_i \perp y_j | y_{V \setminus \{i,j\}}$.

This means an edge $i \text{---} j$ is missing if y_i and y_j are uncorrelated given the remaining components $y_{V \setminus \{i,j\}}$. From the alternative definition of the partial spectral coherence via the inverse of the spectral density, see subsection 3.2.2, we see that the missing edges in the graph can be identified by the zeros in the rescaled inverse of the spectral density.

With the aforementioned notations we are able to state the connection between the separation relations of the partial correlation graph G_{PC} and the conditional independence relations of the form (4.2) for y_V.

Figure 4.3: Partial correlation graph of system (4.3). Vertex 2 separates 1 and 3

Theorem 4.1. *Let y_V be a multivariate stationary process satisfying the assumptions of section 1.1 and let $G_{PC} = (V, E)$ be the corresponding partial correlation graph. Let $A, B, S \subset V$ where S separates A and B. Then we have*

$$y_A \perp y_B | y_S .$$

Proof. See Dahlhaus (2000). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

The use of this theorem is clear. We construct the partial correlation graph for a multivariate stationary process based on the bivariate conditional independence relations, i.e. the non-correlation relations, of the component processes. Then the application of theorem 4.1 yields the independence relations for arbitrary sub-processes. So this is an easy and effective way for the analysis of the dependence structure of a process.

For clarification purposes we give a short example for the application of the last theorem. Consider the following AR process

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}(t) = \begin{pmatrix} 0 & \alpha z & 0 \\ 0 & 0 & 0 \\ 0 & \beta z & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}(t) + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}(t) \tag{4.3}$$

with $\Sigma = \text{Cov}(\varepsilon) = I_3$. Simple calculations yield the spectral density f_{yy} and its inverse

$$f_{yy}^{-1}(\lambda) = 2\pi \begin{pmatrix} 1 & -\alpha e^{-i\lambda} & 0 \\ -\alpha e^{i\lambda} & 1 + \alpha^2 + \beta^2 & -\beta e^{i\lambda} \\ 0 & -\beta e^{-i\lambda} & 1 \end{pmatrix}. \tag{4.4}$$

Based on the zeros of f_{yy}^{-1} we can draw the partial correlation graph in figure 4.3. We see that vertex 2 separates the vertices 1 and 3 in the partial correlation graph, and therefore the application of theorem 4.1 yields $y_1 \perp y_3 | y_2$, i.e. y_1 and y_3 are uncorrelated after the influence of y_2 is removed from both.

This statement is trivial, looking at the system (4.3), but it shall serve as an example for the application of the partial correlation graph. In more complex and higher dimensional systems the computation of sub-systems would be more time-consuming, therefore the partial correlation graph is an effective way to gain insights into the structure of a process.

As already discussed in subsection 3.2.2 one advantage of the partial spectral coherence or respectively the partial correlation graph is that in practical applications the spectral density can be estimated via a non-parametric estimator. The AR representation of the investigated time series is not needed for the analysis. A statistical test for the practical application of the PSC can be found in Dahlhaus (2000).

4.3 Granger causality graph

In this section we extend the idea of the partial correlation graph given above to Granger causality. First we draw a graph based on Granger causalities between the component processes and then we can infer non-causalities for arbitrary sub-processes. As Granger causality is a directed property we have to change the definitions accordingly.

To the best of our knowledge this kind of graphical model or rather this type of analysis was introduced in Eichler (2007). For the explanations of this section we will heavily draw from Eichler (2006a).

It is an intuitive idea to depict the Granger causalities in a graph. In order to incorporate all independence relations in the analysis (and the graph respectively) we also include the (contemporaneous) non-correlation relations. Following these ideas we define the following graph.

Definition 4.2. Let y_V be a multivariate stationary process satisfying the assumptions of section 1.1. The *path diagram* associated with y_V is a graph $G = (V, E)$ with vertex set V and edge set E such that for $i, j \in V$ ($i \neq j$)

(i) $i \longrightarrow j \notin E \iff A_{ji}(m) = 0 \; \forall m$,

(ii) $i \dashrightarrow j \notin E \iff \Sigma_{ij} = 0$.

In other words the path diagrams contains an edge $i \longrightarrow j$ if y_i is Granger causal for y_j with respect to y_V. An edge $i \dashrightarrow j$ is present if y_i and y_j are contemporaneously correlated with respect to y_V. The conditions in (i) and (ii) can be replaced by equivalent criteria as described in section 3.4.

The associated path diagram shows Granger causalities, therefore it is often called *Granger causality graph*, see e.g. Dahlhaus and Eichler (2003).

Following the idea presented in section 4.2 we want to state a connection between the graph separation relations of the associated path diagram and the conditional independence relations of y_V. Now the considered graph is a mixed graph and we will use the notion of m-separation as introduced in section 4.1. The conditional independence relations now consist of non-causality relations of the form $y_B(t + 1) \perp \overline{y_A(t)} | \overline{y_{S \setminus A}(t)}$ for all $t \in \mathbb{Z}$ (which are non-correlation relations in the right spaces) and (contemporaneous) non-correlation relations of the form $y_A(t + 1) \perp y_B(t + 1) | \overline{y_S(t)}$ for all $t \in \mathbb{Z}$, compare criteria (3.17) and (3.18). Using the associated path diagram we are able to state the following theorem.

Theorem 4.2. *Let y_V be a multivariate stationary process satisfying the assumptions of section 1.1 and let $G = (V, E)$ be the corresponding associated path diagram. Let $S \subseteq V$ and let I and J be disjoint subsets of S. If every J-pointing path between I and J is m-blocked given $S \setminus I$, then y_I is Granger non-causal for y_J with respect to y_S.*

Proof. See Eichler (2007).   □

A similar theorem can be stated for the contemporaneous non-correlation relations.

Theorem 4.3. *Let y_V be a multivariate stationary process satisfying the assumptions of section 1.1 and let $G = (V, E)$ be the corresponding associated path diagram. Let $S \subset V$ and let I and J be disjoint subsets of S. If*

  (i) *$i \dashleftarrow\dashrightarrow j \notin E$ for all $i \in I$ and $j \in J$, and*

  (ii) *every bi-pointing path between I and J is m-blocked given S,*

then y_I and y_J are contemporaneously uncorrelated with respect to y_S.

Proof. See Eichler (2007).   □

With these theorems we are able to derive the conditional independence relations, including non-causality and contemporaneous non-correlation relations, for arbitrary sub-processes y_S of y_V. These conditional independence relations can again be depicted in a graph $\tilde{G}_S = (S, E_S)$ for the sub-system y_S. It is important to note, that this new graph for the sub-system generally is not the associated path diagram for y_S. Besides the conditional independence relations for y_S (derived from the associated path diagram) displayed in $\tilde{G}_S$, there could be additional ones. In other words, not all arrows in $\tilde{G}_S$ indicate Granger causality, as they would in the associated path diagram. Thus, theorems 4.2 and 4.3 provide only sufficient and not necessary conditions for Granger non-causality relations in sub-processes y_S of y_V.

For a better understanding of the application of theorems 4.2 and 4.3 we present an example. This example is taken from Eichler (2006a). Consider the following AR system

$$
\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}(t) = \begin{pmatrix} 0 & 0 & 0 & \alpha z^2 \\ 0 & 0 & \beta z & \gamma z \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}(t) + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}(t)
\tag{4.5}
$$

with $\Sigma = \mathrm{Cov}(\varepsilon) = I_4$. The associated path diagram of system (4.5) is shown in figure 4.4 (a).

We want to analyze the non-causality relations between two component processes with respect to arbitrary sub-processes y_S using the associated path diagram (figure 4.4(a)) of y_V. For this analysis it is easier to consider the (not necessary) causal relations than the non-causality relations. In other words, which edges can be drawn in the sub-graph. We will sloppily say y_i is Granger causal for y_j with respect to y_S, denoted by $y_i \rightarrow y_j | y_S$, if the relation $y_i \nrightarrow y_j | y_S$ is not inferred from the associated path diagram of y_V.
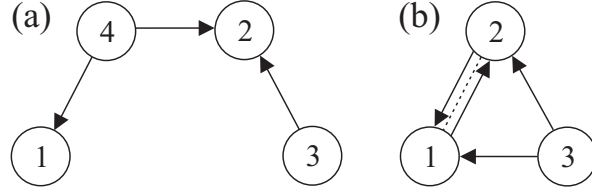
We will start with the three dimensional sub-processes.

Figure 4.4: (a) Associated path diagram of y_V from system (4.5). (b) The use of theorem 4.2 yields the path diagram for $y_{\{1,2,3\}}$.

In the case $S = \{1, 2, 3\}$ we infer the following relations: Clearly we have $y_3 \rightarrow y_2 | y_{\{1,2,3\}}$ because of the direct path $3 \longrightarrow 2$. Via the path $1 \longleftarrow 4 \longrightarrow 2$ we have $y_1 \rightarrow y_2 | y_{\{1,2,3\}}$, $y_2 \rightarrow y_1 | y_{\{1,2,3\}}$ and $y_2 \sim y_1 | y_{\{1,2,3\}}$ because vertex 4 is a non-collider of this path. Considering the path $3 \longrightarrow 2 \longleftarrow 4 \longrightarrow 1$ we get $y_3 \rightarrow y_1 | y_{\{1,2,3\}}$, because the vertex 2 is a collider in S and the vertex 4 is a non-collider. To infer a causality relation from vertex i to vertex j we need a path ending with an $\longrightarrow$ edge pointing to j. In the aforementioned relations we already used all $\longrightarrow$ edges, therefore there are no additional relations in this sub-system. Figure 4.4(b) shows the derived independence relations for the sub-process $y_{\{1,2,3\}}$.

The case $S = \{1, 2, 4\}$ is simple as we only have $y_1 \rightarrow y_2 | y_{\{1,2,4\}}$, $y_2 \rightarrow y_1 | y_{\{1,2,4\}}$ and $y_2 \sim y_1 | y_{\{1,2,4\}}$.

Similarly the case $S = \{2, 3, 4\}$ yields only $y_3 \rightarrow y_2 | y_{\{2,3,4\}}$ and $y_4 \rightarrow y_2 | y_{\{2,3,4\}}$.

In the case $S = \{1, 3, 4\}$ we have the relation $y_4 \rightarrow y_1 | y_{\{1,3,4\}}$, as the vertex 2 (as a collider) separates the vertices 3 from 1 and 4.

We get the causality relations for a 2-element set S by considering only the paths ending with an $\longrightarrow$ edge. Therefore, we have $y_4 \rightarrow y_1 | y_{\{1,4\}}$, $y_4 \rightarrow y_2 | y_{\{2,4\}}$ and $y_3 \rightarrow y_2 | y_{\{2,3\}}$. Additionally we have $y_1 \rightarrow y_2 | y_{\{1,2\}}$, $y_2 \rightarrow y_1 | y_{\{1,2\}}$ and $y_2 \sim y_1 | y_{\{1,2\}}$ as vertex 4 is a non-collider of the path $1 \longleftarrow 4 \longrightarrow 2$.

In this example we have seen that the associated path diagram encodes conditional independence relations for arbitrary sub-processes. It is possible to construct other graphs that encode other independence relations, e.g. the bivariate Granger causality graph, see Eichler (2006a). If a mixed graph G (containing $\longrightarrow$ and $\text{-}\text{-}\text{-}$ edges) encodes certain conditional independence relations of a process y_V, we say that y_V satisfies a Markov property with respect to the graph G.

Definition 4.3. We say that a stationary process y_V (satisfying the assumptions of section 1.1) satisfies the *global Granger causal Markov property* with respect to a mixed graph G if for all $S \subseteq V$ and all disjoint sub-sets I and J of S the following conditions hold

(i) y_I is Granger non-causal for y_J with respect to y_S whenever in the graph G every J-pointing path between I and J is m-blocked given $S \setminus I$,

(ii) y_I and y_J are contemporaneously uncorrelated with respect to y_S whenever

in the graph G the sets I and J are not connected by an undirected edge - - -
and every bi-pointing path between I and J is m-blocked given S.

In other words, all conditional independence relations that can be derived from the
graph G also have to be found in the process itself. So the graph itself may encode
fewer conditional independence relations than can be found in the process itself.
We want to emphasize the construction of this definition, a process satisfies a prop-
erty with respect to a graph. So the graph is given and we check whether the
process satisfies the property or not.
With definition 4.3 theorems 4.2 and 4.3 state that a stationary process (satisfying
the assumptions of section 1.1) satisfies the global Granger causal Markov property
with respect to its associated path diagram.
A saturated graph, i.e. a graph that includes all possible edges, trivially satisfies
the global Granger causal Markov property for all stationary processes, because it
does not entail any separation relations.

To conclude this section we want to mention that different associated path diagrams
encode different conditional independence relations. So there do not exist two
different associated path diagrams that encode the same conditional independence
relations. This is a property of associated path diagrams, this property (generally)
does not hold for more general types of graphs.

4.4   Unobserved variables

In Granger's original definition he assumed that all important information is known.
In modern data analysis it is a valid question whether all important variables are
taken into account, or if an important variable, i.e. a component process, is un-
observed. It is very difficult to answer this question. The application of graphical
modeling yields indications for unobserved variables in certain cases and provides
a tool to (partially) cope with this problem. The ideas discussed in this section are
taken from Eichler (2009).
The term *unobserved variable* normally refers to an unobserved component in a
system, which is only partially observed. In our considerations we know the origi-
nal process and simply leave out one or more component processes, which is nor-
mally not the case in practical applications.

To explain the problem of unobserved variables we continue the example of system
(4.5). We consider the case that y_4 is unobserved. As shown in Eichler (2006a) a
representation of $y_{\{1,2,3\}}$ is given by

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}(t) = \begin{pmatrix} 0 & \frac{\alpha\beta}{1+\beta^2}z & \frac{\alpha\beta\gamma}{1+\beta^2}z^2 \\ 0 & 0 & \gamma z \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}(t) + \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \tilde{\varepsilon}_3 \end{pmatrix}(t) \qquad (4.6)$$
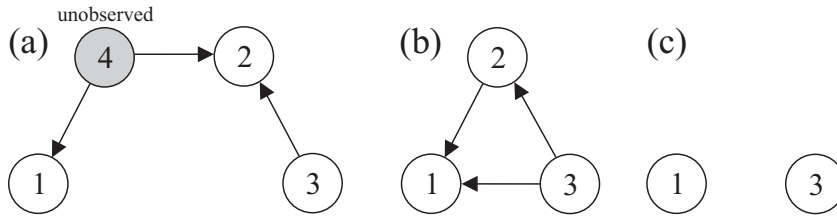
Figure 4.5: Changing causality relations when considering sub-processes
(a) Associated path diagram for the full process of system (4.5), y_4 is unobserved.
We infer $y_3 \nrightarrow y_1 | y_{\{1,3\}}$. (b) Granger causalities for sub-system $y_{\{1,2,3\}}$ derived from
representation (4.6). (c) In the 2-dimensional sub-system $y_{\{1,3\}}$ we have: y_3 is non-
causal for y_1.

with

$$
\begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \tilde{\varepsilon}_3 \end{pmatrix}(t) = \begin{pmatrix} 1 & -\frac{\alpha\beta}{1+\beta^2}z & 0 & \frac{\alpha}{1+\beta^2}z^2 \\ 0 & 1 & 0 & \beta z \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}(t). \tag{4.7}
$$

The Granger causal relations indicated by this representation are displayed in figure
4.5(b). Figure 4.5(a) shows the associated path diagram of the full process, and the
application of theorem 4.2 yields the relation $y_3 \nrightarrow y_1 | y_{\{1,3\}}$ for the bivariate sub-
system $y_{\{1,3\}}$. This relation can be found in the data, but it cannot be derived from
the diagram 4.5(b).

This is an example for changing causality relations when considering sub-systems,
in particular this phenomenon is called spurious causality of type I, where causality
relations with respect to the full process vanish when considering a sub-process, see
Hsiao (1982).

The change of causality relations when analyzing sub-systems is an indicator for
unobserved variables as can be seen in this example.

The literature suggests two different approaches to cope with the problem of un-
observed variables in the iid case. The first approach suggests to include the unob-
served variables as nodes in the graph, see e.g. Pearl (2000). The second approach
focuses on the conditional independence relations of the observed variables. In
Eichler (2006b) the second approach has been discussed for the time series case.

For the interested reader we want to briefly explain the second approach following
Eichler (2006b) and Eichler (2009). By the introduction of an additional type of
edge --→ we extend the class of considered graphs, so called general path diagrams
contain ---, ⟶ and --→ edges. In the class of general path diagrams we are able
to display more causality relations, but we loose the unique identifiability of the
graph, i.e two different graphs may yield the same set of independence relations.
To solve this problem we restrict ourselves to graphs called *dynamic ancestral*

graphs, which are uniquely identifiable in the class of general path diagrams. These ancestral graphs are associated with ARMA processes in comparison to the normal associated path diagrams, which are associated with AR processes.

The main aim of graphical modeling (in this approach) is to display the causality structure of a process as good as possible, i.e. to depict as much conditional independence relations as possible in a single graph.

4.5   Statistical inference

So far we have only talked about the theoretical aspects of graphical modeling. Now we want to briefly discuss the application of graphical models to actual data. Graphical models are used to display the causality structure of observed processes.

The first and straightforward idea is to draw the associated path diagram for the observed data $y_V(t), t = 1, \ldots, T$. In section 3.6 we explained that normally an AR(p) model is fitted to the data using the Yule-Walker equations. The components of the estimated AR polynomial $\hat{A}_{ji}(m)$ are statistically tested if they are jointly zero by an F-test, in order to derive the bivariate Granger causal relations $y_i \overset{?}{\to} y_j | y_V$. Based on these relations we are able to draw the associated path diagram.

The associated path diagram often is a very good description for the causal structure of an observed multivariate system. But in the case of unobserved variables we will encounter changing causality relations when analyzing sub-systems. In section 4.4 we discussed the problem that these changes in the causal relations are indicators for unobserved variables. For a better representation of the causality structure of observations with unobserved variables general path diagrams are introduced.

Following the ideas in Eichler (2005) we briefly describe the practical application of general path diagrams, which include $\longrightarrow$, $\dashleftarrow\dashrightarrow$ and $\dashrightarrow$ edges. Suppose that observations $y_S(t)$ are available, which are a part of a larger system $y_V(t)$. We are searching for a graph $G = (S, E)$ that describes the causality structure of y_S as best as possible, and of course the graph should not imply any independence relations that do not hold in y_S.

Definition 4.4. A general path diagram G is *consistent* with y_S if y_S satisfies the global Granger causal Markov property with respect to G.

So our aim is to determine a consistent graph that depicts the Granger causal relations of the observations as good as possible. A smaller number of edges in the graph leads to a larger set of related conditional independence relations. Therefore, we search for a *minimal consistent* graph in the sense, that the set of related conditional independence relations is maximal (but does not contain independence

relations that are not found in y_S). The idea of minimal consistent graphs in the iid case was introduced in Pearl (2000).

In practice the determination of the minimal consistent graph is based on the pairwise comparison of different consistent graphs by a statistical test on the causality relation distinguishing them.

Unfortunately two different general path diagrams might imply the same set of conditional independence relations. Therefore, the minimal consistent graph in general is not unique.

Chapter 5

Factor models

Analysis and forecasting of high-dimensional multivariate time series is an important issue in various disciplines. A common and well-understood tool for modeling and analysis of these systems is (regular) autoregressive modeling. Up to this point we talked talked about AR-(∞), but for applications you have to use finite order AR-systems. The arising problem when using regular finite order AR-systems is called *curse of dimensionality*:

Let n denote the cross-sectional dimension, let T be time dimension of the observed time series x and let the AR-order p be fixed. Then the dimension of the parameter space (for unrestricted regular autoregressive modeling) is proportional to n^2, but the number of data points is only linear in n for fixed T, which leads to numerical problems for large n.

The idea to solve this problem is to compress the information and to extract the important information from the data. This compression is not only in the time axis but also in the cross section, and leads to a reduction of the dimension of the parameter space.

Nowadays factor models are a widely used technique for dimension reduction.

The main idea behind factor models is to separate the *observations* $x(t)$ into the *latent variables* $\chi(t)$ and the noise $\eta(t)$

$$x(t) = \chi(t) + \eta(t), \tag{5.1}$$

where the latent variables include the *common* information concerning the observations and the noise is simply the rest. In our applications the common parts will reflect the important information of the observations, i.e in ECoG data the co-movement is considered the important information. The aforementioned dimension reduction is accomplished because the latent variables are generated by a small number of factors, therefrom the term *factor model*.

Our main interest in high-dimensional time series is their causal analysis. The curse of dimensionality also impairs this causal analysis, because the causal analysis is usually based on a regular AR model, as discussed in chapter 3.

We are especially interested in the analysis of EEG and ECoG data, see chapter 2. As the EEG data represent potential changes of spatially neighboring electrodes, the data are highly correlated and show co-movement. These properties lead to the aforementioned numerical problems. In chapter 6 we will return to these problems and discuss how we can use Granger causality for factor models.

In the section 5.1 we detail theoretic results concerning factor models and how the separation (5.1) can be carried out, whereas in section 5.2 we discuss a practical method for the aforementioned separation. For the following elaborations we extensively use material from Deistler et al. (2010) .

5.1   Generalized dynamic factor models (GDFMs)

For our purposes we consider generalized dynamic factor models (GDFMs), which have been introduced in Forni et al. (2000) and Forni and Lippi (2001), and in a slightly different form in Stock and Watson (2002a) and Stock and Watson (2002b). The idea of GDFMs is to combine and generalize linear dynamic factor models with strictly idiosyncratic noise, as discussed in Sargent and Sims (1977) and Scherrer and Deistler (1998), and generalized linear static factor models, as introduced in Chamberlain (1983) and Chamberlain and Rothschild (1983).

The basic idea of GDFMs is to separate the n-dimensional observation $x^n(t)$ as in equation (5.1) into [3]

$$x^n(t) = \chi^n(t) + \eta^n(t), \tag{5.2}$$

where this time, the latent variables χ^n are strongly dependent in the cross-sectional dimension, and the noise η^n is weakly dependent in the cross-sectional dimension. The exact definition of these terms is given below.

For the remainder we assume that χ^n and η^n are weakly stationary with absolutely summable covariances, have mean zero and

$$\mathbb{E}\chi_t^n \, \eta_s^{n\prime} = 0 \ \ \forall s, t \tag{5.3}$$

holds, this means that they are uncorrelated. Thus, the spectral densities corresponding to (5.2) can be written as

$$f_{xx}^n(\lambda) = f_{\chi\chi}^n(\lambda) + f_{\eta\eta}^n(\lambda). \tag{5.4}$$

The class of GDFMs is constituted via a set of assumptions. These assumptions are technical in nature and are not important in detail for our final goal, the causal analysis of a factor model. The whole list of constituting assumptions including

---

[3]In this section we want to emphasize the cross section dimension n of the observations $x(t)$ by writing n as a superscript: $x^n(t)$. Later the superscript will be omitted.

additional informations can be found in Deistler et al. (2010) and Anderson and Deistler (2008).

Assumption 5.1. *There is an n_0 such that for all $n > n_0$ the spectral density of the latent variables $f_{\chi\chi}^n$ is rational and has constant rank $q < n$ on $[-\pi, \pi]$.*

For the mathematical analysis a sequence of GDFMs (5.2) is considered. We have the following assumptions.

Assumption 5.2. *The sequence $(y_i(t)|i \in \mathbb{N}, t \in \mathbb{Z})$ corresponds to a nested sequence of models, in the sense that $\chi_i(t)$ and $\eta_i(t)$ do not depend on n for $i \leq n$.*

Assumption 5.3. *The rank q and the McMillan degree, 2ν say (see Hannan and Deistler (2012) for further information), of $f_{\chi\chi}^n$ is independent of n $(n \geq n_0)$.*

In the following we use $\omega_{\eta\eta,r}$ to denote the r-th largest eigenvalue of $f_{\eta\eta}^n$ and define strong and weak dependence according to Forni and Lippi (2001).

Assumption 5.4 (weak dependence). *$\omega_{\eta\eta,1}^n$ is uniformly bounded in λ and n.*

Assumption 5.5 (strong dependence). *The q largest eigenvalues of $f_{\chi\chi}^n$ diverge to infinity for all frequencies λ as $n \to \infty$.*

The problem with GDFMs is, that they are not identifiable for any fixed n, no matter how large n is. This means, that the separation in (5.2) cannot be achieved for finite n. In the idiosyncratic case, where $f_{\eta\eta}^n$ is diagonal the identification (and thus the separation) is possible generically for q sufficiently small in relation to n. Nevertheless, the elements of χ^n and η^n are uniquely determined from x^n for $n \to \infty$ as shown in Forni and Lippi (2001).
From now on we will omit the superscript n for the sake of simplicity.

Our final goal is the causal analysis of a factor model in chapter 6. Therefore, the most important theoretic result for our considerations is the fact, that (theoretically) the separation (5.1) is possible under technical assumptions.
In section 5.2 we will suggest a method for the practical separation.

As already mentioned we commence from assumption 5.1 for a spectral factorization of $f_{\chi\chi}$. The results of this procedure will be used in the next chapter for the causal analysis.
For the remainder of this thesis we assume that n is large enough, so that $f_{\chi\chi}$ is of rank $q < n$. We commence from the population spectral density of the latent variables $f_{\chi\chi}$. Then, according to Hannan (1970) or Rozanov (1967) we have the following result.

Theorem 5.1 (Spectral factorization)**.** *Every $n \times n$ rational spectral density $f_{\chi\chi}$ of constant rank q for all $\lambda \in [-\pi, \pi]$ can be factorized as*

$$f_{\chi\chi}(\lambda) = w(e^{-i\lambda})w(e^{-i\lambda})^*, \tag{5.5}$$

where $w(z)$ ($z \in \mathbb{C}$) is an $n \times q$ real rational matrix of full column rank which has no poles and no zeros for $|z| \le 1$.

Proof. See Rozanov (1967). □

It is easy to show, that $w(z)$ is unique up to post-multiplication by constant orthogonal matrices.

By using the notation $w(z) = \sum_{m=0}^{\infty} W(m)z^m$, $W(m) \in \mathbb{R}^{n \times q}$, we can write the latent variables χ as

$$\chi(t) = w(z)\varepsilon(t) = \sum_{m=0}^{\infty} W(m)\varepsilon(t - m), \tag{5.6}$$

where the q-dimensional input ε is white noise with $Cov(\varepsilon) = I_q$. The components of ε are called *dynamic factors*. As can be seen from this representation (5.6) the q-dimensional dynamic factors ε generate the n-dimensional latent variables χ. As can be easily seen, (5.6) corresponds to the Wold representation.

GDFMs are a large model class and offer a lot of versatility. However, for the causal analysis we will present in chapter 6, we will only need a small portion of the theoretical background from GDFMs. Our causal analysis will be based on the tall transfer function $w(z)$ derived from the spectral factorization (5.5) and shown in equation (5.6). For the practical implementation of our method matters will simplify even more.

5.2 Principal component analysis (PCA)

In this section we want to present a practical method for the separation (5.1) and respectively the derivation of the latent variables in contrast to the theoretic application of factor models in section 5.1. For this reason we use the *principal component analysis* (PCA).
The PCA is a tool for the estimation of the static factors z. Then the static factors are are modeled as a regular AR process.
First, we briefly explain how the PCA functions and give a short introduction using material from Filzmoser (2010). Second, we show its application for our purposes. Third, we model the static factors as a regular AR process.

The principal component analysis roots on Pearson (1901) and was developed by Hotelling (1933). In the analysis the data are linearly transformed in such a way, that the resulting components comprise the most important information of the data

in descending order. In this way the dimension of the data can be reduced to the number of the most important components. For further informations we refer the interested reader to Jackson (2004).

We explain this procedure for the iid case, the time series case will function analogously. Our theoretic considerations are based on the population covariance matrix Ω. Of course, for practical applications the covariance matrix of the observations has to be estimated. Accordingly, the AR polynomial for the modeling of the latent variables has to be fitted in the posterior part of this section.
In the following we discuss how the PCA works.

Let $x = (x_1, \ldots, x_n)'$ be an n-dimensional random vector with mean zero, covariance matrix $\Omega = \mathrm{Cov}(x) = \mathbb{E}xx'$ and $\Omega > 0$. For the explanation of the PCA we use x as a normal multivariate random vector, not a stationary process.
Furthermore, let $O = (o_1, \ldots, o_n)$ be an orthogonal $n \times n$ matrix with fixed components, i.e. $O^{-1} = O'$. We consider the following linear transformation

$$z = O'x \tag{5.7}$$

or written component-wise

$$z_i = o_i'x \quad \text{for} \quad i = 1, \ldots, n . \tag{5.8}$$

The transformation (5.7) yields a new n-dimensional random vector z. The variance of its components is

$$\mathrm{Var}(z_i) = \mathbb{E}\, o_i'xx'o_i = o_i'\Omega o_i \quad \text{for} \quad i = 1, \ldots, n . \tag{5.9}$$

We want to construct the transformation matrix O from (5.7) in such a way, that the variances (5.9) are maximized under the constraint that the o_i have length 1 (or $o_i'o_i = 1$). Mathematically this can be written as a Lagrangian maximization problem with a side constraint

$$\varphi_i = o_i'\Omega o_i - \lambda_i(o_i'o_i - 1) \quad \text{for} \quad i = 1, \ldots, n . \tag{5.10}$$

The derivation with respect to the unknown o_i are set equal to zero, yielding

$$\frac{\partial \varphi_i}{\partial o_i} = 2\Omega o_i - 2\lambda_i o_i = 0 \quad \text{for} \quad i = 1, \ldots, n . \tag{5.11}$$

A slight recalculation leads to

$$(\Omega - \lambda_i I_n)o_i = 0 \quad \text{for} \quad i = 1, \ldots, n \tag{5.12}$$

or written in matrix notation

$$\Omega O = O\Gamma, \tag{5.13}$$

where $\Gamma = \mathrm{diag}(\lambda_1, \dots \lambda_n)$. This result is well known under the name eigenvalue problem. The o_i are the eigenvectors of Ω and the λ_i are the corresponding eigenvalues. The covariance matrix Ω is positive definite and therefore all eigenvalues are real positive numbers. Furthermore, we sort the eigenvalues: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$.

With these definitions the linear transformation (5.7) is known as *principal components transformation*, the i-th component of z is called the i-th principal component. The principal components have expectation zero $\mathbb{E}\, z = 0$ and their covariance is $\mathrm{Cov}(z) = O'\Omega O = \Gamma = \mathrm{diag}(\lambda_1, \dots \lambda_n)$. Therefore, the variance of z_i is equal to λ_i, the i-th eigenvalue of Ω. The total variance of all principal components is equal to the sum of all eigenvalues, which is obviously equal to the total variance of x.

Now we return to the time series case and x will denote a stationary process from now on. We use the PCA on $\Omega = \mathbb{E}\, xx'$ to derive the static factors z as follows

$$
\begin{aligned}
\Omega = \mathbb{E}\, xx' \;&=\; O\Gamma O' \\
&=\; (O_1\ O_2)\begin{pmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{pmatrix}\begin{pmatrix} O_1' \\ O_2' \end{pmatrix} \\
&=\; O_1\Gamma_1 O_1' + O_2\Gamma_2 O_2',
\end{aligned}
$$

where $\Gamma = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ contains the ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$ and O is an orthogonal matrix.

As the latent variables include the important information of the observations, we assume that the important information is contained in the q-dimensional space explaining the most variance of the data. Therefore, we only use the information in the q first principal components of the data. So our transformation matrix from (5.7) will not be square: O_1 contains the first q columns of O corresponding to the q largest eigenvalues, respectively O_2 the remaining $n - q$ columns.

The latent variables $\chi(t)$ are obtained by projecting the observations $x(t)$ onto the space explaining most of the variance, this space is spaned by the columns of O_1.

$$
\chi(t) = O_1 \underbrace{(O_1'O_1)}_{I_q} \underbrace{O_1 x(t)}_{z(t)}. \tag{5.14}
$$

We use the more common notation Λ for O_1 and rewrite the last equation into the system

$$
\begin{aligned}
\chi(t) &= \Lambda z(t) \tag{5.15} \\
z(t) &= \Lambda' x(t), \tag{5.16}
\end{aligned}
$$

where Λ is termed *factor loading matrix* and the q-dimensional principal components $z(t)$ are called *static factors*. Expression (5.15) is called a *static factor model*, because the factor loading matrix is static, i.e. constant. We clearly see, that the n-dimensional latent variables χ are driven by the q-dimensional static factors z.

The second part of the system, equation (5.16), is used for the actual computation of the static factors $z(t)$.

As is well known, one could generate other (minimal) static factors by pre-multiplication with a regular matrix U

$$\chi(t) = \Lambda z(t) = \underbrace{\Lambda U^{-1}}_{\tilde{\Lambda}} \underbrace{U z(t)}_{\tilde{z}(t)} = \tilde{\Lambda}\tilde{z}(t).$$

However, this does not impair our causality analysis, because the spaces spanned by z and $\tilde{z}$ are equal.

The most important difference between modeling the latent variables via equation (5.6) or via equation (5.15) is the fact, that equation (5.15) uses a static (i.e. constant) factor loading matrix.

The next step is to model the derived static factors z as a regular AR process.

Normally the static factors z are modeled by a singular AR-model, which means that the number of dynamic factors is smaller than the number of static factors. For our purposes we assume the number of static factors to equal the number of dynamic factors.

In order to reach a from similar to equation (5.6) we assume the static factors z to satisfy the assumptions of section 1.1. Therefore, they can be modeled as a regular AR process. This is a strong assumption, but data analysis reveals, that it is reasonable for the ECoG data we want to analyze.

The approach to model the latent variables with a static factor model and the static factors as a regular (finite) AR process is sometimes called *quasi-static factor model*, see Deistler and Zinner (2007).

Following this line of thought we get a regular AR model for the static factors

$$a(z)z(t) = \varepsilon(t), \tag{5.17}$$

where $a(0) = I_q$ and $\mathrm{Cov}(\varepsilon(t)) = \Sigma > 0$.

Equations (5.15), (5.17) and the causal invertability of $a(z)$ together yield

$$\chi(t) = \Lambda z(t) = \Lambda a^{-1}(z)\varepsilon(t). \tag{5.18}$$

In this PCA-based approach it can be easily seen, that $\Lambda a^{-1}(z)$ in equation (5.18) is the transfer function $w(z)$ from equation (5.6) from section 5.1. Therefore, ε in equation (5.18) are the dynamic.

The interested reader may have noticed, that up to this point we assumed q, the number of (static) factors, to be known. For theoretical purposes q is determined via the number of diverging eigenvalues of $f_{\chi\chi}^n$ as $n \to \infty$, see assumption 5.5. In practical applications often the number of static factors is estimated first, followed by an estimation of the number of dynamic factors.

A well known method for the determination of the number of factors is the scree test, which was introduced in Cattell (1966). This test is based on a graphical comparison between the explained variance per factor and the number of the factor, this graph is called scree plot, an example is figure 6.5 in section 6.5.
Another simple method is to set a threshold for the explained variance you want to achieve and choose the number of factors accordingly.

For the conclusion we sum up the model described in this section. The latent variables χ, the factor loading matrix Λ and the static factors z are derived via the PCA from the observations x. Then the static factors are modeled by a regular AR process, which leads to the following system

$$\chi(t) \quad = \quad \Lambda z(t), \tag{5.19}$$

$$a(z)z(t) \quad = \quad \varepsilon(t), \tag{5.20}$$

with $a(0) = I_q$ and $\Sigma = \text{Cov}(\varepsilon) > 0$.

Chapter 6

Influence analysis

In this chapter we discuss the causal analysis of high-dimensional time series and the associated problems. In particular we are interested in the causal analysis of ictal ECoG data in order to gain insights concerning the seizure onset zone.

As we will see this causal analysis is not straight forward. To overcome the arising problems we propose a method that is based on the Granger causal analysis of factor models. The application of this methodology to the ECoG data yields results which are in very good accordance with the clinical findings.

In section 6.1 we discuss the problems concerning the causal analysis of co-moving data. To deal with these problems we elaborate the statistical framework for the application of Granger causality to factor models in section 6.2 (based on the knowledge of GDFMs from section 5.1). The practical application of Granger causality to factor models is based on the PCA approach (from section 5.2) as described in section 6.3.

With the gathered knowldege we propose our causal investigation method in section 6.4. To assess our proposed methodology we apply it to simulated data in subsection 6.5.1. The most interesting part of section 6.5 is the application of our proposed method to the ECoG data.

The chapter is concluded by a thorough discussion of the methodolgy and the results in section 6.6. The discussion will highlight the statistical-technical properties of the proposed methodology as well as the clinical view of the results.

This chapter signifies the main contribution of this thesis as it merges the causal analysis, factor models and the neurophysiological aspects. Parts of this chapter have been published in Flamm et al. (2012a).

For background information and notation we refer to chapter 1 for the mathematical introduction, to chapter 3 for Granger causality and to chapter 5 for factor models and the PCA approach.

6.1   Challenge

What is the challenge in using Granger causality for the analysis of high dimensional co-moving data, in particular ECoG data? The naïve approach would be to fit an n-dimensional (regular) AR model to the n-dimensional observations $x(t)$. In practice we typically encounter two problems:

First, the curse of dimensionality impairs the AR-estimation, compare chapter 5. The fitting of an n-dimensional AR(p) model requires the estimation of $n^2 p$ parameters, but the number of observations is only linear in n. Thus, in order to obtain reliable estimation results a sufficient number of data samples $x(t), t = 1, \ldots, T$ is required, i.e. very long data samples are needed for a large cross-sectional dimension n. Normally neurological signals (in particular ECoG and EEG) show a highly non-stationary behavior (the second moments vary significantly within long data windows), so the stationarity assumption for the required long data windows would be violated. Of course, this impairs the estimation of the AR model for the Granger causal analysis.

Second, the ECoG data show strong signs of co-movement, which can be seen visually. This clearly indicates that the n-dimensional observations are originally generated by a small number of factors. Mathematically speaking, the most important part of the observations, the latent variables, is generated only by a small number of factors. As Granger causality analysis is usually based on regular AR estimation, where the number of observations is equal to the number of generating components (see section 3.4), this leads to a poor AR-estimation for the ECoG data. The subsequent causal analysis would also yield misleading results.

In order to cope with these problems we consider factor models. As we have seen in chapter 5, factor models are a useful tool for modelling high-dimensional co-moving systems. The important question is, which causalities can reasonably be analyzed in this context. The price to be paid when using factor models instead of regular AR modeling for data description is that the causal analysis becomes more difficult. A classic Granger causal investigation does not make sense for factor models, because by the definition of a factor model a regular AR-model is not applicable.

In this thesis we assume that the dependence structure of the latent variables χ properly reflects the causal structure of the observations x. Therefore, a causal analysis of the latent variables will reveal the dependencies between the observations themselves. This is a very strong assumption, but it seems meaningful despite the seperation (5.1) into noise and latent variables. We will thoroughly discuss this assumption in section 6.6. This assumption is reasonable, because in ECoG analysis we expect the most important dependencies between the components explaining the most important features of the signal, i.e. the latent variables.

The first idea for a causal analysis in the factor model case would be to consider relations of the form

$$\chi_i \overset{?}{\to} \chi_j | \chi_V. \tag{6.1}$$

In other words, how does χ_i influence χ_j with respect to χ_V. However, the use of the exhaustive set V as the conditioning set leads to problems: Relation (6.1) signifies the contribution of $\overline{\chi_i(t)}$ to $\chi_j(t+1)$ in the projection of $\chi_j(t+1)$ onto $\overline{\chi_V(t)}$, compare criterion (3.12). Although the projection itself is unique, the projection coefficients are not. This means that the contribution of $\chi_i(t), \chi_i(t-1), \ldots$ to the explanation of $\chi_j(t+1)$ is not uniquely determined. This is due to the fact that only a small number of components is needed to span $\overline{\chi_V(t)}$. As these contibutions are not unique the application of criteria (3.12), (3.13) and (3.17) is not reasonable. Thus an analysis involving relations of the form (6.1) is not meaningful and would yield misleading results.
It is very important to note that we consider the general case, because in special cases the above problem might not arise.

Therefore we restrict the conditioning set, instead of V we use *channel selections* $I \subset V$ where we consider only q channels, i.e. $\#I = q < n$.[4] Thus we consider relations of the form

$$\chi_i \overset{?}{\to} \chi_j | \chi_I \quad i, j \in I. \tag{6.2}$$

Of course we are searching for channel selctions I where the projection coefficients from $\chi_I(t+1)$ onto $\overline{\chi_I(t)}$ are unique, compare the aforementioned problem. In the following we will discuss the choice of the channel selection I, such that relations of the type (6.2) yield reasonable results in the causal analysis. For a theoretical discussion on the choice of I we refer to section 6.2, for the practical PCA-based approach we refer to section 6.3.

As we have discussed in chapter 4, all conditional independence relations between arbitrary sub-processes (including non-causality and contemporaneous non-correlation relations) have to be considered for the complete causal description of a system. However, as the bivariate causal relations with respect to the whole process have a special role among the non-causality relations (see section 4.3) our proposed method will be solely based on relations of the form (6.2).

6.2 Theoretical framework

As mentioned in section 6.1 we are interested in causality relations of the form (6.2). Here we discuss the choice of the channel selection I in order to yield reasonable results in the causal analysis.

---

[4]In this context # denotes the number of elements in a set.

In particular we are interested in q-dimensional selections $I \subset V$, where χ_I drives all latent variables χ_V. We propose this kind of channel selections because a simple Granger causal anlysis is possible in this case.

We start our causal considerations by continuing at the end of section 5.1. By reordering (5.6) we obtain

$$\begin{pmatrix} \chi_I(t) \\ \chi_J(t) \end{pmatrix} = \begin{pmatrix} w_I(z) \\ w_J(z) \end{pmatrix} \varepsilon(t) \tag{6.3}$$

where χ_I is a q-dimensional and χ_J is a $(n - q)$-dimensional process, so $w_I(z)$ is a $q \times q$ rational matrix.

We only consider channel selections I, where $w_I(z)$ is *causally invertible*, i.e. all its zeros and poles lie outside the unit circle. Then $w_I(z)^{-1}$ exists inside and on the unit circle, and ε can be expressed as

$$\varepsilon(t) = w_I(z)^{-1} \chi_I(t). \tag{6.4}$$

Using equation (6.4) we rewrite (6.3) as

$$\begin{pmatrix} \chi_I(t) \\ \chi_J(t) \end{pmatrix} = \begin{pmatrix} I \\ w_J(z) w_I(z)^{-1} \end{pmatrix} \chi_I(t), \tag{6.5}$$

which shows that χ_I generates the whole process χ_V. In more mathematical terms this means χ_I and ε span the same space, which is of course equal to the space spaned by the whole process

$$\overline{\chi_I(t)} = \overline{\varepsilon(t)} = \overline{\chi_V(t)}. \tag{6.6}$$

We note, that in the general case it might not always be possible to find a selection I of driving latent variables, e.g. consider the following system

$$\begin{pmatrix} \chi_1(t) \\ \chi_2(t) \end{pmatrix} = \begin{pmatrix} w_1(z) \\ w_2(z) \end{pmatrix} \varepsilon(t) = \begin{pmatrix} 1 + \rho_1 z \\ 1 + \rho_2 z \end{pmatrix} \varepsilon(t) \tag{6.7}$$

where $|\rho_1|, |\rho_2| > 1, \rho_1 \neq \rho_2$. In this example neither $w_1(z)$ nor $w_2(z)$ are causally invertible.

In this example it would be possible to rewrite χ_1 (or alternatively χ_2) as $\chi_1(t) = \tilde{w}_1(z) \tilde{\varepsilon}(t)$ with $\tilde{w}_1(z)$ causally invertible and $\tilde{\varepsilon}_1$ white noise, but χ_2 could not be causally generated by χ_1 in this case.

It is important to note that we only consider the transfer function $w(z)$ derived by the spectral factorization (5.5) according to theorem 5.1. Therefore, we only consider square selections $w_I(z)$ of this transfer function.

Our main reason for requiring channel selections with a corresponding causally invertible transfer function is the following: If there exists a channel selection I,

where $w_I(z)$ is causally invertible, we may investigate conditional Granger causal relations in $\chi_I(t)$ in an easy way.

Interpretation of equation (6.4) as an AR model and premultiplication with $w_I(0)$ (i.e. the leading coefficient of $w_I(z)$) gives

$$\underbrace{w_I(0)w_I(z)^{-1}}_{\breve{a}(z)} \chi_I(t) \;=\; \underbrace{w_I(0)\varepsilon(t)}_{\breve{\varepsilon}(t)} \,,$$

$$\breve{a}(z)\,\chi_I(t) \;=\; \breve{\varepsilon}(t)\,, \tag{6.8}$$

with $\breve{a}(0) = I_q$. This is a regular AR model of the form (1.7) for χ_I. Therefore, the Granger causal relations (6.2) are simply derived by the application of criterion (3.12), which reads the following in this context

$$\breve{a}_{ji}(z) = 0 \iff \chi_i \nrightarrow \chi_j | \chi_I.$$

This relation signifies the influence of χ_i to χ_j with respect to χ_I.

An important aspect we have to consider for our analysis is the following:

If there exist two channel selections $I \supset i, j$ and $\tilde{I} \supset i, j$, with $I \neq \tilde{I}$, where both $w_I(z)$ and $w_{\tilde{I}}(z)$ are causally invertible, the naturally arising question is, whether the causal relations $\chi_i \overset{?}{\rightarrow} \chi_j | \chi_I$ and $\chi_i \overset{?}{\rightarrow} \chi_j | \chi_{\tilde{I}}$ are equal. In other words, how does the choice of the channel selection I influence the causal relation between χ_i and χ_j?

Due to the fact that $\overline{\chi_I(t)} = \overline{\varepsilon(t)} = \overline{\chi_{\tilde{I}}(t)}$ (because they span each other, see equation (6.6)) one could conjecture that the aforementioned causality relations are equal. This would mean that the causal relations (6.2) between channel i and j are independent of the choice of a valid channel selection. This would have the advantage that a single arbitrary channel selection $I^* \supset i, j$ (with $w_{I^*}(z)$ causally invertible) is sufficient to explore the causality between the considered channels χ_i and χ_j.

However, unfortunately this is not the case. The causal relation $\chi_i \nrightarrow \chi_j | \chi_I$ can also be checked via criterion (3.17) which reads

$$\chi_j(t+1) \perp \overline{\chi_i(t)} | \overline{\chi_{I \setminus \{i\}}(t)}$$

in this context. This shows, that the conditioning set is $I \setminus i$ (not I!). Even though $\overline{\chi_I(t)} = \overline{\chi_{\tilde{I}}(t)}$ holds, in general $\overline{\chi_{I \setminus \{i\}}(t)} \neq \overline{\chi_{\tilde{I} \setminus \{i\}}(t)}$, i.e. the conditioning spaces are not equal. Thus, the conditional Granger causal relations are not the same (in general).

We want to emphasize the importance of this statement a second time. The causal relation $\chi_i \overset{?}{\rightarrow} \chi_j | \chi_I$ strongly depends on the channel selection I, because it is (in general) different from the relation $\chi_i \overset{?}{\rightarrow} \chi_j | \chi_{\tilde{I}}$ for valid $I \neq \tilde{I}$.
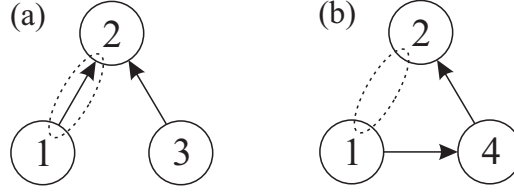
Figure 6.1: Associated path diagrams for the sub-processes $\chi_{\{1,2,3\}}$ and $\chi_{\{1,2,4\}}$ of system (6.9). Different causality relations between χ_1 and χ_2 depending on the considered channels: (a) $\chi_1 \rightarrow \chi_2 | \chi_{\{1,2,3\}}$ holds in sub-system (6.10), but (b) $\chi_1 \nrightarrow \chi_2 | \chi_{\{1,2,4\}}$ holds in sub-system (6.11).

For the interested reader we have a simple example showing that the causality relations depend on the channel selections. We want to investigate the causality from χ_1 to χ_2 in the following system

$$\begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \\ \chi_4 \end{pmatrix}(t) = \begin{pmatrix} 1 & 0 & 0 \\ \alpha z^2 & 1 & \beta z \\ 0 & 0 & 1 \\ \frac{\alpha}{\beta} z & 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}(t), \tag{6.9}$$

with $\alpha \neq \beta$, $\alpha, \beta \neq 0$ and $\mathrm{Cov}(\varepsilon) = I_3$. Simple calculations show that the sub-process $x_{\{1,2,3\}}$ has the following AR representation

$$\begin{pmatrix} 1 & 0 & 0 \\ -\alpha z^2 & 1 & -\beta z \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \end{pmatrix}(t) = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}(t). \tag{6.10}$$

For the sub-process $x_{\{1,2,4\}}$ we have

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\beta z \\ -\frac{\alpha}{\beta} z & 0 & 1 \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_4 \end{pmatrix}(t) = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}(t). \tag{6.11}$$

Figure 6.1 shows the associated path diagrams of these two sub-processes displaying their causality structure.

According to criterion (3.12), in sub-system (6.10) χ_1 is Granger causal for χ_2 due to $a_{21}(z) \neq 0$, hence

$$\chi_1 \rightarrow \chi_2 | \chi_{\{1,2,3\}}.$$

On the contrary, in sub-system (6.11) we have $a_{21}(z) = 0$, thus χ_1 is Granger non-causal for χ_2, hence

$$\chi_1 \nrightarrow \chi_2 | \chi_{\{1,2,4\}}.$$

So in one sub-system χ_1 is Granger causal for χ_2 and in the other sub-system it is not. This short example illustrates the fundamental fact that the causal relations

depend on the channel selection I.

For the sake of completeness we want to mention, that the proposed causal analysis between two sub-processes χ_i and χ_j is not possible, if there does not exist a set $J \supset i, j$ with a causally invertible $w_J(z)$. In this case an analysis of $w(z)$ might give further informations about the connection between these two processes.

Due to the fact that we only consider channel selections I where the corresponding χ_I generates the whole process, the observed causal relations $\chi_i \stackrel{?}{\to} \chi_j | \chi_I$ signify direct causality relations, see chapter 3. That means that the causal effects are direct effects and not mediated via a third variable (or a path of variables). This is a huge advantage of the considered causality relations (6.2).

At the end of this section we want to sum up its most important statements. In order to analyze the causality between two latent variables χ_i and χ_j we choose a channel selection $I \supset i, j$ where the corresponding transfer function $w_I(z)$ is causally invertible. In the context of a simple regular AR model for χ_I we can analyze the desired relation $\chi_i \stackrel{?}{\to} \chi_j | \chi_I$. For different channel selections the resulting causal relation between the two latent variables might differ.

6.3 PCA-based approach

Now we are interested in the practical derivation of the causal relations $\chi_i \stackrel{?}{\to} \chi_j | \chi_I$ discussed in section 6.1. This practical approach is based on the PCA approach presented in section 5.2. We commence from equation (5.18), which reads

$$\chi(t) = \Lambda z(t) = \Lambda a^{-1}(z)\varepsilon(t),$$

where $a(z)$ is the AR polynomial of the static factors z.
For a channel selection $I \supset i, j$, $\#I = q$ we consider the corresponding sub-system

$$\chi_I(t) = \Lambda_I z(t) = \Lambda_I a^{-1}(z)\varepsilon(t), \tag{6.12}$$

where Λ_I is the square sub-matrix of Λ correponding to the selected components χ_I.

As before, we have to search for meaningful channel selections I. In contrast to the previous section this problem simplifies significantly here. In this approach we only consider channel selections I where Λ_I is regular.
This criterion is (in this PCA approach) equivalent to the channel selection restriction to causally invertible $w_I(z)$ from section 6.2. If Λ_I is regular, $\Lambda_I a^{-1}(z) = w_I(z)$ is causally invertible, and vice versa. As a side note we mention that obviously $\overline{\chi_I(t)} = \overline{\chi_V(t)}$ holds in this approach, in other words χ_I generates the whole process χ_V.

For the remainder of this section we call channel selections I *admissible* if Λ_I is regular.

By rewriting (6.12) as an AR representation we obtain

$$\underbrace{\Lambda_I a(z) \Lambda_I^{-1}}_{\breve{a}(z)} \chi_I(t) \;=\; \underbrace{\Lambda_I \varepsilon(t)}_{\breve{\varepsilon}(t)} \tag{6.13}$$

$$\breve{a}(z)\chi_I(t) \;=\; \breve{\varepsilon}(t),$$

where $\det(\Lambda_I) \neq 0$. Note that we premultiply with Λ_I in order that the leading coefficient of the left-hand side polynomial becomes the identity, $\breve{a}(0) = I_q$.

The Granger causal relations for the AR system in representation (6.13) can easily be checked by criterion (3.12), which reads

$$\chi_i \nrightarrow \chi_j | \chi_I \iff \breve{a}_{ji}(z) = 0 \;\; i, j \in I; i \neq j$$

in this context. So in this PCA-based approach it is easy to derive the desired causal relations.

We want to emphasize, that the PCA-based approach is a special case of the general case discussed in section 6.2. It also holds in this case, that distinct channel selections I and $\tilde{I}$ (in general) yield different causal relations between χ_i and χ_j, which can be seen in the following example.
Consider the following system

$$
\begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \\ \chi_4 \end{pmatrix}(t) \;=\; \overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}}^{\Lambda} \overbrace{\begin{pmatrix} 1 & 0 & \alpha z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{a(z)^{-1}} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}(t) \;=\;
$$

$$
= \begin{pmatrix} 1 & 0 & -\alpha z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 - \alpha z \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}(t) \tag{6.14}
$$

with $\mathrm{Cov}(\varepsilon) = I_3$. The sub-process $\chi_{\{1,2,3\}}$ has the following AR representation

$$
\begin{pmatrix} 1 & 0 & -\alpha z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \end{pmatrix}(t) = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}(t), \tag{6.15}
$$

and for sub-process $\chi_{\{1,2,4\}}$ we have

$$
\begin{pmatrix} 1 - \alpha z & -\alpha z & \alpha z \\ 0 & 1 & 0 \\ -\alpha z & -\alpha z & 1 + \alpha z \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_4 \end{pmatrix}(t) = \begin{pmatrix} \breve{\varepsilon}_1 \\ \breve{\varepsilon}_2 \\ \breve{\varepsilon}_3 \end{pmatrix}(t), \tag{6.16}
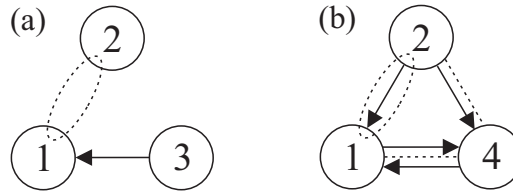$$

Figure 6.2: Associated path diagrams for the sub-processes $\chi_{\{1,2,3\}}$ and $\chi_{\{1,2,4\}}$ of system (6.14). Different causality relations between χ_1 and χ_2 depending on the considered channels: (a) $\chi_2 \nrightarrow \chi_1|\chi_{\{1,2,3\}}$ holds in sub-system (6.15) (b) $\chi_2 \rightarrow \chi_1|\chi_{\{1,2,4\}}$ holds in sub-system (6.16)

with

$$\text{Cov}(\breve{\varepsilon}) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix}. \tag{6.17}$$

According to criterion (3.12), in sub-system (6.15), χ_2 is Granger non-causal for χ_1 due to $a_{12}(z) = 0$. In sub-system (6.16) $a_{12}(z) \neq 0$, thus χ_2 is Granger causal for χ_1. Compare figure 6.2 for a graphical representation of these two sub-systems.

To conclude this section we want to mention one advantage of this PCA-based approach, it is computationally efficient. After the PCA computation only one AR system $a(z)$ for the static factors z has to be fitted. Linear transformations according to (6.13) are sufficient for determining the AR systems for all possible channel selections.

For the remainder of this chapter we are interested in a practical causal analysis of high-dimensional co-moving systems. Therefore, the following sections will all be based on the practical PCA approach discussed in this section rather than the theoretical approach discussed before.

6.4 Methodology

As we have seen in the previous section, a Granger causal investigation between channels i and j depends on the chosen channel set I. Different channel selections I and $\tilde{I}$ (in general) lead to different results. However, we are interested in a single overall statement, whether channel i influences channel j or not.

For this purpose we propose an intuitive methodology, which is based on the ideas the PCA-based approach of section 6.3.

Our methodology consists of three steps. First, we use PCA to separate the observations into the latent variables (explaining the co-movement) and the noise, see section 5.2.

Second, for fixed i and j we analyze the conditional Granger causality relation $\chi_i \overset{?}{\to} \chi_j | \chi_I$, given a fixed channel selection $I \supset i, j$.

Third, we perform this analysis for all admissible channel selections $\tilde{I} \supset i, j$ and derive a heuristic statement for the influence from χ_i to χ_j, condensing the information of all sub-systems.

In detail we proceed as follows:

Initially we perform a PCA on the observations x in order to obtain the factor loading matrix Λ and the static factors z. The dimension of the static factors q is determined via a Scree plot, see Cattell (1966).

Now let i, j, I be fixed. The straightforward application of the PCA approach from section 6.3 yields two problems:

While in theory we can easily distinguish regular and singular matrices Λ_I in equation (6.12) by considering the determinant, the estimator $\hat{\Lambda}_I$ will typically yield $\det(\hat{\Lambda}_I) \neq 0$. The causality relations drawn from systems with very small values of $|\det(\hat{\Lambda}_I)|$ are not meaningful, which is due to the fact that $\breve{a}(z)$ in (6.12) cannot be estimated reliably. The term $|\det(\hat{\Lambda}_I)|$ is a measure for the similarity of the selected channels. Therefore, we only consider channel selections I with $|\det(\hat{\Lambda}_I)|$ exceeding a threshold τ.

A similar challenge arises in the estimation of $\hat{\breve{a}}_{ji}(z)$ (which has a finite order now). In theory $\breve{A}_{ji}(m) = 0 \; \forall m$ signifies that χ_i is Granger non-causal for χ_j. However, in estimation we typically have $\hat{\breve{A}}_{ji}(m) \neq 0$, so we have to statistically test whether the polynomial coefficients $\hat{\breve{A}}_{ji}(m)$ (for all lags m) are significantly jointly different from zero. For this purpose we use an F-test ($\mathcal{H}_0 : \breve{A}_{ji}(m) = 0 \; \forall \; m$), which is implemented in the *GCCA toolbox*, in Seth (2010) the toolbox and the test are described. We consider the p-value of the test as a measure for Granger causality: Rejection of $\mathcal{H}_0$ ($p < 0.03$) signifies Granger causality, acceptance means non-causality. The threshold value $p_0 = 0.03$ was chosen empirically.

In order to sum up, for each channel selection I (for fixed i, j) we obtain two values: $|\det(\hat{\Lambda}_I)|$ as a similarity measure of the channels in I and the p-value as an indicator for the causality from χ_i to χ_j.

As a global influence statement from χ_i to χ_j is our goal, we want to condense the different conditional causality statements based on distinct channel selections I into a single one. For this purpose we propose an intuitive rule: If all statements for distinct channel selections match, we conclude a global influence statement.

In other words: if $\chi_i \to \chi_j | \chi_I$ for all I with $|\det(\hat{\Lambda}_I)| > \tau$, we say that χ_i *influences* χ_j. On the other hand, if $\chi_i \nrightarrow \chi_j | \chi_I$ for all I with $|\det(\hat{\Lambda}_I)| > \tau$, we say that χ_i *does not influence* χ_j. In case of non-conclusive Granger causality statements we do not derive any global influence statement.

Finally, as the causality structures of the observations and the latent variables are equal by assumption, compare section 6.1, we say x_i *influences* x_j if χ_i influences χ_j. The analog reasoning holds in case of non-influence.
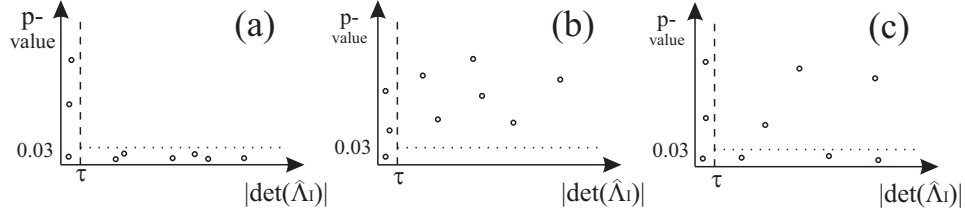
Figure 6.3: Visualization of the proposed method. Analysis for all causality relations $\chi_i \rightarrow \chi_j | \chi_I$ for distinct channel selections $I \supset i, j$. In each plot a point shows the p-value (as a measure of causality) and $|\det(\hat{\Lambda}_I)|$ (as a measure of channel similarity) for the respective channel selection I. Only points with $|\det(\hat{\Lambda}_I)| > \tau$ are considered in the analysis (numerical reasons). (a) All relevant points have an associated p-value < 0.03, i.e. indicate causality (for each respective I). We conclude that x_i *influences* x_j. (b) All relevant points have an associated p-value > 0.03, i.e. indicate non-causality (for each respective I). We conclude that x_i *does not influence* x_j. (c) For different I, causality as well as non-causality statements are indicated. We do not conclude any influence statement.

For a better understanding we want to visualize the described methodology: For i, j fixed we plot a point for each distinct channel selection $I \supset i, j$ into the plane spaned by $|\det(\hat{\Lambda}_I)|$ on the x-axis and the p-value on the y-axis. This procedure yields graphs such as shown in figure 6.3. In such a plot we only consider points with $|\det(\hat{\Lambda}_I)| > \tau$, in other words points which are located to the right of the dashed vertical threshold line. Points to the left of this determinant threshold line are ignored, because the corresponding p-values are not meaningful due to numerical instabilities.

A point situated below the dotted line represents a p-value < 0.03 and therefore indicates Granger causality. Consequently a point lying above the dotted line indicates Granger non-causality.

Figure 6.3 shows three plots constructed in the aforementioned way, it illustrates the three cases we distinguish:

In plot (a) all relevant points are situated below the dotted line, i.e. each point individually indicates causality ($\mathcal{H}_0$ of non-causality rejected due to $p < 0.03$), thus we have global influence.

We observe the opposite situation in plot (b), where all relevant points are above the dotted line, i.e. each point individually indicates non-causality, so we speak of global non-influence.

Plot (c) illustrates a situation where distinct channel selections lead to causality as well as non-causality statements. In this case, we refrain from concluding on global influence.

6.5   Results

In this section we apply our proposed method to data. First, we assesss our methodology with a signal model. Second, we apply the proposed method to the ECoG data. In a subsequent step we use the results to draw conclusions concerning the seizure onset zone.

6.5.1   Simulated data

In order to assess the proposed methodology we apply it to simulated data where we know the imposed dependence structure. In the first part of this subsection we explain the signal model and in the second part we present the results of the application of our methodology to the simulated data.
Consider the following signal model

$$x(t) = \Lambda z(t) + \eta(t) \tag{6.18}$$
$$a(z)z(t) = \varepsilon(t).$$

First we simulate the 3-dimensional static factors z as an AR(2) process with the following AR polynomial

$$a(z) = \begin{pmatrix} 1 - 0.2z & 0 & 0 \\ -0.3z^2 & 1 - 0.5z & 0 \\ -0.7z^2 & 0 & 1 - 0.5z \end{pmatrix}$$

and $Cov(\varepsilon) = I_3$.[5] For the construction of x we choose the factor loading matrix

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and the variance of the noise

$$Cov(\eta) = \text{diag}(0.15, 0.15, 0.61, 1.37,$$
$$0.61, 0.15, 1.37, 1.37, 0.61).$$

The Granger causal structure of the simulated 3-dimensional static factors $(z_{1^*}, z_{2^*}, z_{3^*})'$ is depicted in figure 6.4(a), the resulting influence structure of the 9-dimensional system $(x_1, \ldots, x_9)'$ is shown in figure 6.4(b). Due to the simple structure of the
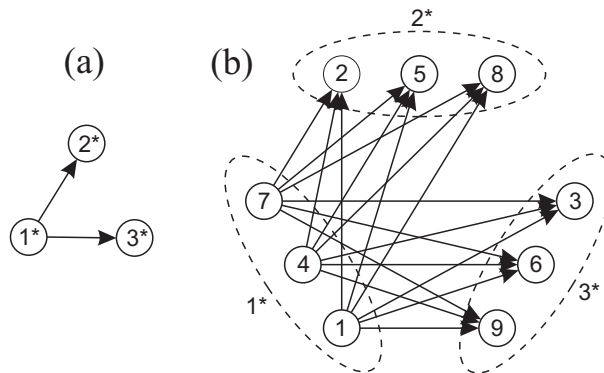
Figure 6.4: Illustration of the imposed dependence structure of signal model (6.18). (a) Associated path diagram for the static factors z. (b) Graph depicting the influence structure of the simulated observations (arrows indicate influence).
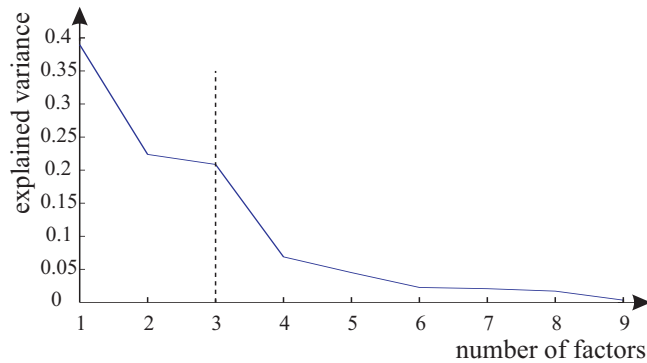

Figure 6.5: Scree plot of the principal component analysis of the simulated data from signal model (6.18). Three factors explain the majority of the variance.

loadings matrix Λ we get a simple influence structure in figure 6.4(b).

We now apply our proposed methology to the simulated data from the signal model (6.18).

For the initial calculation of the PCA, we determine the number of static factors q by considering the Scree plot, see figure 6.5. This figure shows the percentage of the explained variance per factor. We observe that three factors explain the majority of the variance, thus we choose $q = 3$. Furthermore, by application of the BIC criterion we obtain an AR-model order of $p = 2$ (matching the imposed model order).

Proceeding according to our methodology, for fixed source i and target j we obtain causality relations for all channel selections $I \supset i, j$. They are represented as points

---

[5]Simulation is done using the function arsim of the Matlab® package arfit, described in Schneider and Neumaier (2001).
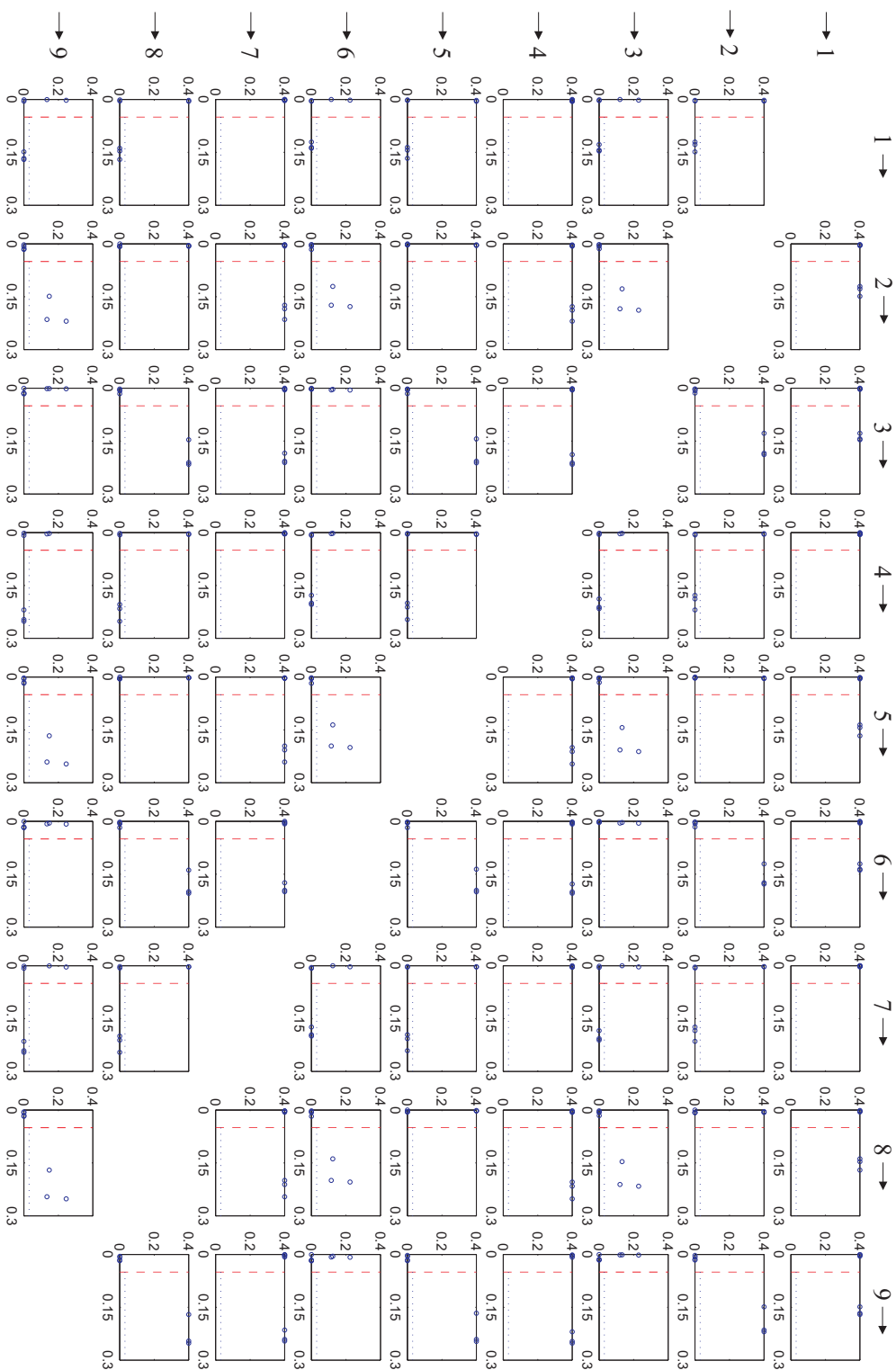
Figure 6.6: Results of the influence analysis of signal model (6.18) in matrix plot form. Columns indicate the source channels x_i and the rows the target channels x_j, the (j, i)-sub-plot quantifies the influence from x_i to x_j. Interpretation of each sub-plot as in figure 6.3 retrieves the imposed influence structure illustrated in figure 6.7.
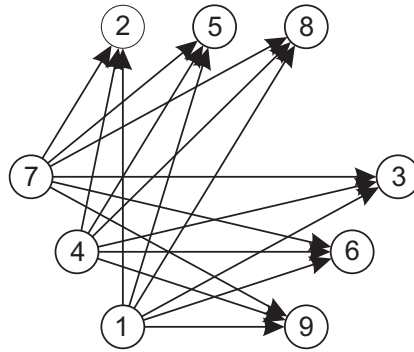
Figure 6.7: Inferred influence structure of the simulated data of signal model (6.18). The influence structure matches the imposed influence structure depicted in figure 6.4(b).

in a graph as described in section 6.4. Hereby points with p-values > 0.4 are displayed with p-value = 0.4, because this does not change the results of the analysis and facilitates the visualizaton.

In figure 6.6 all these plots are arranged in a 9×9 matrix plot, where the columns indicate the source channels x_i and the rows the target channels x_j. Thus, the (j, i)-sub-plot quantifies the influence from x_i to x_j. Diagonal elements are not displayed.

Let us consider the interpretation of three selected sub-plots in figure 6.6 in detail:
In sub-plot (3,1) all points to the right of the determinant threshold line are located below the dotted line, and therefore represent p-values smaller than 0.03 (i.e. the null hypothesis of non-causality is rejected). This means that for all admissible channel selections I, we have $\chi_1 \rightarrow \chi_3|\chi_I$. Thus we say x_1 influences x_3.
In sub-plot (3,2) all points to the right of the determinant threshold line are located above the dotted line. Thus we say x_2 does not influence x_3.
In sub-plot (4,1) all points are located to the left of the determinant threshold line, therefore we do not draw any conclusions. The reason for this behavior is that x_1 and x_4 are both generated by z_{1*} and therefore are highly correlated.

Summing up the information of figure 6.6 we retrieve the influence structure as depicted in figure 6.7. This is exactly the imposed influence structure of system (6.18) as shown in figure 6.4(b). Channel x_1 influences channels $x_2, x_3, x_5, x_6, x_8, x_9$, so do channels x_4 and x_7.
For the simulated data the proposed methodology performed very well.

6.5.2 ECoG data

In order to demonstrate the practicability of our proposed methodology we apply it to the ECoG data described in section 2.7. In this thesis we are interested in the seizure onset zone of an epileptic seizure. Therefore, we analyze the dependence
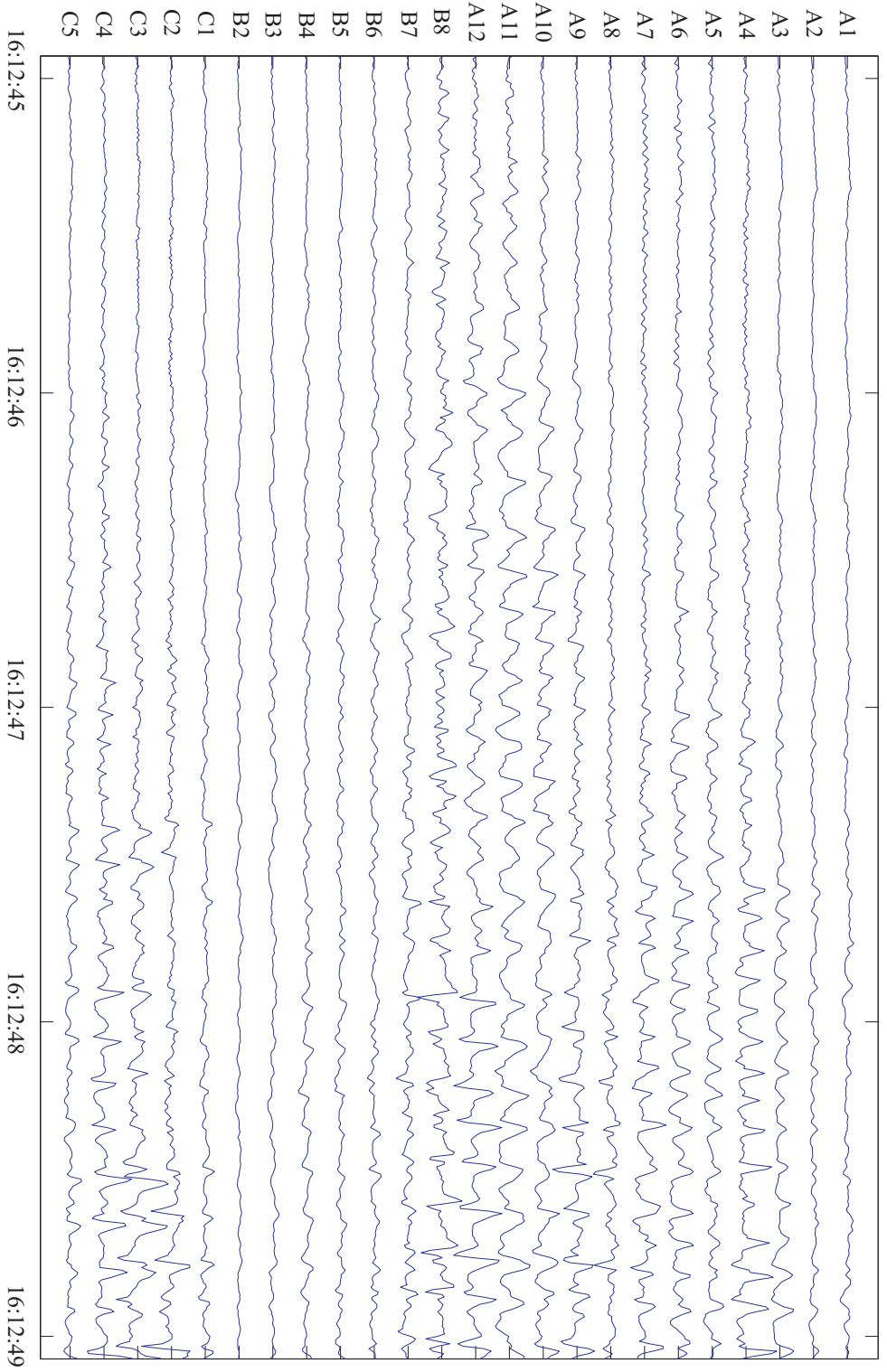
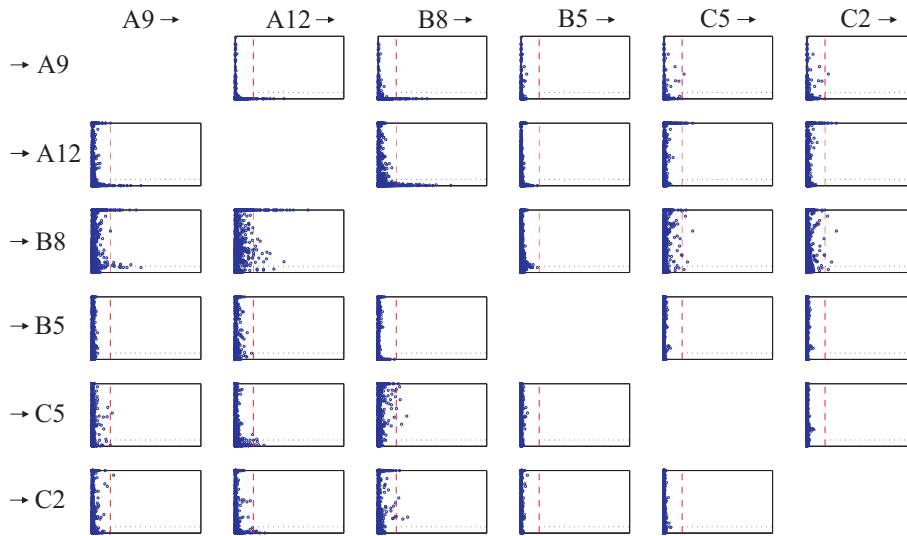Figure 6.8: Ictal ECoG data. Four second segment from the initial phase of the first seizure.

Figure 6.9: Results of the influence analysis of the ictal ECoG data. Six selected channels (out of the whole 24 channel set) are dsiplayed. Columns indicate the source channels and the rows the target channels. Interpretation of each sub-plot as in figure 6.3 yields (a part of) the influence structure depicted in figure 6.10.

structure of the data at the beginning of the seizure. In particular we process a 4 second time window at the beginning of the first recorded seizure of the patient. The starting time is indicated by the clinicians, and the data are shown in figure 6.8.

As we will see, the influence structure of the data at the beginning of the seizure will give indications about the seizure onset zone.

In the first step we calculate the PCA of the data. The number of static factors is determined by a Scree plot as before in subsection 6.5.1. In order to achieve an explained variance greater than 80 % we choose $q = 5$. Furthermore, in accordance with the survey paper Tseng et al. (1995), we set the AR-model order for the Granger causal analysis of the ECoG data to $p = 8$.

In the second step we fix the source channel i and the target channel j and compute the causal relations for all channel selections $I \supset i, j$. Following our proposed methodology we obtain a 24×24 matrix plot. Figure 6.9 shows a 6×6 sub-matrix, corresponding to the channels A9, A12, B8, B5, C5, C2. The whole 24×24 matrix plot would have been to large to properly display it here. For explanatory purposes the 6×6 sub-plot will suffice.

We briefly want to discuss 4 sub-plots of figure 6.9 in detail to highlight important characteristics:

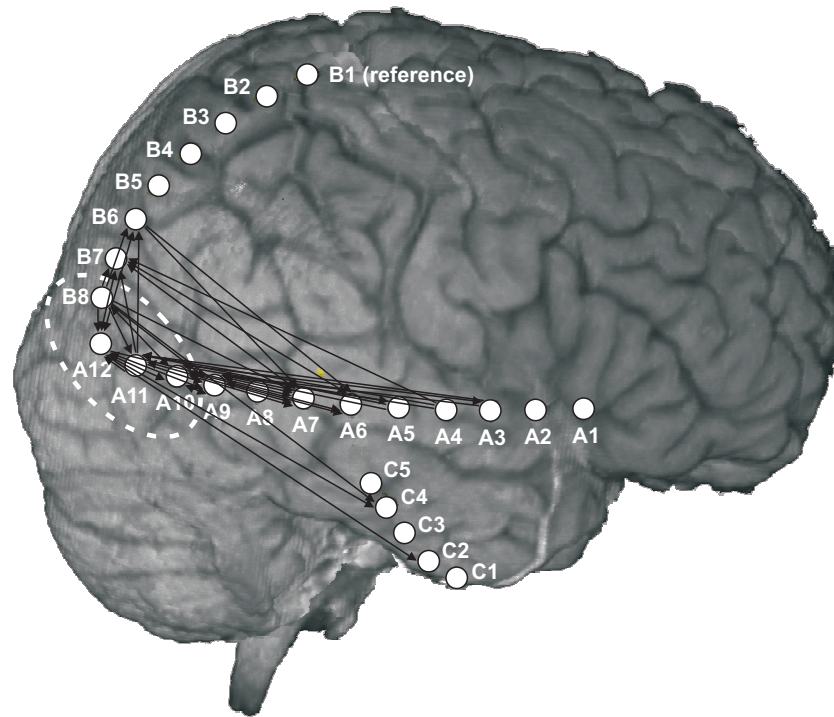Sub-plot (2,3) describes the causality relations from B8 to A12. All points to the

Figure 6.10: Results of the influence analysis of the ictal ECoG data. Illustration of the influence structure of the ECoG data, displayed in figure 6.8. MRI scan of the patient's brain together with the subdural electrode strip positions. Arrows indicate influence between the respective electrodes. Electrodes with the highest number of departing arrows are considered to represent the seizure onset zone (B8, A12, A11, A10), compare the out-degree histogram in figure 6.11.

right of the determinant threshold are located below the dotted line, thus we say channel B8 influences channel A12.

In sub-plot (5,3) all points located to the right of the determinant threshold are above the dotted line. Therefore B8 does not influence C5.

An interesting case occurs in sub-plot (3,2). We have admissible points above and below the dotted line. In this case we refrain from any influence statement.

Finally in sub-plot (4,2) all points are located to the left of the determinant threshold. We do not draw any conclusion in this case, as there are no admissible channel selections.

By interpreting each sub-plot in the 24×24 matrix in this way, we obtain all influence relations. In particular we are interested in the (source, target) channel pairs, where the source does influence the target channel. Figure 6.10 shows the derived influence structure of the ECoG data. For a better understanding the arrows (which indicate influence) are drawn into a MRI scan of the patient's brain together with
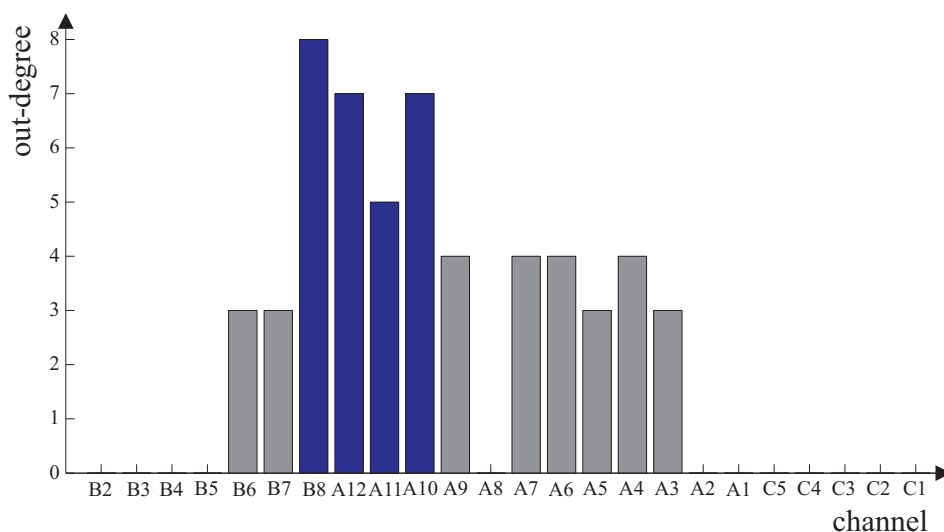
Figure 6.11: Results of the influence analysis of the ictal ECoG data. Histogram of the out-degree of the electrodes shown in figure 6.10. Electrodes with the highest number of departing arrows are considered to represent the seizure onset zone (B8, A12, A11, A10)

the electrode positions.

In this thesis we are interested in the seizure onset zone, in other words the brain area that triggers the epileptic seizure. Analyzing figure 6.10 we see that the channels B8, A12, A11 and A10 have the highest number of outgoing arrows. This can also be seen in figure 6.11, which shows the out-degree per channel. This observation suggests that the seizure onset zone comprises these four electrodes, which is in good accordance with the visual analysis of the clinicians. Of course this is a fundamental statement and in section 6.6 we will discuss the neurophysiological aspects of this result.

6.6 Discussion

In this chapter we proposed a procedure for deriving influence relations in high-dimensional co-moving time series using PCA and Granger causality. In this section we discuss various aspects of this approach. We start with a discussion of the proposed methodology followed by a discussion of its application to simulated data and ECoG recordings.

6.6.1 Theory and methodology

A key assumption of this method is that the latent variables reflect the causality structure of the observations. In other words, we can infer the dependencies be-

tween the channels based on the dependencies from the latent variables, the noise is assumed not to contain any causal information. Although this is a strong assumption, it seems reasonable to us: This assumption is very similar to the one that the important causal information is contained in the high amplitude wave forms, not in the small background noise. We believe that in particular in neurophysiological applications this is meaningful, as we expect these high-amplitude oscillations to carry substantial information about the causality structure of the generating cortical mechanisms.

To assess this assumption we compared the relations based on the observations with the causal relations based on the latent variables. Although small changes in the influence structure occured, the seizure onset zone did not change. This shows that this assumption is justified in this context.

In our approach we investigate all dependencies in all q-dimensional sub-systems of the latent variables. By the application of the PCA we perform this analysis in a computationally efficient way, as we only have to compute one PCA and estimate one single AR-model. Of course it would also be possible to refrain from the use of factor models. In this case one would fit an AR model for each q-dimensional sub-process of the observations, which would result in a higher computational effort.

In contrast, the proposed methodology yields a simple mathematical method for the causal analysis of the whole multivariate system. Another advantage of our method lies in the simple measure for the channel similarity of a channel selection I, $|\det(\hat{\Lambda}_I)|$, which allows for a straightforward comparison for different channel selections.

As mentioned in subsection 6.4, our methodology consists of three steps: PCA, Granger causality analysis for fixed channel selection I and derivation of an influence statement. This modular design allows for an easy adaptation of each single step, i.e. alternative methods could be used in each step independently of the others.

First, the use of sparse PCA would enforce additional zeros in the loading matrix Λ. Thurstone (1947) suggested five criteria for a simple structure and d'Aspremont et al. (2007) gave a direct formulation for sparse PCA.

Furthermore, we proposed the use of a static PCA in this chapter. The application of a dynamic PCA would also be conceivable, see Brillinger (1981). Especially in connection with the physiological frequency bands of the brain this would be a very interesting idea.

In the second step we use two central indicators in the Granger causal analysis for a fixed channel selection: $|\det(\hat{\Lambda})_I|$ as a measure of channel similarity and the p-value of an F-test as an indicator for Granger causality. We employ the latter due to its well-established theory. Note that in neuroscience literature various other directed dependence measures are used. Prominent ones include the lineaer measure of conditional dependence introduced by Geweke (1984), the partial directed coherence (PDC) introduced by Baccala and Sameshima (2001), its numerous mod-

ifications and the directed transfer function (DTF) proposed by Kaminski and Bli-nowska (1991). Compare section 3.5 and Flamm et al. (2012b) for an overview and a discussion of these measures with regard to Granger causality.

Other measures for the channel similarity would also be conceivable, e.g. the determinant of the covariance matrix of the errors $|\det(\hat{\Sigma}(\breve{\varepsilon}))|$ in (6.13). However, we note that $|\det(\hat{\Lambda}_I)|$ revealed good numerical properties in the simulations.

Third, in this chapter we propose an intuitive rule for the derivation of influence statements: If $\chi_i \rightarrow \chi_j | \chi_I$ for all admissible I, we say that x_i influences x_j. One could imagine other rules depending on specific applications. For example only the channel selection with the largest $|\det(\hat{\Lambda})_I|$ could be taken into account for the influence statement. Another possible rule for influence statements might be based on the comparison of the number of points above and below the dotted line, e.g. the majority or a certain percentage.

We focused our attention on the PCA-based approach. It would also be possible to base the proposed methodology on the causal relations directly derived from the transfer function $w(z)$ as described in section 6.2.

A naturally arising question is whether and to which extent our definition of influence is meaningful. In our opinion it is an intuitive and workable procedure for causality analysis in high-dimensional co-moving systems: Intuitively one expects a certain kind of dependence between x_i and x_j if χ_i is causal for χ_j for all (admissible) channel selections. In other words the most important parts of x_i are causal for the most important parts of x_j.

A potential weakness of our definition of influence is the fact that in practical applications one is often confronted with the case where no influence statement can be inferred. Compare sub-plot (3,2) in figure 6.9, where causality as well as non-causality relations are symbolized. In such cases we recommend a more precise investigation which particular channel selections yield causality relations and which do not.

If, in applications, the conditioning on channels of a specific (brain) region yields non-causality between channels i and j, but the conditioning on all other regions indicates causality, further investigation could be performed to explain this. One might conjecture that there is a separating region between i and j.

Furthermore, there could be some clustering in the inference plots. This fact might also yield additional information over the system.

Due to the existence of the aforementioned undecidable cases we avoid the term *causality* and refer to the derived dependence statement as *influence*.

We want to conclude this first part of the discussion with two remarks.

First, the problem of high-dimensional co-moving data often occurs in practical applications. We propose a methodology connecting Granger causal analysis and factor models.

One of the most prominent ideas to cope with co-movement is to analyze only a manually selected sub-system (where the number of observations equals the num-

ber of driving components). Another method would be the causal analysis of so
called extracted atoms, which represent condensed information of the system, see
Eichler (2005).

The selection of such low-dimensional sub-systems can also be performed in an
automatic way, e.g. the channel selection method proposed in Graef et al. (2012b)
based on the the An-algorithm described in An and Gu (1989).

Another possibility is to use penalized regression for an estimation of the AR-
model including all observations, e.g. LASSO as in Tibshirani (1996). This method
enforces additional zeros in the AR polynomial and avoids numerical problems.
Only recently, Chiang et al. (2009) successfully applied this approach to neural
data, calculated the PDC and visualized the indicated brain connectivity.

The idea behind most of these approaches is to compress the information of the
data. But in a subsequent causal analysis one has to be careful how to conclude
causal statements for the observations from the causal statements of the com-
pressed information. In our methodology we explained each step and the reasoning
behind it thoroughly.

Second, for a Granger causal analysis the AR representation of a process is some-
how a natural starting point, because the causalities can be easily checked via cri-
terion (3.12). However, in the factor model case a representation in the form of
(1.6) is often not meaningful (high number of observations, low number of driving
components). For the latent variables an MA representation of the form (5.6) is
more natural. Our methodology suggests to rewrite parts of this MA representa-
tion in order to use the classic AR representation theory. In simple cases causal
conclusions can be drawn from MA representations, see Sims (1972).

We based our methodology on regular AR models for sub-processes. However, it
is conceivable to use a singular AR model for the whole process. In this case one
has to check the uniqueness of $a(z)$ in order that the use of Granger causality makes
sense.

In this thesis we use the concept of Granger causality as the basic concept of causal-
ity. This is reasonable because Granger causality is intuitive, workable and well
understood. However, for the causal analysis of high-dimensional co-moving sys-
tems other definitions of causality might be conceivable.

6.6.2   Results

In this subsection we discuss the results obtained by our method. Especially we
are interested in the clinical interpretation of our results.

Our methodology correctly identifies the dependence structure of the signal model
(6.18). During this analysis we encountered the challenges discussed in subsection
6.6.1.

The application of our methodology to ictal ECoG data shows promising results.
We succesfully localize the seizure onset zone of the analyzed patient by identi-

Table 6.1: Onset zone and initial propagation of the analyzed seizure according to the visual inspection by clinicians.

Investigator	Initial Electrodes	Close follow-up
Expert 1	B8	A10, A11, A12
Expert 2	A11, A12, B8	A9, A10, B7
Expert 3	A10, A11, A12	B8

fying the zone with the highest number of outgoing arrows. In this manner we define the area comprising the electrodes B8, A12, A11 and A10 as the seizure onset zone, compare the MRI scan in figure 6.10 and the out-degree histogram in figure 6.11.

This result correlates very well with the visual inspection of the raw ECoG data by the clinicians. In chapter 7 figure 7.4 shows the initial time of the first seizure, this figure includes the beginning of the epileptic activity per channel marked by three clinical experts. Table 6.1 summarizes the findings of the three clinical experts who independently marked the electrodes initially involved in the epilpetic activity, which characterize the seizure onset zone.

For each of the three investigations the electrodes identified by our methodology are comprised in the set of initial and follow-up electrodes. Electrodes B8, A12 and A11 are specified as initial in two out of three cases and as follow-up in the third case. Inversively, electrode A10 is indicated as initial in one case and as follow-up in two cases.

In our opinion the reason for this good correlation between our results and the clinical findings is the following:

In case of focal epilepsy, the pathological synchronous activity (characterizing the epilepetic seizure) starts at a circumscribed brain area. From this seizure onset zone ictal activity spreads to its immediate vicinity recruiting more and more parts of the neural network. This leads to distinct co-movement of the observations. One could imagine a *focus* located in the seizure onset zone driving the surrounding channels by imposing its oscillatory frequency in the course of the recruiting process. This could be interpreted as a kind of information transfer or causal interaction. The electrodes in the focus causally influence the behavior of the surrounding electrodes in the initial phase of the seizure.

Therefore, we expect to obtain indications for the seizure onset zone by applying a Granger causal analysis to factor models during the initial seconds of the seizure. We think that the aforementioned results strengthen this hypothesis.

In the course of this recruiting process we obviously expect feedback mechanisms between the channels (besides unidirectional dependence). This can be observed in figure 6.10, consider e.g. channels A9 ↔ A10, A10 ↔ A12 and B6 ↔ A12.

However, in the seizure onset zone the departing arrows dominate, in other words the channels in the conjectured seizure onset zone have a high out-degree. This situation is reflected in the out-degree histogram in figure 6.11. Channels with the highest out-degrees coincide with the seizure onset zone (as indicated by the clinical experts), and with increasing distance to the seizure onset zone the respective out-degree decreases. Channel A8 with an out-degree of zero is an exception, as we cannot infer any influence statement for this source channel (only non-admissible channel selections for all target channels).

6.7   Conclusion

In this chapter we proposed a causal analysis of high-dimensional co-moving data, connecting the topics of Granger causality and factor models. The application of Granger causality to factor models is not straightforward, and we proposed a natural extension termed influence. Besides the theoretical considerations we also applied our methodology to ECoG data.

The application of our proposed method shows good first results in the detection of the seizure onset zone, because the results correlate well with the visual inspection of the clinicians.

Concluding we think that our proposed methodology might have the potential to assist clinicians in the presurgical evaluation by objectivating their visual ECoG examination.

Chapter 7

Band power measure

In this chapter we present an alternative method for the detection of the seizure onset zone of epileptic seizures. In chapter 6 we presented a rather theoretical-based approach with an application to the ECoG data, whereas in this chapter the introduced methodology will be more applied.

The seizure onset zone will be determined by the channels showing the first rhythmic ϑ-activity (4-9Hz). The application of this method yields results which are in good accordance with the visual characteristics of the ECoG data as well as the clinical findings based on the visual analysis of the clinicians.

After we motivate the method in section 7.1, we explain its two steps in section 7.2. Then we apply the procedure to the ECoG data in section 7.3 and discuss the results in section 7.4.

In contrast to the preceding chapters this methodology is only meant for application to ECoG data, there will be no theoretical discussions. Parts of this chapter have been published in Graef et al. (2012a).

7.1 Motivation

As we mentioned in section 2.4 the ECoG analysis in order to locate the seizure onset zone is done visually by the clinicians. Our methodology is partly based on the ideas of this visual analysis, therefore we will explain its basic concepts.

In the first step the clinicians temporally locate the seizure and mark its beginning and ending. Then they backtrace the beginning of the epileptic activity on each channel. For a patient suffering temporal lobe epilepsy (TLE) this normally means to backtrace the ϑ-activity to its temporal origin on each channel. Following this procedure the clinicians get the temporal onset of the epileptic activity for each observed channel. This information shows the propagation of the epileptic seizure. The seizure onset zone is considered as the area containing the channels with the earliest epileptic activity in this context.

However, the visual analysis is not straightforward due to practical reasons. The finding and the identification of epileptic activity (ϑ-activity for TLE patients) on each channel is a difficult and time-consuming task. Especially marking the temporal beginning of the epileptic activity is complicated, because there are a variety of different characteristics for the epileptic activity. Furthermore, the starting point is often not clearly identifiable because the normal brain activity interferes with the first epileptic signs and therefore impairs the analysis. Due to these reasons the visual analysis has to be carried out by experienced persons.

In the case of TLE the visual analysis simplifies slightly because the main epileptic activity are ϑ-waves, compare section 2.6 for more information on on temporal lobe epilepsy. The temporal onset of ϑ-activity is easier to identify than in the general case.

As we already mentioned our aim is to identify the seizure onset zone for TLE patients. Therefore, we propose a method that is based on the ideas of the visual analysis. It is physiologically motivated and focuses on the propagation of the ϑ-waves.

Our method consists of two consecutive steps:

1. In the first step we segment the ECoG data based on relative frequency contributions. The ECoG data and especially their spectral frequency contributions show instationary behavior. We propose a statistics, termed band power measure (BPM), which we employ for the segmentation. In subsection 7.2.1 we explain the segmentation in detail.

2. Based on this segmentation we classify each segment with respect to its epileptic character. Segments showing dominant ϑ-activity are said to be epileptic.

   The temporal delay of the onset of epileptic activity on the different channels is an indicator for the seizure propagation. Therefore, the channels showing the first epileptic activity are said to belong to the seizure onset zone. In subsection 7.2.2 we detail the segment classification and the analysis of the onset zone.

So in the first step we segment the data, whereas in the second step we classify the segments and infer the seizure onset zone. The first step is called segmentation and the second step is called onset zone analysis in the following.

For both steps or rather the concepts behind them, the literature offers a wide variety of related ideas:

In order to cope with the instationarities of EEG and ECoG signals, various segmentation methods have been developed in the last decades: The prominent *spectral error measure (SEM)* was introduced in Bodenstein and Praetorius (1977), a non-linear energy operator was used in Wu and Gotman (1998) and a generalized Kolmogorov-Smirnov-statistics in Brodsky et al. (1999). Other segmentation

approaches involve the use of information criteria in Inouye et al. (1995) or the Itakura distance in Kong et al. (1997). The various segmentation methods are normally rooted on a statistics based on ECoG properties. Large changes in the statistics indicate the beginning of a new segment.

The classification of the epileptic character is closely related to seizure detection analysis. This term refers to the temporal detection of epileptic seizures in long-term EEG recordings by evaluating the epileptic character of the combined channel set. For our purposes we evaluate each segment individually, then we infer the seizure onset zone. We refer to van Putten et al. (2005) and Khamis et al. (2009) for an overview of common detection approaches. Other methods include non-linear approaches like entropies in Acharya et al. (2011) or the recurrence quantification analysis (RQA) in Thomasson et al. (2001).

7.2 Methodology

In our analysis we consider multivariate signals $x(t), t = 1, \ldots, T$ (t denotes the time index) consisting of K channels $x_k(t), k = 1, \ldots, K$. We refrain from making rigorous assumptions on the data like in the previous chapters, we only assume the signals to be zero-mean and stationary in a sufficiently short data window.

The multivariate signals we are going to analyze with our methodology are the data presented in section 2.7. For a recapitulation the data consist of three seizures of a 43 year old male patient obtained during a long term ECoG monitoring. The multivariate signal consists of $K = 24$ channels and has a sampling frequency of 128Hz.

As aforementioned our methodology consists of two consecutive steps: the segmentation of the ECoG data, see subsection 7.2.1, and the classification of the segments' epileptic character followed by the seizure onset analysis, see subsection 7.2.2.

As the segmentation and classification are applied channel-wise, we will explain them for an arbitrary single channel $x_k(t)$ in the following.

7.2.1 Segmentation

A simple idea for the segmentation of a channel $x_k(t)$ is to derive a statistics based on the properties of the channel and segment according to it. We follow this idea and construct an univariate statistics based on the spectral properties of the (univariate) channel $x_k(t)$ and from its temporal evolution we derive the segments.

Now we detail the construction of the statistics, in the following sections we present the results of the application and discuss its advantages and shortcomings.

The ECoG data are only stationary over short time periods and in order to analyze

the temporal evolution of a statistics we have to define how to deal with the temporal changes. We simply use a sliding window approach, where a window of length T_{win} seconds is moved over the data in T_{res} second-steps. Within these windows we assume the data to be stationary.

The considered statistics is based on the spectral properties of the channel. Therefore, we calculate the spectral density of the channel x_k for each window, $(f(\lambda))(\tau)$ say, where τ is the new time index associated with the new temporal resolution $1/T_{res}$. The spectral density of a window is associated with the window center on the time axis.

Due to neurophysiological considerations we consider the following frequency bands:

- the δ_{low}-band from 1.0 - 1.5 Hz,

- the δ_{up}-band from 2.0 - 3-5 Hz,

- the ϑ-band from 4.0 - 8.5 Hz,

- the α-band from 9.0 - 13.5 Hz and

- the β-band from 14.0 - 30.0 Hz.

These are the most important frequency bands for EEG signals and therefore they are also important for ECoG analysis. The split of the δ-band was due to technical reasons.

For each window, i.e. each time-step τ in the new temporal resolution, we calculate the power of these bands, e.g. the power of the α-band is calculated by

$$\mathbb{F}_\alpha(\tau) = \int_{9.0}^{13.5} (f(\lambda))(\tau)\, d\lambda, \tag{7.1}$$

as well as the total power

$$\mathbb{F}(\tau) = \int_{1.0}^{64.0} (f(\lambda))(\tau)\, d\lambda.$$

In order to analyze the temporal evolution of a statistics we can either look at the absolute statistics itself or its changes relative to a reference point. Both methods have their advantages, but for our considerations we compare the statistics at a running point to the statistics at a reference point.

Our segmentation method is based on the temporal changes of the relative frequency contributions of the aforementioned physiological bands. We choose an initial reference point τ^*, and for increasing time-steps $\tau > \tau^*$ we calculate our segmentation statistics termed *band power measure (BPM)* as

$$\mathbb{B}(\tau) = \left(\frac{\mathbb{F}_{\delta low}(\tau)}{\mathbb{F}(\tau)} - \frac{\mathbb{F}_{\delta low}(\tau^*)}{\mathbb{F}(\tau^*)} \right)^2 + \ldots + \left(\frac{\mathbb{F}_{\beta}(\tau)}{\mathbb{F}(\tau)} - \frac{\mathbb{F}_{\beta}(\tau^*)}{\mathbb{F}(\tau^*)} \right)^2 . \tag{7.0}$$

If the band power measure exceeds a given threshold *th*, i.e. $\mathbb{B}(\tau) > th$, we start a new segment by updating the reference point $\tau^* = \tau + 1$ and continue the calculation of $\mathbb{B}(\tau)$ for increasing τ. The set of reference points obtained by this algorithm are the boundary points of our segments, i.e. each resulting segment is limited by two subsequent reference points.

We will discuss the properties of the BPM in the discussion section. Right now we emphasize only one important point. As we already discussed in section 2.6, ictal ECoG data of temporal lobe epilepsy patients are often characterized by distinct rhythmic ϑ-activity according to Foldvary et al. (2001). As our algorithm tends to yield segments with one predominant frequency band by construction, we consider it to be appropriate for the segmentation of the ictal ECoG signals of TLE patients.

7.2.2 Onset zone analysis

In the second step we decide whether the data in a segment show epileptic activity or not. For this purpose we propose a simple and intuitive rule which focuses on neurophysiological aspects of TLE patients: A segment is classified as epileptic if ϑ-activity is dominant within this segment. We say that the ϑ-activity is dominant within a segment, i.e. we classify a segment as epileptic, if one of the following two requirements is fulfilled:

- the ϑ-contribution is the largest among the frequency contributions in more than half of the time-steps of the segment or

- the maximal spectral peak occurs in the ϑ-band more often than in each other single spectral band.

This procedure is based on the good localizing value of ϑ-activity, see Foldvary et al. (2001). Following this line of thought the start of the epileptic activity on a single channel is indicated by the beginning of the first epileptic segment.

In order to draw conclusions on the initial seizure propagation we apply the segmentation and subsequent classification channel-wise to all channels. The temporal delay of the start of the epileptic ϑ-activity over the different channels is used for describing seizure propagation. The first channels showing epileptic ϑ-activity indicate the seizure onset zone.

7.3 Results

As we are interested in the initial spread of the rhythmic ϑ-activity, we investigate the first 20 seconds of the three seizures presented in section 2.7. We start the in-

vestigation one second prior to paroxysmal fast activity (30Hz) or high-frequency oscillations (75Hz), which are in a certain sense the first signs of the seizures. The synchronization phase and the rhythmic ϑ-activity start approximately 10 seconds after these first signs.

We apply our methodology to the full 24-channel set $x_k(t), k = 1, \ldots, 24$, of all three seizures. Instead of presenting the segmentation and classification results for all seizures, we want to focus on the results of seizure 1 and discuss them thoroughly. This discussion will deepen the understanding of the methodology.

7.3.1   Segmentation

For the segmentation we use the following set of parameters: $T_{win} = 1.5$s, $T_{res} = 1/16$s. Power spectral densities are estimated using the non-parametric Welch method. Furthermore, we employ an empirically determined threshold $th = 0.07$ and an initial reference point τ^* at 0.75s.

In figure 7.1 we display the segmentation of an exemplary channel in detail. For presentation purposes we choose the channel A11 from seizure 1, this channel will also occur in the other figures.

On top, figure 7.1(a), the ECoG data are shown. For a better understanding of the band power measure we display the temporal evolution of the relative frequency contributions in the middle, figure 7.1(b). The corresponding BPM statistics is shown at the bottom, figure 7.1(c). A new segment starts when the statistics exceeds the threshold. The resulting segment boundaries are marked in all three graphs by vertical lines.

In this example a significant change of the BPM statistics can be observed in case of frequency shifts from one physiological band into another, e.g. at 16:12:50 when a rise of the α-contribution disrupts the ϑ-contribution.

In figure 7.2 we show the segmentation of all channels of seizure 1 for the reader to get a grasp how the segmentation algorithm behaves. The different phases discussed in section 2.7 can clearly be found and we note that the length of the segments is largest in the phase of distinct ϑ-activity. Furthermore, we see that (spatially) adjacent channels show a similar behavior.

As it will turn out in the next section, the most interesting channels with respect to the seizure propagation are A9, A10, A11, A12, B8, B7 and B6. Therefore, we show an enlarged graph of these 7 channels in figure 7.3. The ϑ-activity seems to start at around 16:12:45 at all of these channels.
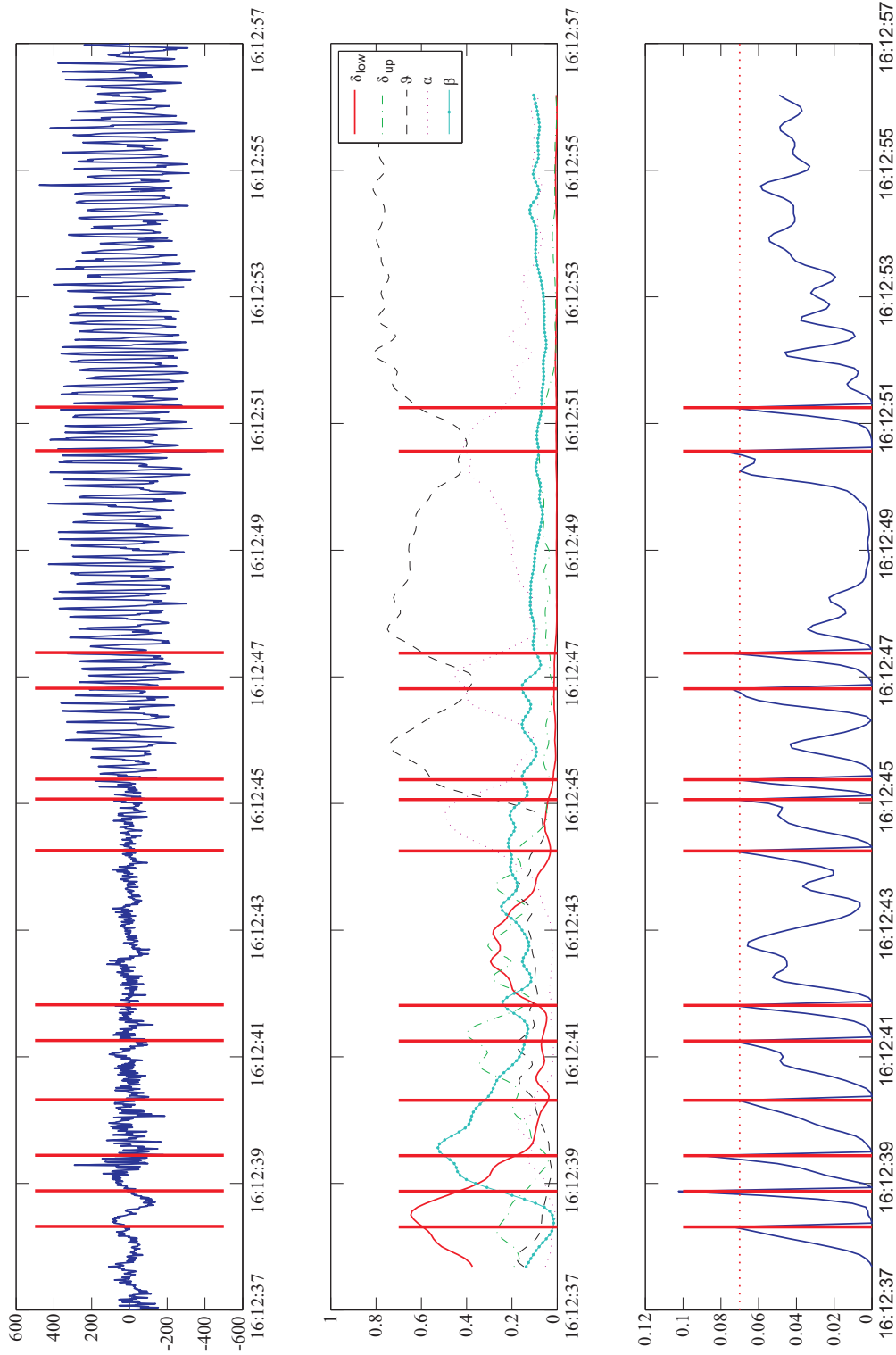
Figure 7.1: Segmentation of channel A11, seizure 1. (a) ECoG data, (b) Relative contributions of the physiological frequency bands, (c) Band Power Measure and imposed threshold (dotted). Segment boundaries are indicated by vertical lines in all three graphs.
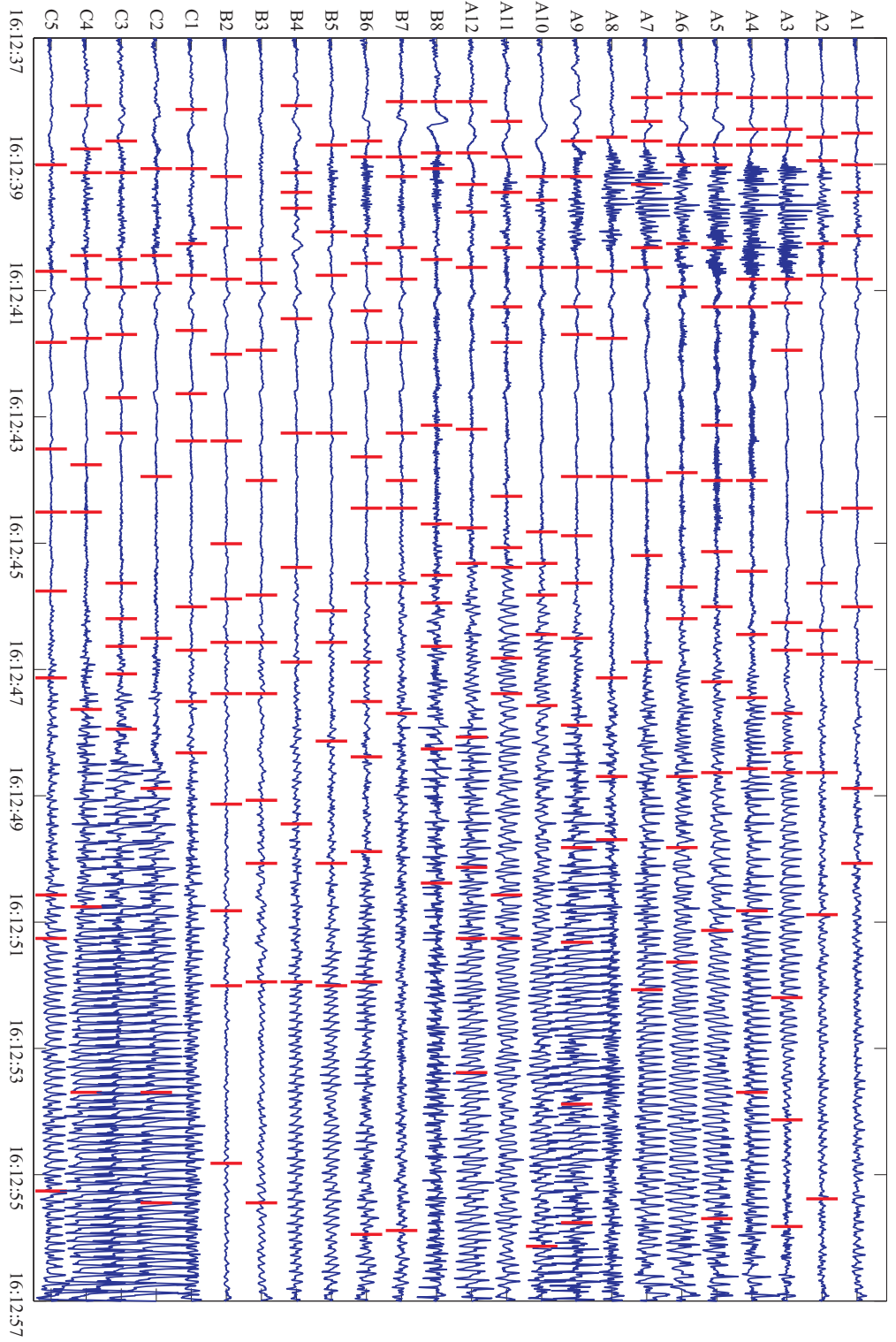
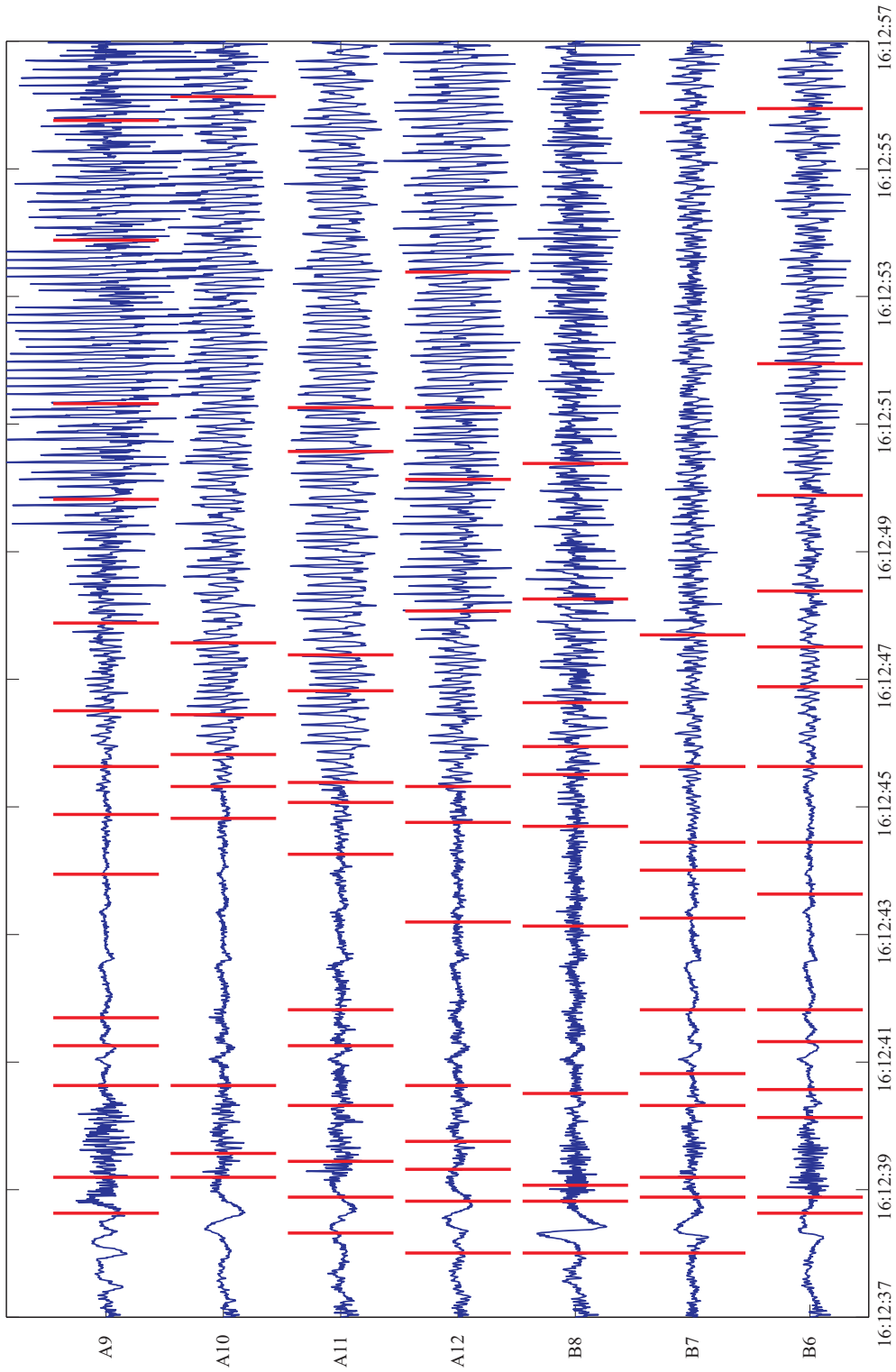Figure 7.2: Segmentation of all channels, seizure 1. Segment boundaries are indicated by vertical lines.

Figure 7.3: Segmentation of 7 selected channels, seizure 1. Segment boundaries are indicated by vertical lines.
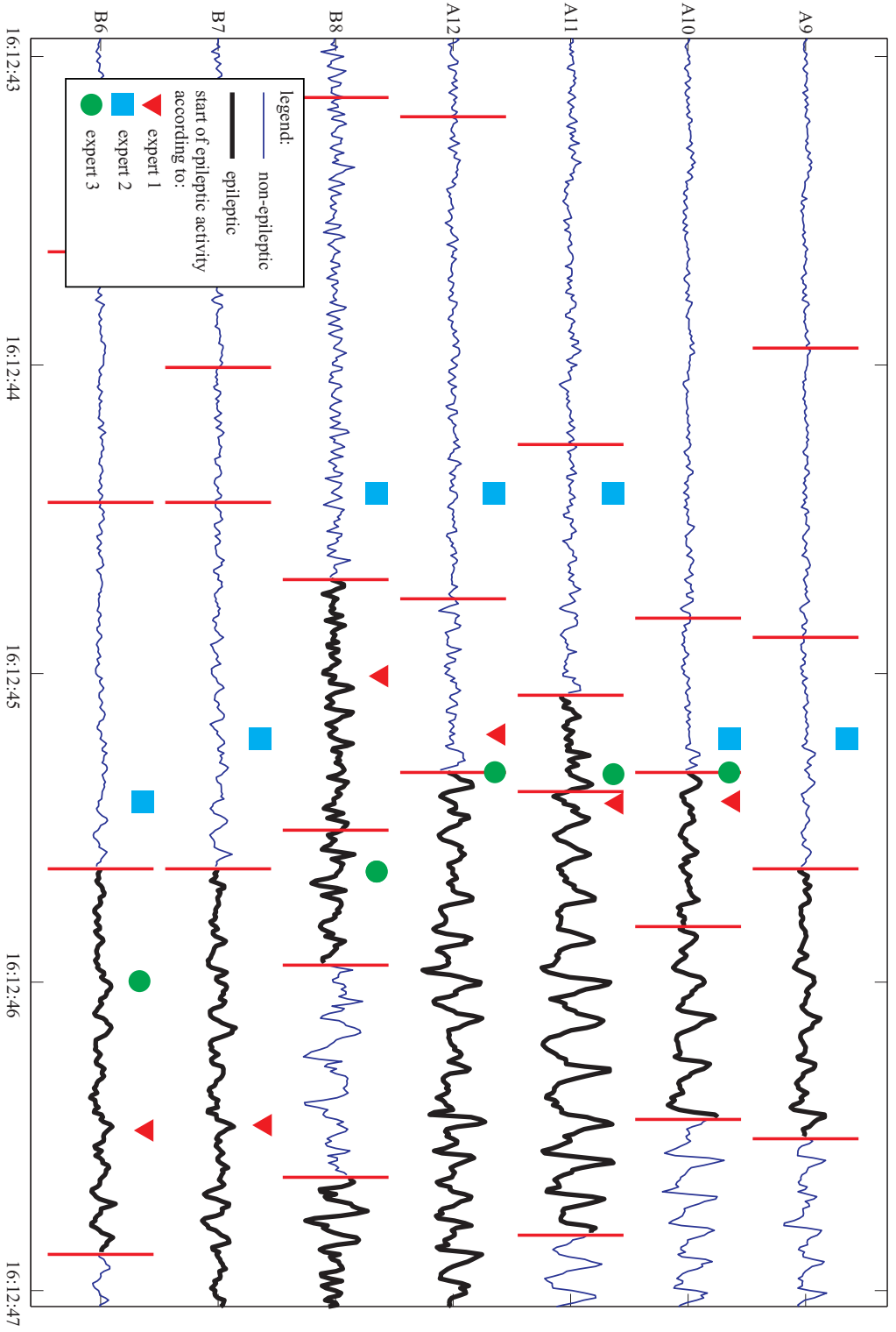
Figure 7.4: Propagation of ϑ-activity during four seconds in 7 selected channels, seizure 1. Segment boundaries are indicated by vertical lines. Epileptic segments according to our classification are highlighted. The starting time of the epileptic activity per channel according to the visual analysis of the experts are marked.

Seizure	Investigator	Initial Electrodes	Close follow-up
1	*Algorithm*	*B8*	*A10, A11, A12*
	Expert 1	B8	A10, A11, A12
	Expert 2	A11, A12, B8	A9, A10, B7
	Expert 3	A10, A11, A12	B8
2	*Algorithm*	*A10, A11, A12*	*B6, B8*
	Expert 1	A11, A12	A9, A10
	Expert 2	A11, A12	A10
	Expert 3	A11, A12	B8
3	*Algorithm*	*A10, A11, A12*	*B6, C1, C2, C5*
	Expert 1	A9, A10	A8, A11, A12, B6, B7, B8, C1, C4, C5
	Expert 2	A9	A1, A2, A3, C2, C3
	Expert 3	A8, A9	A1, C3, C4, C5

Table 7.1: Onset zone of all three seizures. Results based on our method and the visual inspection by clinicians

7.3.2 Onset zone analysis

Subsequent to the segmentation we classify the segments as epileptic or not according to the rules discussed in subsection 7.2.2. This allows us to investigate the initial propagation of the ϑ-activity and derive the localization of the seizure onset zone. The channels showing the first ϑ-waves are considered the seizure onset zone.

Following up the last subsection we present our findings for seizure 1 in figure 7.4. The figure shows a 4-second-zoom of the channels displayed in figure 7.3. The epileptic segments are drawn with a bold line, non-epileptic segments with a normal line. The depicted channels show the earliest epileptic activity among all considered channels of seizure 1.
In figure 7.4 channel B8 shows the first occurrence of dominant ϑ-activity, closely followed by electrodes A10, A11 and A12. Therefore, we consider these channels the seizure onset zone in seizure 1.
Furthermore, we mark the beginning of the epileptic activity according to the visual analysis of three experts, see the legend in figure 7.4. The comparison of these times with the computed start of the epileptic activity serves as an assessment for our methodology.

We applied the segmentation and classification to all three seizures and table 7.1
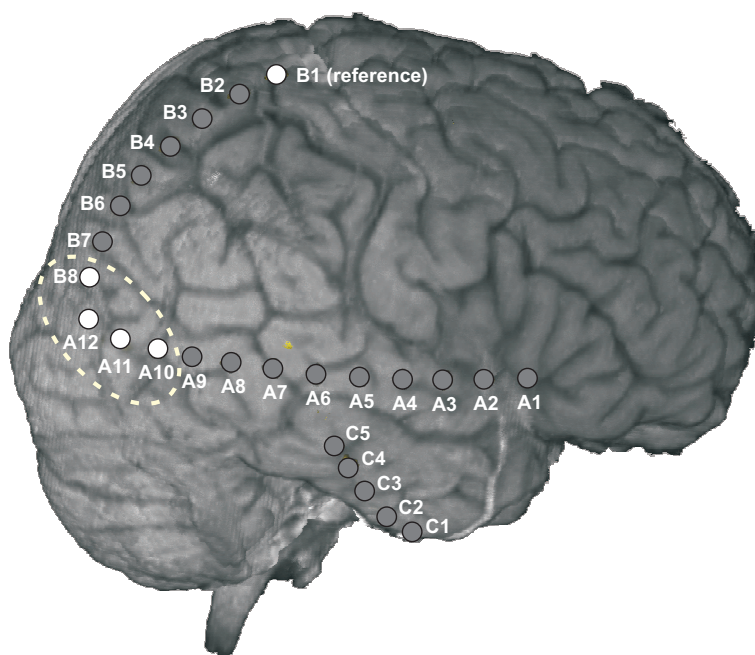
Figure 7.5: MRI scan with electrode positions. The electrodes of the seizure onset zone indicated by the proposed methodology are marked.

summarizes our findings as well as the visual analysis. Our algorithm indicates that the seizure onset zone comprises the electrodes B8, A10, A11 and A12. This coincides well with the visual analysis of the experts. In figure 7.5 we marked the electrodes of the seizure onset indicated by our proposed methodology.

7.4   Discussion

In this section we discuss the properties of the segmentation and the onset zone analysis of our methodology.

7.4.1   Segmentation

The design goal of our segmentation method was to achieve long segments during ictal periods and short segments in non-ictal periods. The BPM-based segmentation method has shown to achieve this requirement.

In figure 7.1 we see the characteristics of the BPM statistics: Prior to the rhythmic ϑ-activity (starting at 16:12:45) we observe quickly interchanging frequency contributions, see plot (b). This results in a BPM statistics with high variations and frequent threshold exceedings, see plot (c). Therefore, our algorithm yields short segments in this period. On the other hand, during the distinct rhythmic activity, only small power shifts occur within the physiological frequency bands. Thus, the

frequency contribution of the respective bands show a constant behavior, namely the ϑ-band on a high level. This results in longer segments.

As can be seen in the segmentation overview in figure 7.2, the segments are long in ictal periods, i.e. during the phase with distinct ϑ-activity, and short otherwise. This is exactly the desired behavior.

The segmentation behavior is rooted on the construction of the BPM in formula 7.2.1. Our first segmentation approach involved a statistics of the form $\int ((f_{xx}(\lambda))(\tau) -(f_{xx}(\lambda))(\tau^*)) d\lambda$, but as it turned out, this statistics was too sensitive to shifts within frequency bands, therefore we introduced the frequency contributions $\mathbb{F}_{\cdot}$.

Furthermore, we initially used absolute differences of the relative frequency contributions, e.g. $|\mathbb{F}_\alpha(\tau)/\mathbb{F}(\alpha) - \mathbb{F}_\alpha(\tau^*)/\mathbb{F}(\tau^*)|$, in contrast to the squared differences in the final statistics. As the squared differences showed a better segmentation behavior, they are now used in the BPM statistics.

The results of the segmentation depend on the parameters used in the segmentation algorithm:

- the window length T_{win},

- the time-steps T_{res},

- the parameters for the calculation of the spectral density within the sliding windows,

- the initial reference point τ^* and

- the threshold *th*.

The parameters can be arranged in two groups. The window length T_{win}, time-steps T_{res} and the spectral density parameters belong to the first group. These parameters have to be adjusted in order for the spectral estimation and its temporal change to yield reasonable results. They are in a certain sense independent of the parameters in the second group, which signify the main segmentation parameters: the threshold *th* and the starting point τ^*. We call the first group the spectral estimation parameters and the second group the segmentation parameters.

We want to briefly discuss the segmentation parameters. Due to the low threshold $th = 0.07$ we obtain a reactive segmentation behavior and short segments. As a consequence of this fact the influence of the initial reference point τ^* is almost negligible in this setting.

Furthermore, the presented segmentation is robust in respect to the segmentation parameters. Small parameter changes of *th* and τ^* only lead to slightly different segments. This is an advantage of the method.

7.4.2 Onset zone analysis

In this study we assume that the initial ϑ-spread represents a valid indicator (among others) for seizure propagation in TLE, in particular for the determination of the seizure onset zone. The proposed method delivers a conclusive seizure onset zone for the three seizures. Furthermore, the resulting onset zone is in good accordance with the clinical findings, see Table 7.1, which supports our assumption that ϑ-activity is a good indicator for epileptic TLE activity.

As we mentioned in the previous subsection the proposed segmentation method is robust with respect to the segmentation parameters. The seizure onset zone identification inherits the robustness, because the segmentation is the first step in the methodology. This means that small changes of th and τ^* do not change the resulting seizure onset zone. This is a very good property of the proposed methodology and we want to emphasize it.

Although the proposed methodology is capable of identifying the first channel showing epileptic activity, this channel should not be called the focus of the seizure. The identification of this first channel is not robust, and even the expert opinions differ on the first epileptic channel, see table 7.1. Therefore, it is far more reasonable to speak of a channel set showing the first epileptic signs, we call this channel set the seizure onset zone.

In our analysis we intentionally limit ourselves to tracking ϑ-activity as an indicator of epileptic activity. However, the observed patient also shows intermittent epileptic α-activity. The segments showing this α-activity are not characterized as epileptic, see figure 7.4 (channels A9, A10, A11 at 16:12:47). For future works we would propose patient-specific epileptic frequency bands indicated by the clinicians to enhance the onset zone analysis.
Another possible amelioration of the proposed methodology would be the improvement of the classification rules. The additional consideration of the signal amplitude or the entropy as measures of rhythmicity, see e.g. van Putten et al. (2005), could enhance the segment classification and therefore the whole seizure onset zone analysis.

To conclude the discussion we want to say it is remarkable that the combination of two simple ideas delivers results which are well correlated with the clinical findings. This might be due to the close relation between the method and neurophysiology.

7.5 Conclusion

In this chapter we proposed a novel method for the detection of the seizure onset zone of epileptic seizures based on the segmentation and subsequent classification

of ictal ECoG data. The application shows promising first results in tracking the initial propagation of ictal ϑ-activity as an indicator for seizure propagation. The identified seizure onset zone correlates well with the visual inspection of the clinicians.

It therefore has the potential for an objectivation in the presurgical clinical evaluation of therapy-resistant patients. However, this requires further research and an application to a broader data basis.

Chapter 8

Conclusion

The main contribution of this thesis are two methods for the detection of the epileptic seizure onset zone. The method presented in chapter 6 has a detailed theoretic background and focuses on the causal analysis of the most important parts of the observations. In chapter 7 we propose a method closely connected to the visual analysis performed by the clinicians.

In this chapter we conclude this work. We present a general discussion as well as ideas how to enhance the proposed methods. We focus on aspects of the methods going beyond the scope of their description. Furthermore, we propose possible next steps.

8.1 General discussion

Although both of our methods yield good results we have to keep some limiting factors for ECoG analysis in mind.

ECoG data are normally obtained by surgically implanted subdural strip electrodes. There is no default scheme for the placement of these electrodes, like the 10-20 system for the EEG electrodes. Thus the placement of the electrodes as well as the corresponding electrical referencing scheme is unique for each patient. Therefore, a generalization or automatization of our methods for a larger data basis will be difficult.

The proposed methods are intended for the use with ECoG data obtained by subdural strip electrodes. Of course it is possible to apply the methods to regular (i.e. surface) EEG data. As the processed ECoG data of the analyzed patient did not show (technical or natural) artifacts, we do not know if our methods are robust with respect to data artifacts. This means it is unclear if the proposed methods also yield good results in the case of disturbing artifacts. In fact the band power measure was originally developed for the use with EEG and we expect it to be robust for EEG data.

Summing up, both methods yield results which are in good accordance with the

clinical findings. But we have to consider the methods as what they are. The proposed methods are useful tools to aid the clinicians in the visual analysis. The best analyzer for the data is still a human.

8.2   Outlook

Both presented methods show promising first results. The logical next step is to apply the methods to a larger data basis.
In this section we discuss ways to possibly enhance the proposed methods and to develop them further.

8.2.1   Band power measure

The methodology presented in chapter 7 is a very simple and yields surprisingly good results.
One of the main problems of the method is its dependency on the ϑ-band. Although rhythmic ϑ-activity normally is an indicator for epileptic activity in TLE, the ictal activity could also comprise faster waves, e.g. α-activity (9-13 Hz), or slower activity, e.g. δ-activity (1-3 Hz). In fact we had a lot of problems with the processed ECoG data because some channels showed epileptic activity at 9 Hz (which is faster than normal ϑ-activity). We propose two enhancements of the method to cope with this problem.
First, we propose to use patient specific frequency bands, which can differ from the physiological frequency bands presented in section 7.2. These new bands are used for the segmentation as well as the classification.
Second, in addition to the two classification rules for the epileptic character of a segment we propose a third rule based on rhythmicity. If a segment shows distinct rhythmic activity it is also classified as epileptic.

8.2.2   Influence analysis

In the proposed method only one static 4 seconds data window is considered. Based on this investigation we draw conclusions regarding the seizure onset zone. The naturally arising question is, whether it is also possible to draw conclusions regarding the seizure propagation using the proposed causal analysis. The simplest idea would be to use a sliding data window and conduct the proposed methodology for each data window. The temporal change of the influences could give indications over the seizure propagation. Unfortunately there is no clear interpretation for this procedure like in the case of the detection of the seizure onset zone (comprising the channels with the highest out-degrees). Perhaps the temporal change of the zones with the highest in-degrees and the zones with the highest out-degrees might give indications which zones emit information and which zones receive information.

As we already mentioned in section 6.6 it would be possible to use different causality measures to quantify the strength of causality for the influence analysis. In this point we see the largest potential for an enhancement of the method.

Another promising way to develop the influence analysis further is the application of graphical models. In chapter 4 we discussed the application of graphical models to regular AR systems. In the authors opinion it should be possible to incorporate graphical models to infer the influences between the latent variables. It has to be checked whether it is reasonable to apply graphical models to sub-systems (analog to the original influence analysis) or if graphical models should be applied to the whole system. For the latter approach a generalization of graphical models to the singular AR case would be necessary. From a theoretical point of view this is the most challenging enhancement.

8.2.3 General outlook

The logical next step is the application of our methods to a broader data basis. Up to this point the influence analysis was applied to one seizure and the band power measure was applied to three seizures.

In the previous subsections we discussed ways to enhance the proposed methods, but their original purpose is to aid clinicians in the visual analysis of the ECoG (or EEG) data. They shall serve as a second opinion in the difficult visual analysis, because the proposed methods provide an objective view on the data.
In the authors opinion the implementation of these methods in a clinical environment is more important than the further development of the methods. The most sophisticated methods are useless if they are not used by the clinicians. Therefore, the next step after the application of the methods to a broader data basis should be the implementation in clinical software.

8.3 Concluding remarks

Both presented methods show promising first results and in the authors opinion they should be pursued further. The authors see the most benefit for the patients in the implementation of the proposed methods in clinical software.
Seizure propagation analysis is an interesting and challenging topic. It will stay a vital research topic in the future, because the understanding of epileptic seizures and their mathematical description is far from complete.

The main aim of this thesis was to develop methods which aid clinicians in the visual analysis in order raise the chance of a seizure-free surgical outcome for patients. In other words we wanted to help curing epilepsy. We hope to have contributed a small part to reach this goal by providing this thesis.

Furthermore, we hope to have given the reader new insights and an enjoyable time reading this work.

Bibliography

Acharya, U., Molinari, F., Sree, S., Chattopadhyay, S., Ng, K.-H., Suri, J., 2011. Automated diagnosis of epileptic EEG using entropies. Biomedical Signal Processing and Control.

An, H., Gu, L., 1989. Fast stepwise procedures of selection of variables by using AIC and BIC criteria. Acta mathematicae applicatae Sinica 5 (1), 60–67.

Anderson, B., Deistler, M., 2008. Generalized linear dynamic factor models. a structure theory. Proceedings of IEEE CDC 2008, Cancun, 1980–1985.

Atwood, H., MacKay, L., 1989. Essentials of Neurophysiology. Mosby-Year Book.

Baccala, L., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. Biological Cybernetics 84, 463–474.

Baccala, L., Takahashi, D., Sameshima, K., 2007. Generalized partial directed coherence. In: Proceedings of the 15th International Conference on Digital Signal Processing.

Barrett, A., Barnett, L., Seth, A., 2010. Multivariate Granger causality and generalized variance. Physical Review E 81, 1–14.

Baumgartner, C., 2001. Handbuch der Epilepsien. Springer.

Behrens, E., Zentner, J., van Roost, D., Hufnagel, A., Elger, C. E., Schramm, J., 1994. Subdural and depth electrodes in the presurgical evaluation of epilepsy. Acta Neurochirurgica 128, 84–87.

Berger, H., 1929. Über das Elektroenkephalogramm des Menschen. European Archives of Psychiatry and Clinical Neuroscience 87, 527–550.

Blinowska, K., 2011. Review of the methods of determination of directed connectivity from multichannel data. Medical and Biological Engineering and Computing 49, 521–529.

Bodenstein, G., Praetorius, M., 1977. Feature extraction from the electroencephalogram by adaptive segmentation. Proceedings of the IEEE 65 (5), 642–652.

Bressler, S., Seth, K., 2011. Wiener-Granger causality: A well established methodology. NeuroImage 58(2), 323–329.

Brillinger, D., 1981. Time series, data and analysis. Holden Day.

Brillinger, D., 1996. Remarks concerning graphical models for time series and point processes. Revista de Econometria 16, 1–23.

Brockwell, P., Davis, R., 1991. Time series: Theory and methods. Springer.

Brodsky, B., Darkhovsky, B., Kaplan, A., Shishkin, S., 1999. A nonparametric method for the segmentation of the EEG. Computer Methods and Programs in Biomedicine 60, 93–106.

Cattell, R. B., 1966. The scree test for the number of factors. Multivariate Behavioral Research 1 (2), 245–276.

Chamberlain, G., 1983. Funds, factors and diversification in arbitrage pricing models. Econometrica 51(5), 1305–1324.

Chamberlain, G., Rothschild, M., 1983. Arbitrage, factor structure and meanvariance analysis on large asset markets. Econometrica 51(5), 1281–1304.

Chavez, M., Martinerie, J., Le Van Quyen, M., 2003. Statistical assessment of nonlinear causality: application to epileptic EEG signals. Journal of Neuroscience Methods 124, 113–128.

Chiang, J., Wang, Z., McKeown, M., 2009. Sparse multivariate autoregressive (MAR)-based partial directed coherence (PDC) for electroencephalogram (EEG) analysis. IEEE ICASSP Proceedings 2009, 457–460.

Cowell, R., Dawid, P., Lauritzen, S., Spiegelhalter, D., 2007. Probabilistic networks and expert systems. Springer.

Dahlhaus, R., 2000. Graphical interaction models for multivariate time series. Metrika 51, 157–172.

Dahlhaus, R., Eichler, M., 2003. Causality and graphical models in time series analysis. In: P. Green, N. and Hjort and S. Richardson (eds.), Highly structured stochastic systems , Oxford University Press. pp. 115–137.

Dahlhaus, R., Eichler, M., Sandkuehler, J., 1997. Identification of synaptic connections in neural ensembles by graphical models. Journal of Neuroscience Methods 77, 93–107.

d'Aspremont, A., Ghaoui, L. E., Jordan, M., Lanckriet, G., 2007. A direct formulation for sparse PCA using semidefinite programming. SIAM Review 49 (3), 434–448.

Deistler, M., Anderson, B. D. O., Filler, A., Zinner, C., Chen, W., 2010. Generalized dynamic factor models - An approach via singular autoregressions. European Journal of Control 16 (3), 211–224.

Deistler, M., Zinner, C., 2007. Modelling high-dimensional time series by generalized linear dynamic factor models: An introductory survey. Communications in Information and Systems 7, 153–166.

Dietsch, G., 1932. Fourier-Analyse von Elektroenkephalogrammen des Menschen. Pflügers Archiv: European Journal of Physiology 230, 106–112.

Edwards, J., 2000. Introduction to graphical modelling. 2nd edn. Springer.

Eichler, M., 2005. A graphical approach for evaluating effective connectivity in neural systems. Philosophical Transactions of the Royal Society B 360, 953–967.

Eichler, M., 2006a. Graphical modeling of dynamic relationships in multivariate time series. In: B. Schelter, M. Winterhalder and J. Timmer (eds.), Handbook of Time Series Analysis, Wiley-VCH. pp. 335–372.

Eichler, M., 2006b. Graphical modeling of multivariate time series with latent variables. preprint University of Maastricht.

Eichler, M., 2006c. On the evaluation of information flow in multivariate systems by the directed transfer function. Biological Cybernetics 94, 469–482.

Eichler, M., 2007. Granger causality and path diagrams for multivariate time series. Journal of Econometrics 137, 334–353.

Eichler, M., 2009. Causal inference from multivariate time series: What can be learned from Granger causality. In: C. Glymour, W. Wang, D. Westerstahl (eds.), Proceedings from the 13th International Congress of Logic, Methodology and Philosophy of Science, King's College Publications, London.

Engel, J., 1996. Surgery for seizures. New England Journal of Medicine 334, 647–652.

Engel, J., da Silva, F. L., 2012. High-frequnecy oscillations - Where we are and where we need to go. In press: Progress in Neurobiology, 3.

Faes, L., Nollo, G., 2010. Extended causal modeling to assess partial directed coherence in multiple time series with significant instantaneous interactions. Biological Cybernetics 103, 387–400.

Filzmoser, P., 2010. Multivariate Statistik, Vorlesungsskriptum.

Flamm, C., Graef, A., Pirker, S., Baumgartner, C., Deistler, M., 2012a. Influence analysis for high-dimensional time series with an application to epileptic seizure onset zone detection. Submitted to the Journal of Neuroscience Methods.

Flamm, C., Kalliauer, U., Deistler, M., Waser, M., Graef, A., 2012b. Graphs for dependence and causality in multivariate time series. In: Wang, L., Garnier, H. (Eds.), System Identification, Environmental Modelling, and Control System Design. Springer, pp. 133–157.

Florens, J., Mouchart, M., 1985. A linear theory for non-causality. Econometrica 53, 157–175.

Florin, E., 2010. Causality measures between neural signals from invasively and non-invasively obtained local field potentials in humans. Vol. 29 of Health. Schriften des Forschungszentrum Jülich.

Florin, E., Gross, J., Pfeifer, J., Fink, G., Timmermann, L., 2011. Reliability of multivariate causality measures for neural data. Journal of Neuroscience Methods 198, 344–358.

Foldvary, N., Klem, G., Hammel, J., Bingaman, W., Najm, I., Lüders, H., 2001. The localizing value of ictal EEG in focal epilepsy. Neurology 57, 2022–2028.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2000. The generalized dynamic factor model: identification and estimation. The Review of Economics and Statistics 65, 453–473.

Forni, M., Lippi, M., 2001. The generalized dynamic factor model: representation theory. Econometric Theory 17, 1113–1141.

Freiwald, W., Valdes, P., Bosch, J., Biscay, R., Jimenez, J., Rodriguez, L., Rodriguez, V., Kreiter, A., Singer, W., 1999. Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex. Journal of Neuroscience Methods 94, 105–119.

Friston, K., Harrison, L., Penny, W., 2003. Dynamic causal modeling. NeuroImage 19, 1273–1302.

Gabor, D., 1946. Theory of Communication. Journal of IEEE 93 (26), 429–457.

Gath, I., Feuerstein, C., Pham, D., Rondouin, G., 1992. On the tracking of rapid dynamic changes in seizure EEG. IEEE Tansactions on Biomedical Engineering 39 (9), 952–958.

Geweke, J., June 1982. Measurement of linear dependence and feedback between multiple time series. Journal of the American Statistical Association 77 (378), 304–313.

Geweke, J., 1984. Measurement of conditional linear dependence and feedback between multiple time series. Journal of the American Statistical Association 79, 907–915.

Graef, A., 2008. Non-stationary autoregressive modeling for epileptic seizure propagation analysis. Master's thesis, Vienna University of Technology.

Graef, A., Flamm, C., Pirker, S., Deistler, M., Baumgartner, C., 2012a. A physiologically motivated ECoG segmentation method for epilpetic seizure onset zone detection. IEEE EMBC 2012.

Graef, A., Hartmann, M., Flamm, C., Baumgartner, C., Deistler, M., Kluge, T., 2012b. A novel method for the identification of synchronization effects in multichannel ECoG with an application to epilepsy. Submitted to Biological Cybernetics.

Graef, A., Kluge, T., Baumgartner, C., Deistler, M., Hartmann, M., 2008. A novel method for the characterization of epileptic seizure propagation in multichannel EEG and ECoG. Epilepsia 49 (S7), 13–14.

Granger, C., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.

Granger, C., 1980. Testing for causality, a personal viewpoint. Journal of Economic Dynamics and Control 2, 329–352.

Granger, C., 1988. Some recent developments in a concept of causality. Journal of Econometrics 39, 199–211.

Götz-Trabert, K., Hauck, C., Wagner, K., Fauser, S., Schulze-Bonhage, A., 2008. Spread of ictal activitiy in focal epilepsy. Epilepsia 49 (9), 1594–1601.

Guo, S., Seth, A., Kendrick, K., Zhou, C., Feng, J., 2008. Partial Granger causality - Eliminating exogenous inputs and latent variables. Journal of Neuroscience Methods 172, 79–93.

Hannan, E., 1970. Multiple time series. Wiley.

Hannan, E., Deistler, M., 2012. The statistical theory of linear systems. Classics in Applied Mathematics. SIAM.

Hartmann, M., Graef, A., Perko, H., Baumgartner, C., Kluge, T., 2008. A novel method for the characterization of synchronization and coupling in multichannel EEG and ECoG. WASET Proceedings 2008 34, 6–11.

Hirtz, D., Thurman, D., Gwinn-Hardy, K., Mohamed, M., Chaudhuri, A., Zalutsky, R., 2007. How common are the "common" neurologic disorders? Neurology 68, 326–337.

Hosoya, Y., 1977. On the Granger condition for for non-causality. Econometrica 45, 1735–1736.

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology 26 (6), 417–441.

Hsiao, C., 1982. Autoregressive modeling and causal ordering of econometric variables. Journal of Economica, Dynamics and Control 4, 243–259.

Inouye, T., Toi, S., Matsumoto, Y., 1995. A new segmentation method of electroencephalograms by use of akaike's information criterion. Cognitive Brain Research 3, 33–40.

Jackson, J., 2004. A user's guide to principal components. John Wiley & Sons.

Jenssen, S., Roberts, C., Gracely, E., Dlugos, D., Sperling, M., 2011. Focal seizure propagation in the intracranial EEG. Epilepsy Research 93, 25–32.

Kaminski, M., Blinowska, K., 1991. A new method of the description of the information flow in the brain structures. Biological Cybernetics 65, 203–210.

Kaminski, M., Ding, M., Truccolo, W., Bressler, S., 2001. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biological Cybernetics 85, 145–157.

Khamis, H., Mohamed, A., Simpson, S., 2009. Seizure state detection of temporal lobe seizures by autoregressive spectral analysis of scalp EEG. Clinical Neurophysilogy 120, 1479–1488.

Kim, J. S., Hwan, C., Jung, Y. J., Kim, E. Y., Lee, S. K., Chung, C. K., 2010. Localization and propagation analysis of ictal source rhythm by electrocorticography. NeuroImage 52, 1279–1288.

Kong, X., Lou, X., Thakor, N., 1997. Detection of EEG changes via a generalized Itakura distance. Proceedings of the IEEE EMBC, 1540–1542.

Korzeniewska, A., Manczak, M., Kaminski, M., Blinowska, K., Kasicki, S., 2003. Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. Journal of Neuroscience Methods 125, 195–207.

Lauritzen, S., 1996. Graphical models. Oxford University Press.

Lüders, H., 1992. Epilepsy Surgery. Raven Press.

Lütkepohl, H., 2007. New introduction to multiple time series analysis. Springer.

Marinazzo, D., Liao, W., Chen, H., Stramaglia, S., 2011. Non-linear connectivity by Granger causality. NeuroImage 58, 330–338.

Matysiak, A., Durka, P., Montes, E., Barwinski, M., Zwolinski, P., Roszkowski, M., Blinowska, K., 2005. Time-frequency space localization of epileptic EEG oscillations. Acta Neurobiologiae Experimentalis 65, 435–442.

Ombao, H., von Sachs, R., Guo, W., 2005. SLEX analysis of multivariate non-stationary time series. Journal of the American Statistical Association 100 (470), 519–531.

Osterhage, H., Mormann, F., Wagner, T., Lehnertz, K., 2007. Measuring the directionality of coupling: Phase versus state space dynamics and application to EEG time series. International Journal of Neural Systems 17, 139–148.

Paulsen, F., Waschke, J. (Eds.), 2010. Sobotta: Atlas der Anatomie des Menschen - Kopf, Hals und Neuroanatomie, 23rd Edition. Elsevier.

Pearl, J., 2000. Causality. Cambridge University Press.

Pearson, K., 1901. On lines and planes of closest fit to systems of points in space. Philosophical Magazine Series 6 2 (11), 559–572.

Pondal-Sordo, M., Diosy, D., Téllez-Zenteno, J. F., Sahjpaul, R., Wiebe, S., 2007. Usefulness of intracranial EEG in the decision process for epilepsy surgery. Epilepsy Research 74, 176–182.

Richardson, T., 2003. Markov properties for acyclic directed mixed graphs. Scandinavian Journal of Statistics 30, 145–157.

Rozanov, Y., 1967. Stationary random processes. Holden Day.

Sanei, S., Chambers, J., 2009. EEG signal processing. Wiley.

Sargent, T. J., Sims, C. A., 1977. Business cycle modelling without pretending to have too much a proiri economic theory. Seminar paper on new methods in business cycle research.

Schelter, B., Timmer, J., Eichler, M., 2009. Assessing the strength of directed influences among neural signals using renormalized partial directed coherence. Journal of Neuroscience Methods 179, 121–130.

Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lücking, C., Dahlhaus, R., Timmer, J., 2005. Testing for directed influences among neural signals using partial directed coherence. Journal of Neuroscience Methods 152, 210–219.

Scherrer, W., Deistler, M., 1998. A structure theory for linear dynamic errors-in-variables models. SIAM Journal on Control and Optimization 36, 2148–2175.

Schneble, H., 2003. Heillos, heilig, heilbar - Geschichte der Epilepsie. de Gruyter.

Schneider, T., Neumaier, A., 2001. Algorithm 808: Arfit, a matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. ACM Transactions on Mathematical Software 27 (1), 58–65.

Schreiber, T., 2000. Measuring information transfer. Physical Review Letters 85 (2), 461–464.

Schuele, S., Lüders, H., 2008. Intractable epilepsy: Management and therapeutic alternatives. The Lancet Neurology 7 (6), 514–524.

Schuster, T., Kalliauer, U., 2009. Localizing the focus of epileptic seizures using modern measures from multivariate time series analysis. Master's thesis, Vienna University of Technology.

Seth, A. K., 2010. A matlab toolbox for Granger causal connectivity analysis. Journal of Neuroscience Methods 186, 262–273.

Sims, C., 1972. Money, income, and causality. The American Economic Review 62 (4), 540–552.

Sommerlade, L., Thiel, M., Platt, B., Planod, A., Riedel, G., Grebogic, C., Timmer, J., Schelter, B., 2012. Inference of Granger causal time-dependent influences in noisy multivariate time series. Journal of Neuroscience Methods 203, 173–185.

Stock, J., Watson, M., 2002a. Forecasting using principal components from a large number of predictors. Journal of the Americal Statistical Association 97, 1167–79.

Stock, J., Watson, M., 2002b. Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics 20, 147–62.

Thomasson, N., Hoeppner, T., Webber, C., Zbilut, J., 2001. Recurrence quantification in epileptic EEG. Physics Letters A 279, 94–101.

Thurstone, L., 1947. Multiple factor analysis. University of Chicago Press.

Tibshirani, R., 1996. Regression shrinkage and selection via the LASSO. Journal of the Royal Statistical Society B 58 (1), 267–288.

Tseng, S.-Y., Chen, R.-C., Chong, F.-C., Kuo, T.-S., 1995. Evaluation of parametric methods in EEG signal analysis. Medical Engineering and Physics 17, 71–78.

van Putten, M., Kind, T., Visser, F., Lagerburg, V., 2005. Detecting temporal lobe seizures from scalp EEG recordings: A comparison of various features. Clinical Neurophysilogy 116, 2480–2489.

Whittaker, J., 2000. Graphical models in applied multivariate statistics. Wiley.

Wiebe, S., Blume, W., Girvin, J., Eliasziw, M., 2001. A randomized, controlled trial of surgery for temporal-lobe epilepsy. The New England Journal of Medicine 345, 311–318.

Wiener, N., 1956. The theory of prediction. Modern Mathematics for engineeers Series 1, Chapter 8.

Wiener, N., Masani, P., 1957. The prediction theory of multivariate stochastic processes II. The linear predictor. Acta Mathematica 99 (1), 93–137.

Wilke, C., Ding, L., He, B., 2008. Estimation of time-varying connectivity patterns through the use of an adaptive directed transfer function. IEEE Transactions on Biomedical Engineering 55 (11), 2557–2564.

Winterhalder, M., Schelter, B., Hesse, W., Schwab, K., Leistritz, L., Klan, D., Bauer, R., Timmer, J., Witte, H., 2005. Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems. Signal Processing 85, 2137–2160.

Wu, L., Gotman, J., 1998. Segmentation and classification of EEG during epileptic seizures. Electroencephalography and Clinical Neurophysiology 106, 344–356.

Zijlmans, M., Jacobs, J., Kahn, Y., Zelmann, R., Dubeau, F., Gotman, J., 2011. Ictal and interictal high frequency oscillations in patients with focal epilepsy. Clinical Neurophysiology 122, 664–671.

Zumsteg, D., Wieser, H. G., 2000. Presurgical evaluation: Current role of invasive EEG. Epilepsia 41, S55–S60.

Index

Appendix A

Curriculum vitae

Christoph Flamm

Date of Birth:	July 25, 1981
Place of Birth:	Vienna, Austria
Citizenship:	Austrian

Education

2000	Matura (corresponding to A levels) at HTBLuVA Wien 3
2002 - 2008	Dipl.-Ing. (M.Sc.) in Technical Mathematics, Vienna University of Technology, with focus on Mathematics in Natural Sciences Master thesis: *A model for the length distribution of Actin filaments* supervised by O.Univ.Prof. Dipl.-Ing. Dr.techn. Christian Schmeiser
2009 - 2012	Ph.D. in Mathematics in Natural Sciences, Vienna University of Technology Ph.D. thesis: *Novel methods for epileptic onset zone detection* supervised by Em.O.Univ.Prof. Dipl.-Ing. Dr.techn. Manfred Deistler

Research interests

Applied mathematics, biomathematics, time series analysis, causal analysis, factor models, graphical models

Professional career

2009 - 2012	(FWF) Project Assistant at the Institute for Econometrics and System Theory
2006 - 2008	Tutor at the Institute for Analysis and Scientific Computing
2007	Accenture: Member of the task force for a software roll-out
2007	Austrian Institute of Technology: Trainee in the functional material division
2006	Austrian Institute of Technology: Trainee in the mobility division

List of publications

1. Flamm, C., 2008. A model for the length distribution of Actin filaments.
    Master's thesis, Vienna University of Technology.
2. Flamm, C., Kalliauer, U., Deistler, M., Waser, M., Graef, A., 2012b. Graphs
    for dependence and causality in multivariate time series. In: Wang, L.,
    Garnier, H. (Eds.), System Identification, Environmental Modeling
    and Control System Design. Springer, pp. 133-157.
3. Graef, A., Flamm, C., Pirker, S., Deistler, M., Baumgartner, C., 2012a. A
    physiologically motivated ECoG segmentation method for epileptic seizure
    onset zone detection. IEEE EMBC 2012.
4. Graef, A., Hartmann, M., Flamm, C., Baumgartner, C., Deistler, M., Kluge, T.,
    2012b. A novel method for the identification of synchronization effects
    in multichannel ECoG with an application to epilepsy.
    Submitted to Biological Cybernetics.
5. Flamm, C., Graef, A., Pirker, S., Baumgartner, C., Deistler, M., 2012a.
    Influence analysis for high-dimensional time series with an application
    to epileptic seizure onset zone detection. Submitted to the Journal of
    Neuroscience Methods.