



## DIPLOMARBEIT

# Using EEG waveform complexity to characterize the progress of Alzheimer's Disease

ausgeführt durch

ANDREJ KLIMAN, BSc

Johann Blobnergasse 1/7  
1120 Wien  
Österreich

am

Institut für Analysis and Scientific Computing  
der Technischen Universität Wien

in Kooperation mit dem  
Austrian Institute of Technology

---

unter der Anleitung von

Ao.Univ.Prof. i.R. Dipl.-Ing. Dr.rer.nat. Dr.techn. Dr.scient.med. Frank Rattay

sowie

Univ.Doiz. DI Dr. Heinrich Garn, MSc (Austrian Institute of Technology)

---

Datum

---

Unterschrift

Andrej Kliman: *Using EEG waveform complexity to characterize the progress of Alzheimer's Disease*, Master Thesis in Medical Physics & Imaging , © March 2014

## DANKSAGUNG

---

Ich möchte mich aufrichtig bei meinen Eltern bedanken, deren unterstützende Kraft mir das Studium ermöglicht hat. Danke an meine Geschwister, denn wir haben uns gegenseitig nicht nur ungeheuer großen Rückhalt gegeben, sondern wahrlich beflügelt und uns damit zu jenen liebevollen und umsichtigen Menschen gemacht, die wir heute sind. Ausserdem möchte ich mich für all die unglaublich bereichernden Momente die ich mit meinen Freunden im In- und Ausland teilen durfte bedanken.

Diese Arbeit wurde im Rahmen eines Praktikums am Austrian Institute of Technology (AIT) durchgeführt (Projektnummer 827462), gefördert durch die Österreichische Forschungsförderungsgesellschaft FFG. Die Aufnahmen sind Teil einer großangelegten Studie, der *Prospective Dementia Registry Austria* (PRODEM-Austria), welche von der österreichischen Alzheimer Gesellschaft<sup>1</sup> ins Leben gerufen wurde. Alle Electroencephalogram (EEG) Aufnahmen, die in dieser Arbeit analysiert wurden, stammen somit von Personen, die an Morbus Alzheimer leiden. Ich bedanke mich auch bei Ihnen für ihren Mut und Willen einen wertvollen wissenschaftlichen Beitrag zu leisten.

---

<sup>1</sup> <http://www.alzheimergesellschaft.at/index.php?id=27>



## ZUSAMMENFASSUNG

---

Das Interesse an der Verwendung der Elektroenzephalographie (EEG) in der Diagnostik hat im vergangenen Jahrzehnt stark zugenommen. Gründe dafür sind ein tieferes Verständnis über die Signalentstehung sowie neue Möglichkeiten in der Artefaktbereinigung von Biosignalen. Dies öffnete auch die Tür für neue mathematische Methoden zur Signalanalyse, darunter nichtlineare Maße für diskrete Zeitreihen (siehe Stam (2005)). Aus einer Reihe von Studien, welche die EEG-Signale von Alzheimer-Patienten untersuchten, gingen bisher 3 konsistente Trends hervor: demnach haben Alzheimer-Patienten ein zu niedrigeren-Frequenzen verschobenes Spektrum, eine geringere Komplexität der Wellenform, sowie Differenzen in der Synchronität zwischen den Signalen der verschiedenen Elektroden (siehe Review Paper: Jeong (2004) sowie Dauwels et al. (2010)).

Diese Arbeit untersucht den zweiten Trend, nämlich die Komplexität von EEG-Signalen. Die Auswahl der verwendeten Methoden erfolgte auf Basis einer Meta-Analyse kürzlich erschienener wissenschaftlicher Arbeiten, welche Komplexitätsmaße an Alzheimer-Patienten untersuchten. Es kristallisierten sich jene Maße heraus, welche auf dem informationstheoretischen Prinzip von Entropie basieren (Shannon, 1948), und welche die Signale auf mehreren Zeit-Skalen untersuchen (Costa et al., 2005). Die Grundidee ist dabei folgende: Das Gehirn wird als Blackbox betrachtet, die zu jedem abgetasteten Zeitpunkt einen messbaren Output (nämlich die elektrische Spannung an den Elektroden) produziert. Dieser wird wiederum maßgeblich durch die interne funktionelle- sowie anatomische Gehirnstruktur des Patienten beeinflusst und sollte somit sensibel auf dessen Änderungen reagieren. Zur Analyse wird die EEG-Aufnahme artefaktbereinigt und in 4s lange Epochen geteilt. Die Maße werden dann auf die einzelnen Epochen angewandt, anschließend gemittelt und zwischen den 19 Elektroden und allen 116 *probable AD*<sup>2</sup> Patienten der PRODEM-Datenbank verglichen. Mehrere Biomarker wurden extrahiert und dessen diagnostischer Nutzen in Form von Korrelationskoeffizienten ausgedrückt, welche durch ein lineares bzw. quadratisches Regressionsmodell zwischen Marker und *MMSE*<sup>3</sup> Wert der Patienten ermittelt wurden. Die Resultate der verwendeten Methoden sind konsistent und bestätigen alle den oben genannten Trend: konkret haben Patienten mit niedrigerem *MMSE*, also höherem Alzheimer-Schweregrad, tendenziell eine geringere Komplexität. Jedoch gilt dieser Trend nur für die Komplexität von Datenpunktreihen, welche sich über einen Zeitbereich der Inversen Abtastfrequenz (ca. 4ms) erstrecken. Da nur ein schwacher Trend erkennbar ist, und aufgrund der vergleichsweise

---

<sup>2</sup> *probable AD* ist der medizinische Terminus, der zur Klassifikation von Morbus Alzheimer verwendet wird, siehe Kapitel 3

<sup>3</sup> neuropsychologischer Test, der den Schweregrad der Demenz und somit der Alzheimerkrankheit reflektiert Folstein et al. (1975)

großen Kohorte zeigen die Ergebnisse eindeutig die Grenzen komplexitätsbasierter Methoden auf. Der Hauptgrund dafür liegt in der starken Varianz der Maße, selbst innerhalb eines einzelnen Patienten.

## ABSTRACT

---

The measurement of electric-potential differences on the human scalp, also known as EEG, has regained a lot of interest in the scientific community throughout the last decade. Improved understanding of the origin of brain-oscillations and rapid progress in the fields of artifact removal and Discrete Time-Series (DTS) analysis have yielded new possibilities to use nonlinear methods in search of diagnostic biomarkers from EEG signals (see Stam (2005)). For people suffering from Alzheimer's Disease (AD) three common trends concerning the nature of EEG signals have been identified by a wide amount of research groups: a shift of the power spectrum to lower frequencies, a decreased complexity of the signal waveforms and altered synchrony between the EEG records (see Jeong (2004) and Dauwels et al. (2010) for reviews).

This thesis investigates the second trend, i.e. the irregularity or complexity of the EEG signals. The choice of methods reflects a selection of the most promising markers according to recent articles and reviews that used complexity based measures. The final implemented algorithms are based on the information theoretic concept of entropy (Shannon, 1948) and analyze the signals across multiple temporal scales (Costa et al., 2005). The main idea behind the approach is to consider the brain as a black box which - for each sampled timepoint - produces measurable outputs (electric voltages on the scalp). The form of these outputs should be therefore sensible to changes in the internal functional- and anatomical structure. Each EEG record is preprocessed to remove artifacts and split into 4s segments called epochs. The methods were applied to- and averaged over the resulting DTS of all epochs, and compared between 19 electrodes of all 116 subjects from the PRODEM database classified as *probable AD*. Several biomarkers were extracted and their predictive strength evaluated by the correlation coefficient (coefficient of determination) of a linear (quadratic) regression model between marker values and the patients' Mini-Mental-State-Examination (MMSE)<sup>4</sup> scores. The results of different methods are consistent and confirm the observed trends mentioned above: patients with lower MMSE score and hence more severe AD tend to have less complex waveforms in their EEG records. However this trend is only observable when looking at the regularity of datapoint-successions which span over a time in the order of the inverse sampling frequency (4ms). Since only a weak trend can be observed, but a comparatively big study cohort was used, this thesis clearly demonstrates the limits of complexity based measures. The main reason for their shortcoming is the high variability of the measure between epochs of a single patient.

---

<sup>4</sup> a neuropsychological test reflecting the severity of dementia and therefore the disease, see Folstein et al. (1975)





# CONTENTS

---

<b>i</b>	<b>INTRODUCTION AND BACKGROUND</b>	<b>1</b>
1	INTRODUCTION	3
2	FROM NEURONS TO EEG SIGNALS	5
3	ALZHEIMER'S DISEASE	11
4	EEG IN AD	13
<b>ii</b>	<b>METHODOLOGY</b>	<b>15</b>
5	ABOUT THIS STUDY - DATABASE INFO & ORIGIN OF EEG SIGNALS	17
6	COMPLEXITY OF DISCRETE TIME-SERIES	21
7	STATISTICAL ANALYSIS	31
<b>iii</b>	<b>RESULTS AND INTERPRETATION</b>	<b>33</b>
8	MSE RESULTS	35
9	AMI RESULTS	45
10	CONCLUSIONS	53
<b>iv</b>	<b>APPENDIX</b>	<b>55</b>
A	CMI FOR INTER-CHANNEL ANALYSIS	57
B	BURST ANALYSIS	63
	BIBLIOGRAPHY	66

## ACRONYMS

---

<b>AD</b>	Alzheimer's Disease
<b>AMI</b>	Auto Mutual Information
<b>AP</b>	Action Potential
<b>ApEn</b>	Approximate Entropy
<b>CMI</b>	Cross Mutual Information
<b>ECG</b>	Electrocardiogram
<b>EEG</b>	Electroencephalogram
<b>EOG</b>	Electrooculogram
<b>HC</b>	Healthy Controls
<b>IAF</b>	Individual Alpha Frequency
<b>MCI</b>	Mild Cognitive Impairment
<b>MI</b>	Mutual Information
<b>MMSE</b>	Mini-Mental-State-Examination
<b>MSE</b>	Multi Scale Entropy
<b>DTS</b>	Discrete Time-Series
<b>SaEn</b>	Sample Entropy
<b>ShEn</b>	Shannon Entropy

## Part I

### INTRODUCTION AND BACKGROUND

The first part of this thesis starts with the motivation and the basic physiological background, hence the origin of EEG signals (chapters [1](#) and [2](#)). This is followed by a short overview about Alzheimer's Disease (chapter [3](#)). Finally we summarize empirical findings of how the disease affects EEG signals (chapter [4](#)).



## INTRODUCTION

---

Recent advancements in signal processing and a new generation of biosensor devices have led to the emergence of a new fascinating field of biomarker analysis from biosignals. The human body, from an engineering point of view, is a true masterpiece concerning efficiency and variety of various signaling pathways and information transport. A good example for a very accessible (i.e. easily measurable) biosignal is body temperature, which has been used for ages in order to make diagnostic conclusions about the human well-being (such as fever detection through a medical thermometer). However, the human body, as a highly complex dynamic system, inherits far more physically measurable quantities that can help us to understand its current physical state. Due to the rapid progress in electrical engineering and other disciplines producing new sensor devices, a big palette of bodysignals has now become available to be measured and investigated. Among them are those signals, that are based on the measurement of electrical activity or to be more precise, differences in electrical potential at different locations of the body. Electric signaling pathways play a significant functional role in our body, since the regulation and functioning of many body-parts is coordinated by the central nervous system, which consists of the spinal cord and the brain. Thus our body produces electrical fields of varying strength at a variety of different locations, non-stop.

An example for a biosignal based on electrical activity is the well known Electrocardiogram (ECG), which became established as a valuable tool for investigation of the functional state of the heart due to the simplicity and low cost of measurement. The signal is recorded by placing electrodes (i.e. electrically conductive tips that are connected to a voltage difference amplifier) directly on the skin of different body-parts. During various phases of heart contraction, strong electric currents arise<sup>1</sup> which change not only in magnitude but also direction. These currents produce electrical fields, which are (although strongly weakened) still present at the most outer layer of the human body, the skin. Thus comparing the electrical field strength at different locations on the body-skin via electrodes yields patterns that rhythmically change over time in accordance with the heart-beat. The result of this measurement is the famous QRS complex which, when compared with other body parameters, can give astonishingly rich information on possible malfunctioning of the heart.

EEG signals are recorded in a very similar way, hence the method exists already

---

<sup>1</sup> A unique property of muscular tissue of the heart is that it is electrically conductive. Thus, the electrical currents going through this tissue trigger its own contraction and produce strong electrical fields that are well measurable as potential differences on the skin

for quite a while, but does not enjoy the same popularity and clinical relevance as the ECG. To be more precise, Hans Berger was able to detect rhythmic brain activity as early as 1924 by placing electrodes on different locations on the scalp of his son. He observed a clear, rhythmic electrical pattern that changed with the current state of mind of the subject, a typical feature that is nowadays known as *alpha activity*. The principle of measurement here is the same, but the origin of the signals is different: the voltages arise due to many, many small electrical currents in the most outer layer of the brain (the neocortex). However one has to keep in mind, that even on the human skull there are small electrical field differences between nearby electrodes, which are produced by heart contraction current and thus "contaminate" the purely neuronal signal. In addition, various nerve fibres stimulating skeletal muscles in the head can also produce electrical activity that is added to the signals. This leads us to a very important point in biosignal processing. A biosignal always contains a lot of noise, which arises due to many other simultaneous processes which happen in the body during the time of recording. Thus in order to get a "clean" signal, it is desirable to identify and remove influences of other sources as much as possible.

Biosignals give us live feedback on the current state of a patient's bodypart or body and represent an essential part of modern clinical diagnosis. With new tools in signal processing & artifact removal and due to a better understanding of the mechanisms that lead to the origin of the signals, the human EEG - a complex, multivariate type of biosignal - has regained popularity and is nowadays again widely used in the Neurosciences with the aim to possibly detect global alterations of the brain on a functional level. Some people hope that this might enable us to detect certain types of diseases that affect the whole brain such as AD.

This thesis aims to contribute to this vision by exploiting biomarkers<sup>2</sup> that reflect the cognitive decline due to AD. This is a challenging task, since it is known that strong individual differences between EEG records of patients make it generally hard to find robust biomarkers that retain its validity when being compared to a big number of other individuals.

---

<sup>2</sup> a biomarker is a single number or quantity that is deduced from an individual physiological record and hence can be compared among various patients using statistical analysis.

## FROM NEURONS TO EEG SIGNALS

---

### 2.1 TINY POTENTIALS WITH BIG POTENTIAL

It is of fundamental importance to understand the origin of the EEG signals. However, before we go into detail, let's first take a look at the magnitude of the signals we are actually measuring. EEG signals are electrical potential differences, between different spots on the human scalp, each measured by an electrode. But these potential differences are tiny, in fact they are in the range of just a few *micro Volts*! When compared to other biosignals, such as the ECG<sup>1</sup> (which also reflects electric potential differences on the human skin, but now produced by cardiac contraction rather than neuronal activity), it becomes clear that the measured EEG signals are by far smaller in magnitude (approximately by a factor of 1000). Hence noise becomes a major issue in EEG analysis. However, due recent advances in signal processing, artifact removal and also more sensitive and robust electronic equipment, EEG has regained a lot of popularity during the last decade. New open source software and hardware make EEG analysis available for everyone showing interest and a decent amount of computational knowledge. Thus EEG sensors are becoming part of the new generation of personal, affordable, wearable biosensors.

### 2.2 THE ORIGIN OF RHYTHMIC BRAIN ACTIVITY

What do we actually see, when looking at *brain waves* of EEG signals? Interestingly, the last decades of Neuroscience have shed light on the various working mechanisms of the human brain and as it turns out, rhythmic activity or *neuronal oscillations* are now believed to play a major role in information exchange (i.e. communication) between different regions or substructures of the brain. Hence, they seem to be of fundamental importance in order to tackle more complex phenomena like directed attention, perception, memory, consciousness, and all other essential pieces which together shape human behavior. Lets consider the brain as a highly interconnected, dynamic network, with neurons, i. e. living cells as the basic building blocks. Any single firing neuron (that is, an Action Potential (AP)) evokes electrical potential differences at the end of its branches (axons) through the release of neurotransmitters. This in turn influences the probability of triggering another AP in that region either by inhibiting or exciting the postsynaptic

---

<sup>1</sup> ECG signals can be derived in many different ways, by placing electrodes at different body positions. Depending on the placement, voltage magnitudes differ but approximately lie in the *mV* range

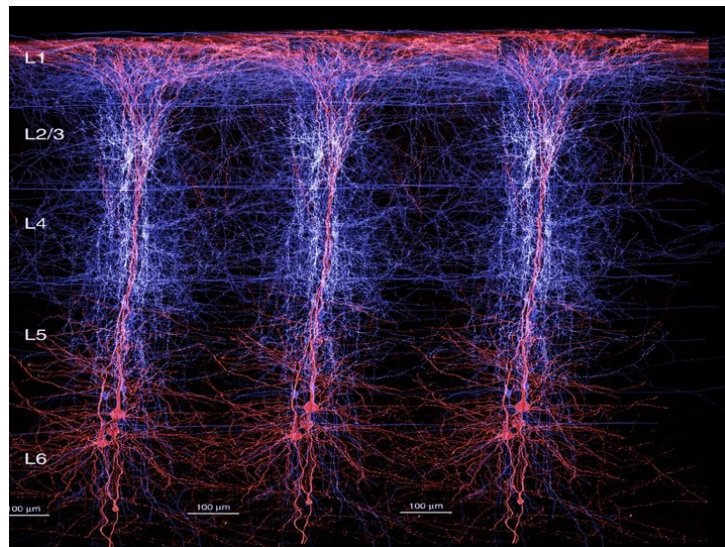


Figure 1: a cross section revealing the six layers of the cerebral cortex as well as the columnar structure created by pyramidal neurons. Pyramidal neurons in red; inhibitory fibers in blue. Image credit: Blue Brain Project; <http://bluebrain.epfl.ch/>

neuron. Since the postsynaptic neuron is also connected to and thus influenced by several other neurons, there is a fine line between triggering or inhibiting an AP. This strong dependence on the conditions<sup>2</sup> at a certain time-instant makes successful signal transduction (at a more global scale) a matter of precise timing between the arrival of the signals from different regions. Successful information exchange between different neuronal assemblies within the brain heavily relies on the timed firing of AP. Rhythms of electrical activity emerge from the interaction of different regions in the brain, where the frequency of the various rhythms are determined by time-delays that arise due to distance between regions, finite conduction speed, but also decay time of postsynaptic potentials. We can imagine that certain frequencies naturally arise in the various networks and sub-networks within the brain depending on those factors (which are not fixed but rather change in different cognitive states - e.g. alertness, sleep-awake, states, etc ...). Small network resonances will depend more on local synaptic properties, whereas in large-scale networks resonance depend more on the wiring itself. This interesting topic of investigating the functional role of various rhythms in the brain has led to a rich body of new methodological approaches that consider the brain as a complex network architecture (see (Buzsaki, 2006)). However, although EEG signals also show clear rhythmic behavior, we have to be very careful in interpreting the functional role of these rhythms.

<sup>2</sup> i. e. neurotransmitter concentration, synapse type (inhibitory or excitatory), membrane potential of the branches (dendrites) of other cell-bodies,...



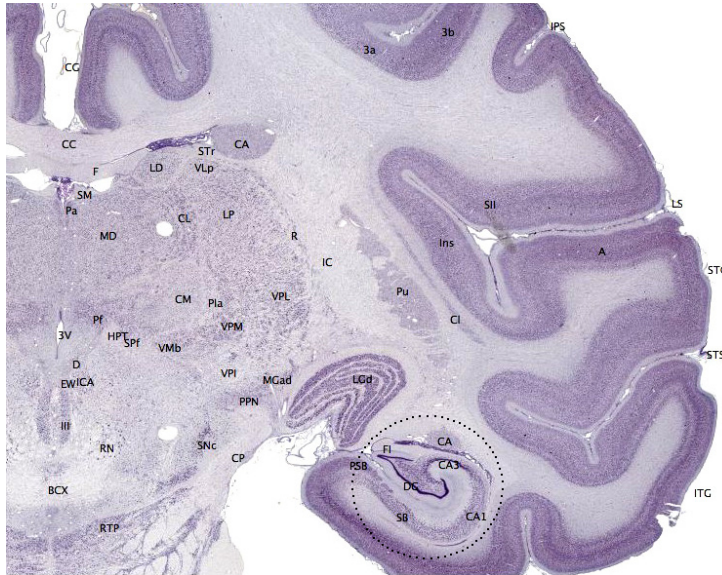


Figure 2: The brain's outer layer, the cerebral cortex, consists of gray matter (depicted in dark violet) and thus acts as a source of dynamic electrical activity which can be measured via EEG. Image credit: [http://en.wikipedia.org/wiki/Cerebral\\_cortex](http://en.wikipedia.org/wiki/Cerebral_cortex)

### 2.3 CORTEX STRUCTURE AND RHYTHMIC SCALP POTENTIALS

An EEG record measures scalp potentials, which show highly oscillatory behavior and hence led to the emergence of the term *brain waves*. However, we have to be aware of the origin of these oscillations in order to interpret them the right way:

The central nervous system (CNS), and hence the brain, has two kinds of tissues: *gray matter*, which contains the cell bodies including dendrites and its synapses; and *white matter* which mainly consists of (myelinated) axons. Thus white matter basically constitutes the (long range) connections between different cell clusters, which are embodied in the gray matter areas (see Fig 2). Within gray matter tissue, the firing of an enormous amount of nerve cells generates millions of tiny electrical impulses (nerve AP) and graded potentials (excitatory and inhibitory post-synaptic potentials). Since the most outer layer of the brain, the cortex, consists of gray matter, we are able to detect electrical activity originating from these regions, even through the skull. Thus a major part of electric potential differences on the scalp is a result of the projection of electric fields from active regions in the cortex. An illustration for a single electrically active site and its projected scalp potential can be seen in Fig 3. But how can we explain the rhythmic, wave-forms in EEG records?

The cortex shows a strongly gyrificated<sup>3</sup> structure (see Fig 2). A closer look at the internal structure of this landscape reveals a layered architecture with different types of cells being dominant in the different domains. Axons from so called *pyramidal cells*, are usually aligned normal to the local cortical surface and thus lead to columnar structures (see Fig 1). Modern approaches of modeling the sources of EEG signals consider the cortical sheet as being composed out of such cortical columns (with sizes varying from micro-columns to macro columns). Each column forms a small electrical dipole (perpendicular to the surface) and thus can be considered as a microsource of electrical activity of the cortex. Consequently the values for each recorded amplitude of an EEG signal are the result of the temporal and spatial summation of all these microsourses, which are eventually projected on the scalp and measured via contact electrodes. The continuous on- and off switching of numerous microsourses, combined with the fact that each source forms a directed dipole results in very complex but smoothly varying waveform. The key is that each electrode measures the electrical potential difference between a fixed location on the scalp and a *single reference* electrode usually positioned at the right or left earlobe which is mainly composed of fatty tissue<sup>4</sup>. One can immediately conclude that although there are strict guidelines for the precise placement of scalp electrodes, the actual sources and thus cortical regions that contribute to the final signal of each electrode may vary significantly between different patients due to differences in the anatomical structure alone. In addition, the original signals spread from the origin through the tissue until they eventually reach the scalp (see Fig 3) This is a very important point to keep in mind, when analyzing and comparing inter-electrode differences in the EEG signals.

---

<sup>3</sup> gyrification refers to the extent of folding of the cerebral cortex in mammals as a consequence of brain growth during embryonic and early postnatal development. The external surface of the brain shows various ridges and grooves, termed *gyri* and *sulci* respectively.

<sup>4</sup> a common alternative to the *single reference* method is to use an *average reference*, that is each potential is measured in comparison to the average of all electrodes. This distributes the responsibility over all electrodes, rather than assigning it to only one of them

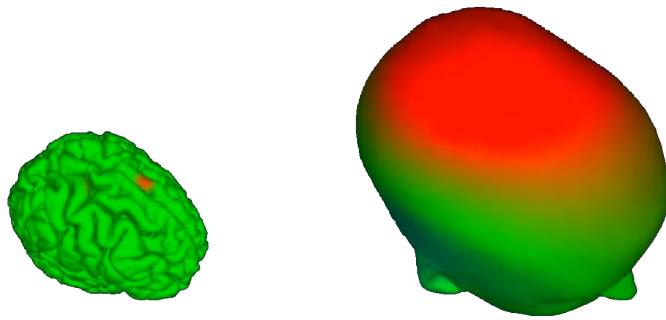


Figure 3: Illustration of how a single cortical microsource influences the electrical potential landscape of the whole scalp through volume conduction. Image credit: <http://sccn.ucsd.edu/eeglab/whatis EEG.html> - extracted from the animation video; see (Acar and Makeig, 2010) for the underlying methodology



## ALZHEIMER'S DISEASE

## 3.1 CHARACTERISTICS

Alzheimer's Disease is a neurodegenerative disease, i. e. the brain of patients suffering from AD slowly degenerates throughout the progress of the disease. This primarily leads to cognitive deficiencies and dementia. Cognitive deficits include impairment of learning and memory, semantic difficulties, deficits in judgement, abstract or logical reasoning, planning and organizing, and, in the late stage of AD, impaired motor functions including chewing and swallowing. Dementia itself is a disorder of cognitive abilities that has increasing prevalence with age. AD is estimated to account for 60-80% of all dementia cases, where hybrid forms with other dementia types occur frequently.

The characteristics of AD are neuronal cell loss, caused by a rapid spreading of neurofibrillary tangles and cortical *amyloid- $\beta$  plaques*, often starting in the hippocampus. Additionally, alterations in transmitter-specific markers including forebrain cholinergic systems are prevalent in AD. The exact reasons for the onset of the disease are still not clarified, however the accumulated plaques and tangles are known to be responsible for the dysfunctioning of neural networks (see e. g. Palop and Mucke (2010)). As from AD diagnosis, the average survival time ranges from 5 to 8 years. Figure 4 illustrates the structural cerebral changes that occur in advanced AD.

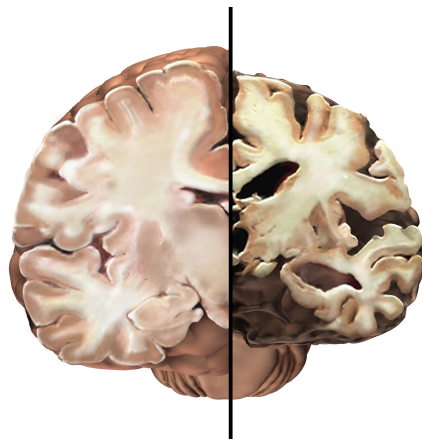


Figure 4: Cerebral slice of a healthy brain and a brain in advanced AD: strong shrinkage, especially in the hippocampus, can be observed in AD. Image credit: ©2015 Alzheimer's Association®alz.org®All rights reserved. Illustrations by Stacy Jannis.

### 3.2 DIAGNOSIS

It is important to note that due to the complex underlying mechanisms, there is no definite in-vivo diagnosis of AD up to this moment. Instead, the disease is classified either as *possible AD* or *probable AD* according to a bunch of criteria (see [McKhann et al. \(2011\)](#)). Therefore, in clinical practice, patients have to undergo a battery of clinical investigations prior to be classified according to the guidelines. Those tests include the assessment of the neurological, internistic and psychiatric status, neuropsychological tests, a complete blood count, and cerebral magnetic resonance imaging (MRI). Additionally, clinical studies suggest genotyping, liquor analysis, serology, imaging procedures such as positron emission tomography (PET) and functional MRI. One can see that it would be desirable to find a biomarker that is able to predict Alzheimer's through a less expensive procedure. This is one of the reasons why the quest of AD biomarkers has gained lots of attention, although up till now no scientific breakthrough could be recorded.

#### 4.1 EEG ABNORMALITIES IN AD PATIENTS

As mentioned in the introduction, recent advances in signal processing resulted in a rich body of literature that discusses EEG as a diagnostic tool. Especially in the field of AD diagnosis, various research groups have tried to increase the sensitivity and specificity of their detection algorithms. In fact, a few clear measurable trends, i. e. detectable alterations in the EEG signals of an average AD patient in comparison with Healthy Controls (HC), have been consistently reported by various research groups over the years (see Jeong (2004) and Dauwels et al. (2010) for very good reviews on the field and (Takahashi, 2013) and (McBride et al., 2014) for recent multi-marker complexity studies). In concrete terms, the most common EEG abnormalities in AD patients are

1. a shift of the power spectrum to lower frequencies
2. a decreased complexity of the signals
3. altered synchrony between the EEG records of different electrodes.

There are several ideas about why and how AD might change EEG signals. The most popular is the *dysconnection* hypothesis of dementia. The loss of neurons in degenerative brain disease may disrupt anatomical connectivity at the level of functional sources and functional networks. Consequently one might expect that brain dynamics in disorders like AD, is characterized by a lower level of synchronization of ongoing brain activity, and that this loss of functional connectivity interferes with normal information processing (see Delbeuck et al. (2003)).

This thesis focuses on a nonlinear methods to measure possible changes in signal complexity related to AD severity. This investigation was initiated by previous findings within the Austrian Institute of Technology (AIT) research group (see Garn et al. (2014)). Their analysis revealed a high correlation between a complexity-based marker (Auto Mutual Information (AMI)) and the scores of a neuropsychological test most commonly used to assess the cognitive decline of AD patients (MMSE, Folstein et al. (1975)). This thesis uses the same approach, that is we compare the evaluation of a patient's cognitive state with complexity of the patients EEG record, rather than looking for significant differences between two groups, AD patients and HCs.

#### 4.2 LINKING THE MMSE SCORE TO AD SEVERITY

Although becoming of age, the MMSE score, is still commonly used to refer to the cognitive decline of AD patients. However, drawing a link between the actual progress of a disease and a psychological test score is dangerous. Furthermore, the MMSE test received criticism due to a possible lack of sensitivity to discriminate mild AD patients from HCs as well as insensitivity to progressive changes occurring with severe AD. On top of that MMSE scores can be strongly affected by demographic factors like age and education (for a review see Tombaugh and McIntyre (1992)). This thesis, being based on the findings of (Garn et al., 2014), tries to exploit biomarkers that correlate to the MMSE values of AD patients. A direct translation from biomarkers of this type to biomarkers being valid for AD classification would therefore intrinsically assume a strong correlation between MMSE and AD severity. Due to an approximate correlation between the two in conjunction with a big study cohort, the results of this thesis can be still placed in the context of AD diagnosis. Hence we expect our findings to be in accordance to the above mentioned EEG hallmarks, but should keep in mind that we have a psychological assessment as a basis for our results.

#### 4.3 GROUP STUDIES - NORMALIZING THE BRAIN

Individual differences in brain anatomy inevitably lead to different types of biosignals, and in the case of EEG, already the shape of the underlying cortex has a strong influence on the electric potential of an accurately placed electrode. In addition, since electric activity reflects the functional state of the brain, differences in the *functional connectivity*<sup>1</sup> also determine the time-course of the measured scalp potentials. The modern neuroscientific way to perform group studies usually includes a so-called *normalization* step of the individual records. Hence the individual records are transformed into a reference space that is made up by a template brain. This process is already well-established for various imaging methods such as fMRI, PET, SPECT,.. studies. However, EEG based studies, including this work, lack a normalization procedure due to the complex genesis of its signals.

To conclude, both anatomical and functional structure of the brain contribute to EEG biosignals. This, on the one hand leads to strong individual differences despite accurate electrode placement. On the other hand, it shows that the full potential of EEG can only be achieved when being combined with records of different modalities.

---

<sup>1</sup> functional connectivity refers to the functionally integrated relationship between spatially separated brain regions, unlike structural connectivity which reflects the anatomical connections (neuron-synapse-neuron network) within the brain.



## Part II

### METHOTODOLOGY

The second part starts with a presentation of the study design, including detailed information about the database, the recording and the preprocessing procedure of the signals (chapter 5). Then an introduction into discrete time series analysis and the information-theoretic concept of entropy is given, followed by a detailed mathematical description of the measures used in this thesis (chapter 6). The last part quickly reviews the concept of linear regression models for biomarker analysis (chapter 7).



## ABOUT THIS STUDY - DATABASE INFO & ORIGIN OF EEG SIGNALS

### 5.1 EEG RECORDINGS

EEG data were recorded from 19 gold cup electrodes placed according to the International 10-20 system (Jasper, 1958). Fig 5 illustrates this unifying approach for electrode placement on the scalp. Electrodes at both mastoids were used as reference electrodes and the ground electrode was located between channels FZ and CZ. Additionally, both horizontal and vertical electro-oculogram (EOG) channels were recorded by electrodes placed above/below the left eye and at the outer corners of both eyes. A wrist clip electrode acquired an ECG signal. The signals were amplified, band-pass (0.3-70 Hz), and notch (50 Hz) filtered by an *AlphaEEG* amplifier (alpha trace medical systems) and digitized at 256 Hz with a resolution of 16 bits. Impedances were kept below 10 k.

All EEG recordings were conducted according to a strict clinically predefined paradigm consisting of two phases. Initially, the subjects were positioned upright in armchairs with integrated neck support in a resting but awake condition with closed eyes (180 seconds). This was followed by a cognitive task with open eyes (130 seconds). Throughout this work, the two recording stages are referred to as *resting phase* and *active phase* respectively. However, the active phase records were not analyzed in this work.

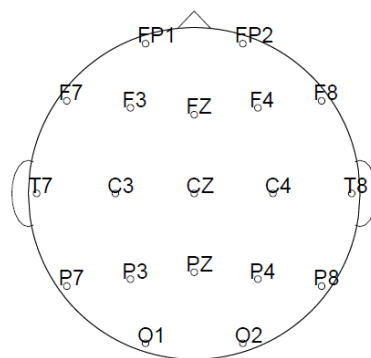


Figure 5: Schematics with the placement and labeling of the 19 electrodes on the scalp, as seen from above (acc. to 10-20 system).

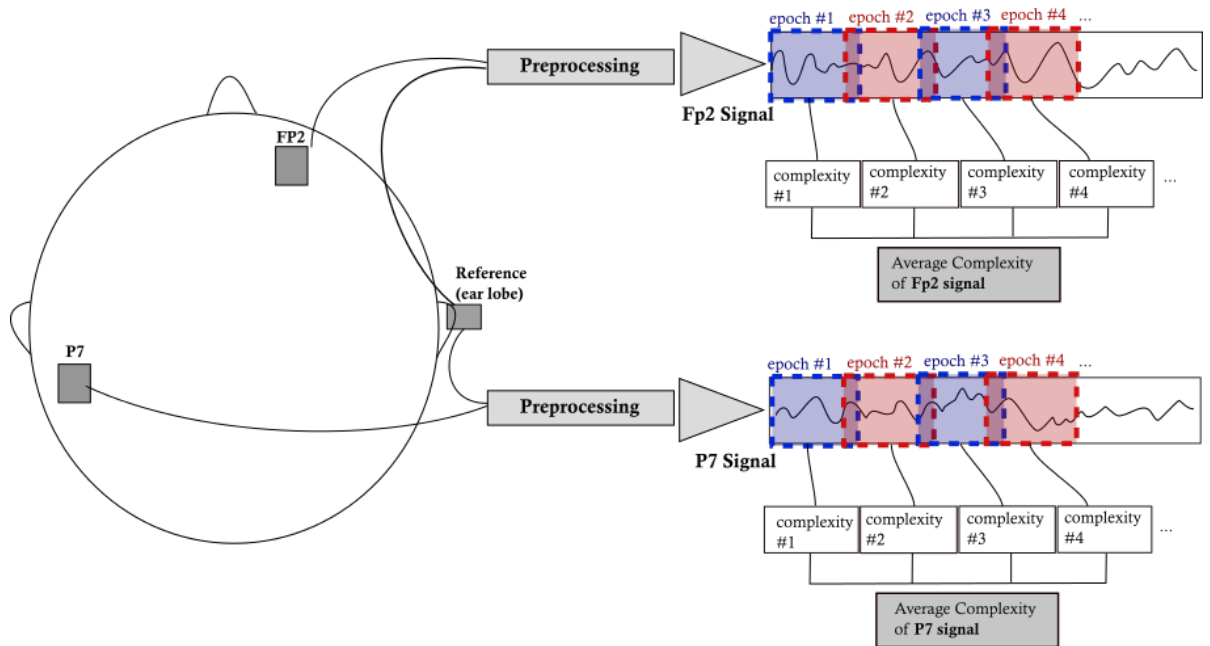


Figure 6: Illustration of the workflow to analyze the 19 channel-EEG records. Each recorded signal is first preprocessed and then segmented into epochs of 4s length (1024 datapoints), with an overlap of 1s (256 datapoints). All complexity measures are applied to all available epochs and then averaged to obtain the final estimate.

### 5.1.1 EEG preprocessing

The original EEG recordings are usually still corrupted by electrical signals of non-neuronal origin. These artifacts arise from either physiological or technical sources. Physiological sources include eye movements and blinking, muscular tension, movement, transpiration, cardiac activity, and talking. Technical artifacts are caused by spurious noise from electronic devices, induction from the mains supply (at 50 Hz), or poor electrode contacts. EEG preprocessing aims at removing these artifacts and obtaining purely neuronal signals. The following preprocessing steps were applied prior to complexity analysis in order to achieve this goal:

**PRE-SELECTION** At first, EEG segments corrupted by non-removable artifacts, e.g. from poor electrode contacts, were visually identified and excluded from further analysis. This led to a reduced, patient-dependent length of the original 180s long *resting phase* records.

**SEGMENTATION INTO EPOCHS** The methods used in this study can only be applied to stationary<sup>1</sup> time-series. However, EEG signals are in general non-

<sup>1</sup> a stationary process is a stochastic process whose joint probability distribution does not change when shifted in time. In other words, the mean does not drift over time.

stationary. This is why for each patient, *quasi-stationary* 4-second segments with a 1-second overlap were extracted from the resting phase record. All methods were applied to these epochs. (see Fig 6)

**HIGH-PASS FILTERING** The remaining EEG, EOG and ECG signals were then digitally high-pass filtered using a stable, direct-form finite impulse response (FIR) filter with linear phase and a border frequency of 2 Hz. Here, any non-neuronal trends and low-frequency artifacts - e.g. from transpiration - were removed from the signals.

**REMOVING CARDIAC ARTIFACTS** Next, artifacts originating from cardiac activity were approached. These artifacts appear mostly in multiple EEG channels as near-periodic spikes, affecting the EEG signals in a broad frequency range due to their non-sinusoidal waveform and the resulting harmonics. The cardiac artifacts were removed by applying the so-called modified Pan-Tompkins algorithm that makes use of the ECG signal for detecting the locations of the cardiac spikes (Waser and Garn, 2013).

**REMOVING OCULAR ARTIFACTS** Eye-induced artifacts result from blinking and ocular movements and affect mostly the signals of frontal and fronto-temporal EEG channels. The eye-induced artifacts were removed by utilizing the EOG channels that captured blinking and ocular movements.

**LOW-PASS FILTERING** Finally, the EEG signals were digitally low-pass filtered using a stable, direct-form FIR filter with linear phase and border frequency 15 Hz. In this way, high-frequency artifacts, e.g. from muscle tension, were removed from the EEG. The border frequency of 15 Hz was determined due to the observation that muscular induced artifacts altered the EEG signals from 15 Hz upwards.

## 5.2 STUDY SUBJECTS

All patients in the database were participants in the multi-centric cohort study Prospective Dementia Registry Austria (PRODEM-Austria) of the Austrian Alzheimer Society. Enrollment criteria included the availability of a caregiver, written informed consent of each participant and caregiver, as well as the absence of comorbidities affecting the conduction of the study. Clinical assessments including EEG recordings were conducted at the Medical Universities of Graz, Innsbruck, Vienna, and the General Hospital Linz, each of them complying with a homogeneous study protocol.

The methods used in this study were applied to all patients from the PRODEM database who were diagnosed with *probable AD* according to NINCDS-

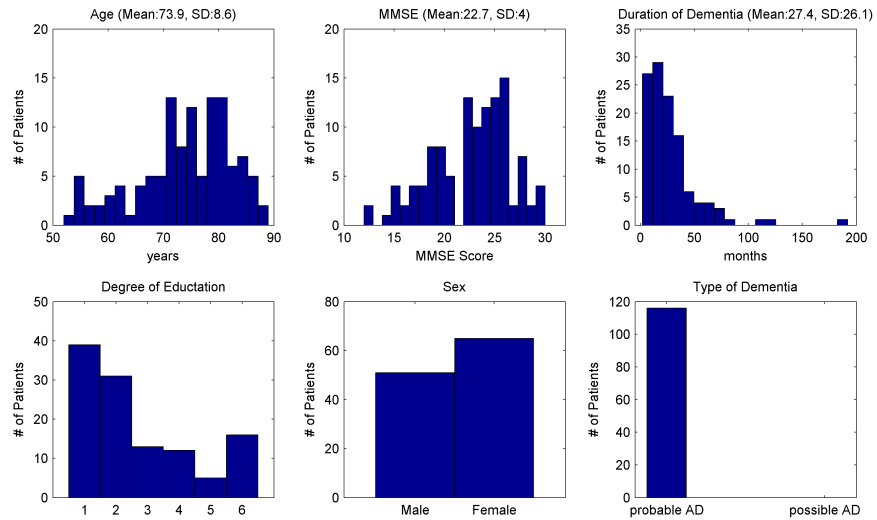


Figure 7: Study subject Information. The group shows all patients from the PRODEM database, that were classified as *probable AD* with a resting phase EEG record available. This resulted in a total of 116 patients used throughout the analysis.

ADRDA criteria (McKhann et al., 2011) and who had an artifact free<sup>2</sup> resting phase EEG record available. This resulted in a total of 116 patients (65 female, 51 male). The subjects were aged between 52 and 89 years (Mean:73.8, SD:8.6) with a duration of probable AD ranging from 2 to 192 months (Mean=27.4, SD=26.1). Additionally, each subject's highest completed level of education was classified on a scale of 1 (primary school) to 6 (tertiary institution). Cognitive deficits were evaluated by MMSE on a scale of 0 to 30 with lower scores indicating more severe cognitive impairment (Folstein et al., 1975). The study subjects reached MMSE scores between 12 and 30 (Mean = 22.7, SD = 4.0). Fig 7 shows a graphical presentation with more details about the study cohort.

<sup>2</sup> artifact free in the sense that all preprocessing steps have been applied to the original signals

## COMPLEXITY OF DISCRETE TIME-SERIES

---

### 6.1 EEG SIGNAL AS DISCRETE TIME-SERIES - FINGERPRINTS OF A COMPLEX SYSTEM?

A full EEG record contains the signals from all electrodes, which were recorded for a certain time  $T_{\text{rec}}$  with a certain sampling rate  $f_{\text{samp}}$ . The original records of each channel are non-stationary time series, with finite-length and high level of noise, since they are biological signals. After preprocessing those signals, as described in the previous chapter, we obtain so called EEG epochs of length  $N_T = T_{\text{epoch}} \cdot f_{\text{samp}}$ , which are - mathematically speaking - discrete sets of data-points:

$$\mathbf{z} = (z_1, z_2, \dots, z_k, \dots, z_{N_T})$$

, hence DTS. It is obvious that all measures that intend to analyze these epochs as representatives of the original EEG signal, will intrinsically depend on  $T_{\text{epoch}}$  and  $f_{\text{samp}}$ .

The idea behind trying to quantify the complexity of those signals is that they originate from the same complex common system: the brain, or to be more precise the neocortex (i.e. the most outer layer of the brain). The cortical network can be regarded as a stochastic, nonlinear system (Stam, 2005) with an enormous number of interacting units, neurons. It can take on various states, which (now looking at it from a statistical physics point of view) could be theoretically described by a point in  $m$ -dimensional phase space. Of course there is no way to determine the value of those  $m$  variables apart from the fact that a prior characterization of what those variables actually are would be necessary.

However, several methods have been specifically developed for DTS analysis of complex systems, which can be applied without having any knowledge about the actual  $m$ -variables. These methods allow to infer information about the system by considering the measured signals (which are DTS) as a set of fingerprints. In other words, we consider the system (i.e. the brain) as a black-box, with measurable outputs that might give us clues of what is actually happening inside. In this work we wish to infer possible signs of cognitive dysfunctions of the brain related to the severity of Alzheimer's Disease.

A lot of progress has been made in the field of nonlinear time series analysis and a big toolbox has emerged, with several new methods that focus especially on analysis of biological signals. The following sections introduce the concepts of different measures that have already been applied by various research groups to investigate possible alterations in the complexity of EEG records for patients

with AD. The methods presented here were chosen after performing an in-depth meta-study about recent research results concerning the use of nonlinear- and complexity-based measures in AD related studies. Interestingly the final choice of methods are all based on the concept of entropy from information theory<sup>1</sup>. Thus the following section comprises a short, intuitive introduction to this concept.

## 6.2 THE CONCEPT OF SHANNON ENTROPY

Entropy is a central quantity in physics and information theory. In information theory, the concept was introduced by Shannon (Shannon, 1948), defining entropy as a measure of the uncertainty associated with a random variable. The concept goes deep and the interested reader should consult the above paper or start with the article on Wikipedia for a more detailed presentation of this topic. To get a rough idea imagine a fictitious machine (i. e.a black box) that automatically shuffles a previously inserted coin, resulting in two different final outcomes: heads or tails. We don't know anything about the underlying architecture, but we can analyze the outcome itself concerning its information content: Each output constitutes an event. Each event has different outcomes with different probabilities. If an outcome has high probability, it is very likely to observe it more often than an outcome that is less probable. Let's define a random variable  $X$  as the amount of observed heads outcomes. Moreover, let's assume that we don't know if the machine is fair or not, thus it has a certain probability  $p(x_1) = p$  for heads and  $p(x_2) = 1 - p$  for tails ( $p$  could take on any value between 0 and 1). If we define the information content of the event  $i$  as  $I(x_i) = \ln\left(\frac{1}{p_i}\right)$ , then the *average information content* can be written as

$$H = \sum_i p(x_i) \cdot I(x_i) = \sum_i p(x_i) \cdot \ln\left(\frac{1}{p(x_i)}\right) = - \sum_i p(x_i) \cdot \ln(p(x_i)) \quad (1)$$

also known as *Shannon Entropy*, denoted by  $H$ . Depending on the base of the logarithm used in equation (1), the unit of this quantity can be measured in *bits* or *shannon* (base 2), *nat* (base e) or *hartley* (base 10). In the simple example from above, we have only 2 outcomes  $x_1$  (heads) or  $x_2$  (tails). Now, depending on the type of machine (i. e. for different values of  $p$ ),  $H$  will vary according to

$$H(p) = -\left(p \cdot \ln(p) + (1 - p) \cdot \ln(1 - p)\right) \quad (2)$$

. Plotting  $H$  for all possible values of  $p$  (see Fig. 8), we can see that the entropy is maximized when the machine is a fair coin tosser ( $p=0.5$  reflects equal probabilities for heads and tails), and minimized for an unfair system, hence if one of the two outcomes becomes very small. This example gives a nice intuitive picture to understand the information theoretic concept of entropy. It can be understood

<sup>1</sup> The concept of entropy actually originates from statistical physics, where it is used to describe the behavior of thermodynamical systems via the entropy-maximization principle. It can be seen as a fundamental scalar quantity, just like energy.



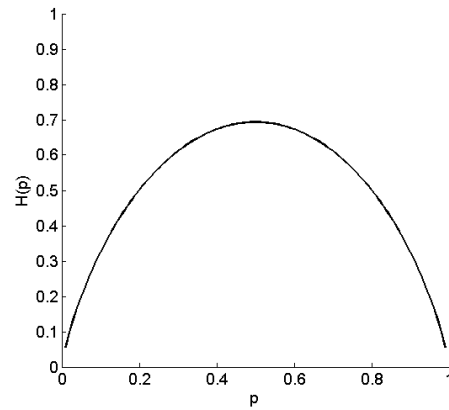


Figure 8: Shannon Entropy  $H(p)$  for different probabilities  $p$  of heads, reflecting the unpredictability for different types of machines characterized by  $p$ .

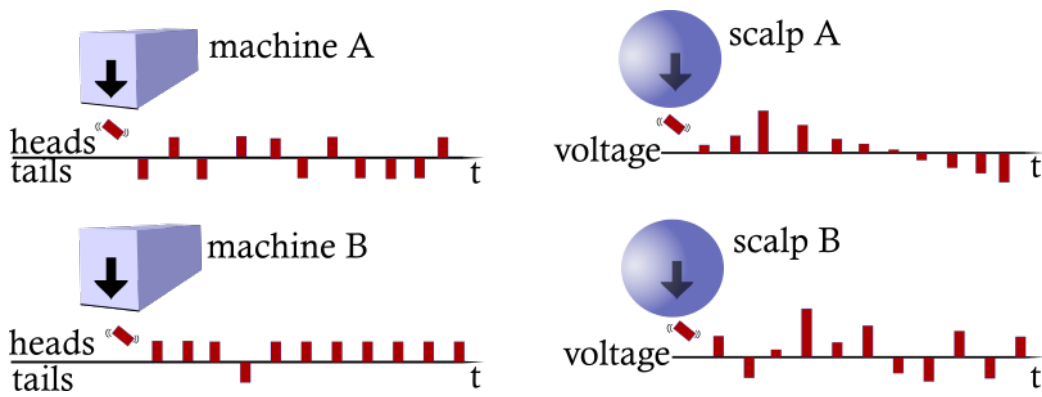


Figure 9: Illustration of how DTS analysis can give information about the function of an unknown underlying system. Here system A and B reflect a fair and unfair coin tossing machine (see text). In the case of EEG analysis, we hope to infer information about the well-being of the patient, i.e. patient A and B reflect patients with and without AD

as a measure of uncertainty or *unpredictability*. A machine with perfectly equal probabilities is completely unpredictable (high entropy), whereas a machine that favors a certain outcome over the other leaves room to predict the outcome (low entropy).

In the framework of DTS, we can use the same concept. The unpredictability, or randomness of a system's output (i.e. a signal which can be measured as a DTS), is often associated with the underlying system *complexity*. In other words, if a simple system gives predictable outputs, whereas a complex system is very unpredictable. Although this association is not entirely correct, it is the reason why the two terms are often used interchangeably and the terminology *complexity analysis* was used by various research groups investigating EEG signals with entropy-based measures.

## 6.3 METHODS FOR ESTIMATING THE UNPREDICTABILITY OF DTS

6.3.1 Shannon entropy *ShEn*

To calculate *ShEn* of a given *DTS* of finite length, we need to know the probabilities of the certain outcomes. In the case of a single channel *EEG* signal, we consider the observed voltage amplitudes  $z_i$  as the random variable, which can take  $N_{\text{bit}}$  possible values (depending on the digitization rate, in our case 16bit, thus  $N_{\text{bit}} = 2^{16} = 65536$ ). Thus at each each sampling-point the system puts out a certain voltage  $z$  with a certain probability  $p(z)$ . However, since we do not know the actual probabilities, we have to estimate them from the observed data. A straightforward possibility is to simply determine the relative occurrence of each possible  $z_i$ . Of course, the longer the *DTS*, the better the estimation. Additionally it is important to make sure that the *DTS* is stationary, i. e. doesn't contain any drifts where the mean amplitude changes slowly over time. This would lead to a bad estimation of the probabilities and thus a bad estimate for the entropy  $H$ . For this reason, the *EEG* is split up into quasi-stationary epochs, i. e. short time segments of the original record. The length of these sections should be short enough to make sure they are stationary, but also long enough in order to guarantee decent probability estimates. Since our random variable can take on  $N_{\text{bit}} = 2^{16}$  different possible values/voltages, but the number of observations is relatively short ( $1024 \leq 65536$ )<sup>2</sup> it is unlikely to get good probability estimates via a counting procedure. The reason lies in the fact that most of the possible outcomes might not occur at all or only very few times due to the comparably small number of observations within an epoch. Thus in order to increase the relative number of observations, it is aswell possible to reduce the number of possible outcomes via another discretization of the measured voltage amplitudes. A straightforward approach is to partition the range of observed voltage amplitudes into a discrete number of equally spaced bins  $i = (1 \dots N_{\text{bin}})$ , with  $N_{\text{bin}} \ll N_{\text{bit}}$ . Then the probabilities  $p(z_i)$  of each outcome  $i$  is just the relative occurrence within the observation  $\mathbf{z}$ , or the height of the normalized histogram (see Fig 10). Thus, given a certain *EEG* epoch, its corresponding *ShEn* can be calculated via the following algorithm:

## SHEN ALGORITHM

1. Choose a number of bins  $N_{\text{bin}}$  to create a discrete histogram of observed voltages  $z_i$  during an *EEG* epoch (compare with Fig 10). The  $p(z_i)$  are then estimated by the relative height of the  $i$ -th bin in the histogram.

<sup>2</sup> In our case the epochlength was chosen to be 4s. Given a  $f_{\text{sampling}} = 256\text{Hz}$ , 4s correspond to 1024 observation

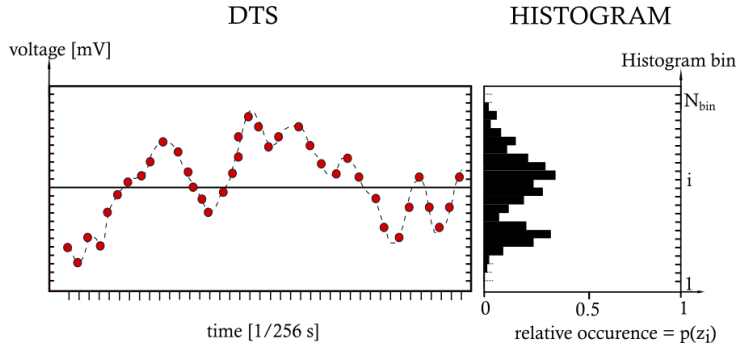


Figure 10: Illustration of the histogram counting procedure. The (quasi) continuous voltages of the DTS are partitioned into  $N_{\text{bin}}$  discrete regions and their relative occurrence estimates the probabilities  $p(z_i)$ . (the figure serves just as an illustration, hence the length of the signal and the created histogram do not match)

2. Calculate Shannon Entropy as

$$H = - \sum_{i=1}^{N_{\text{bin}}} p(z_i) \cdot \ln(p(z_i)) \quad (3)$$

### 6.3.2 Mutual Information MI

Mutual Information (MI) provides a measure of both the linear and non-linear statistical dependencies between two DTS (Jeong et al., 2001). The MI between observations of the two random variables  $z_i^{(1)}$  and  $z_j^{(2)}$ , is the amount of information that the former provides about the latter. This can be expressed formally using the concept of entropy. Since entropy reflects the unpredictability of a signal, we can calculate how much information one signal provides about the other via

$$MI = H(\mathbf{z}^{(1)}) - H(\mathbf{z}^{(1)} | \mathbf{z}^{(2)}) = H(\mathbf{z}^{(2)}) - H(\mathbf{z}^{(2)} | \mathbf{z}^{(1)}) \quad (4)$$

where  $H(\mathbf{z}^{(1)} | \mathbf{z}^{(2)})$  is a *conditional entropy*, hence the entropy of signal  $\mathbf{z}^{(1)}$  given  $\mathbf{z}^{(2)}$ . Using the definition of entropy from eq 1 and using the shorthand notation  $p_i = p(z_i^{(1)})$ ,  $p_j = p(z_j^{(2)})$  and  $p_{ij} = p(z_i^{(1)}, z_j^{(2)})$  eq 4 can be expressed as

$$MI = \sum_i^N \sum_j^N p_{ij} \cdot \ln \left( \frac{p_{ij}}{p_i \cdot p_j} \right) \quad (5)$$

, where  $N$  is the number of possible outcomes of the random variables  $z_i^{(1)}$  and  $z_j^{(2)}$  <sup>3</sup>. If there is no relationship between the compared DTS, the joint probability

<sup>3</sup> Again, in the case of EEG signals,  $z_i$  is the measured voltage. But due to high digitization rate  $N_{\text{bit}} = 2^{16}$ ,  $z_i$  should be restricted to  $N$  possible outcomes by counting all outcomes within a

of observing a pair of random variables  $p_{ij} = p(\mathbf{z}^{(1)}, \mathbf{z}^{(2)})$  will be just the product of both single variable distributions  $p_{ij} = p(z_i^{(1)}) \cdot p(z_j^{(2)}) = p_i \cdot p_j$ . In this case the argument of the logarithm becomes 1 and the expression in the sum yields 0, hence the 2 signals are completely independent. However, any linear or nonlinear correlations between an observed variable pair leads to a positive expression in the sum and thus adds up to the total MI.

In addition one can calculate the time-resolved MI of signal pairs in order to see how the MI changes over time. This way one can estimate, on average, the degree to which the delayed series can be predicted from the original series. Mathematically speaking, this can be achieved by calculating the MI between an EEG epoch  $\mathbf{z}^{(1)} = (z_1, z_2, \dots, z_k, \dots)$  and the time series obtained by shifting  $\mathbf{z}^{(2)}$  by  $\tau$  sampling points  $\mathbf{z}^{(2)} = (z_{1+\tau}, z_{2+\tau}, \dots, z_{k+\tau}, \dots)$ . This way we obtain a MI value for each time-shift  $\tau$ , leading to the MI function  $MI(\tau)$ . In the case of EEG signal analysis, there are basically two different ways to apply this measure. If  $\mathbf{z}^{(1)}$  and  $\mathbf{z}^{(2)}$  are simultaneously recorded epochs from *different* electrodes, one obtains the so-called Cross Mutual Information (CMI), whereas the term AMI is used to describe the MI between epochs from the *same* electrode. This leads to the following interpretation of the two measures:

AMI measures the predictability of the time-course in a given signal. In a very regular signal that shows the same patterns over and over, it is easy to predict how the signal will look like in the future. Thus when calculating the the AMI between the original and time-shifted signal, the mutual information stays high for a large number timeshifts  $\tau$ , since the two versions of the signal still share the same characteristics and the same patterns for a big time-window. In contrast, for a signal that quickly changes its characteristics over time, it is very hard to predict the upcoming signal shape even for a few timepoints, and therefore the AMI function values drop quickly for increasing  $\tau$ . AMI is a tool to estimate how fast a signal loses its linear and nonlinear correlations with the time-shifted versions of itself. The main characteristics of the AMI function is the strength of decrease which tells us how fast mutual information within a single DTS is lost. Hence it contains valuable information about the dynamics of the signal-generating system.

CMI measures the change of correlations between a signal pair of different origin. In the case of EEG the CMI function describes how the linear and nonlinear correlation between signals of different electrodes changes over time. Thus the shape can be more complex than the AMI function. The time-course of this function can contain valuable information about characteristic time-delays and synchronous activity (see A).

---

certain range as a single outcome. A proper choice of N allows decent probability estimations from the observation

## AMI/CMI ALGORITHM

1. Separately determine the probability densities  $p(z_i^{(1)})$  and  $p(z_j^{(2)})$  from the observed occurrences of random variables  $z_i^{(1)}$  and  $z_i^{(2)}$  via histogram counting (like in 6.3.1).
2. Determine the joint probability density  $p_{ij} = p(z_i^{(1)}, z_j^{(2)})$  from the observed occurrences of variable pairs via histogram counting.
3. Calculate the current AMI/CMI value via

$$\text{AMI/CMI} = \sum_i^{N_{\text{bin}}} \sum_j^{N_{\text{bin}}} p_{ij} \cdot \ln\left(\frac{p_{ij}}{p_i \cdot p_j}\right)$$

4. Repeat the above steps with a time-shifted version of  $\mathbf{z}^{(2)}$  for all desired  $\tau$

## 6.4 APPROXIMATE ENTROPY (APEN)

A (ApEn) is a technique for DTS used to quantify the amount of regularity and the unpredictability of fluctuations over time. It reflects the likelihood that patterns in a given signal will not be followed by similar patterns, thus also describes (un)predictability of the signal time-course. In contrast to the above measures, it is based on comparing patterns between discrete datapoints so it avoids probability estimation. To understand the idea behind ApEn, it is best to go through the algorithm steps listed below:

## APEN ALGORITHM

1. Given a DTS  $\mathbf{z} = (z_1, z_2, \dots, z_i, \dots, z_{N_T})$ , construct a  $m$ -dimensional vector  $\mathbf{u} = (z_i, z_{i+1}, \dots, z_{i+m})$  for each sequence of  $m$  successive points  $z_i, z_{i+1}, \dots, z_{i+m}$ . Depending on the choice of pattern-length  $m$ , one ends up with  $N_T - m + 1$  vectors  $\mathbf{u}_i$ , which represent all the patterns of length  $m$  within the signal.
2. For each  $\mathbf{u}_i$ , count the relative occurrence of similar patterns within the signal  $\mathbf{z}$  using

$$C_i^m(r) = \frac{1}{N - m + 1} \sum_{j=1}^{N-m+1} \Theta(r - d(\mathbf{u}_i - \mathbf{u}_j))$$

where

- $d(\mathbf{u}_i - \mathbf{u}_j) = \max_{k=1,2,\dots,m} (u_{i+k-1} - u_{j+k-1})$  is just the maximum distance between two corresponding amplitudes within the two different patterns
- $r$  is the tolerance, i.e. the maximum allowed difference for counting two sequences to be similar
- $\Theta(\cdot)$  is the Heaviside theta function that will yield 0 if  $r < d(\cdot)$  and 1 if  $r > d$ , i.e. add one count to the sum if the patterns are similar.

3. Calculate the average logarithm of all previously determined  $C_i^m(r)$  via

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln(C_i^m(r))$$

4. Repeat the above steps for  $m+1$

5. Finally calculate the [ApEn](#) via

$$\text{ApEn}(m, r) = \phi^m(r) - \phi^{m+1}(r)$$

Thus [ApEn](#) is proportional to the difference between repetitive occurrences of patterns of length  $m$  and length  $m+1$ . The measure depends on 2 parameters, pattern length  $m$  and tolerance level  $r$ .

## 6.5 SAMPLE ENTROPY (SAEN)

Sample entropy is based on the same concept of [ApEn](#) but excludes self matches. This improvement makes Sample Entropy ([SaEn](#)) largely independent of the signal length and relatively consistent under circumstances where [ApEn](#) is not.

### SAEN ALGORITHM

1. Given a [DTS](#)  $\mathbf{z} = (z_1, z_2, \dots, z_i, \dots, z_{N_T})$ , construct a  $m$ -dimensional vector  $\mathbf{u} = (z_i, z_{i+1}, \dots, z_{i+m})$  for each sequence of  $m$  successive points  $z_i, z_{i+1}, \dots, z_{i+m}$ . Depending on the choice of pattern-length  $m$ , one ends up with  $N_T - m + 1$  vectors  $\mathbf{u}_i$ , which represent all the patterns of length  $m$  within the signal.
2. For each  $\mathbf{u}_i$ , count the relative occurrence of similar patterns within the signal  $\mathbf{z}$  using

$$C_i^m(r) = \frac{1}{N-m+1} \sum_{j=1}^{N-m+1} \Theta(r - d(\mathbf{u}_i - \mathbf{u}_j)), \quad \mathbf{j} \neq \mathbf{i}$$

where  $j \neq i$  excludes self matches and

- $d(\mathbf{u}_i - \mathbf{u}_j) = \max_{k=1,2,\dots,m} (u_{i+k-1} - u_{j+k-1})$  is just the maximum distance between two corresponding amplitudes within the two different patterns
- $r$  is the tolerance, i.e. the maximum allowed difference for counting two sequences to be similar
- $\Theta(\cdot)$  is the Heaviside theta function that will yield 0 if  $r < d(\cdot)$  and 1 if  $r > d$ , i.e. add one count to the sum if the patterns are similar.

3. Calculate the average logarithm of all previously determined  $C_i^m(r)$  via

$$\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln(C_i^m(r))$$

4. Repeat the above steps for  $m + 1$

5. Finally calculate the SaEn via

$$\text{SaEn}(m, r) = -\ln\left(\frac{\phi^{m+1}(r)}{\phi^m(r)}\right)$$

## 6.6 MULTISCALE ENTROPY (ME)

The tools [ApEn](#) and [SaEn](#) both compare neighboring or sequential datapoints of [DTS](#) and therefore automatically inherit a strong dependence on the sampling frequency of the original signal. Of course, since people have been working on biosignal processing for a while now, this fundamental issue is well known within the community and was issued by ([Costa et al., 2005](#)), where the concept of *multiple scales* was introduced. The *scale* of a [DTS](#) can be imagined as the inverse discretization detail of the signal. Given a [DTS](#) recorded with a certain  $f_{\text{sampling}}$ , we can easily increase the scale of a given signal via *resampling* the existing one and end up with a new [DTS](#), that looks like a signal we would have obtained using a lower sampling frequency in the first place. A signal of scale 1 represents the original [DTS](#), whereas the same signal of higher scale represents a coarse-grained version of the signal. Mathematically speaking, at scale 2 each point of the new signal is the average of 2 neighboring points of the original signal, at scale 3 each point is the average of 3 neighboring points, . . . This procedure does not only remove the strong influence of the sampling frequencies, but it also provides new insights since biological signals usually contain information across various scales ([Costa et al., 2005](#)).

## MSE ALGORITHM

1. Create multiple coarse-grained (time-scaled) versions  $\mathbf{z}^{(\tau)} = (z_1^{(\tau)}, z_2^{(\tau)}, \dots, z_i^{(\tau)}, \dots, z_{N/\tau}^{(\tau)})$  from the original series  $\mathbf{z} = \mathbf{z}^{(1)}$ , by gradually increasing the scale parameter  $\tau$

$$z_i^{(\tau)} = \frac{1}{\tau} \sum_{j=(i-1)\tau+1}^{i\tau} z_j^{(1)}, \quad 1 \leq i \leq \frac{N}{\tau}$$

2. For each  $\mathbf{z}^{(\tau)}$ , calculate the desired Entropy measure (e.g. SaEn) to obtain the Multi Scale Entropy (MSE) function

$$\text{MSE}(\tau) = \text{SaEn}(\mathbf{z}^{(\tau)})$$

The actual implementation of the code, was modified slightly in order to increase the length of the upscaled signals, as suggested by (Wu et al., 2013). This allows better and more robust MSE estimation at higher scales.

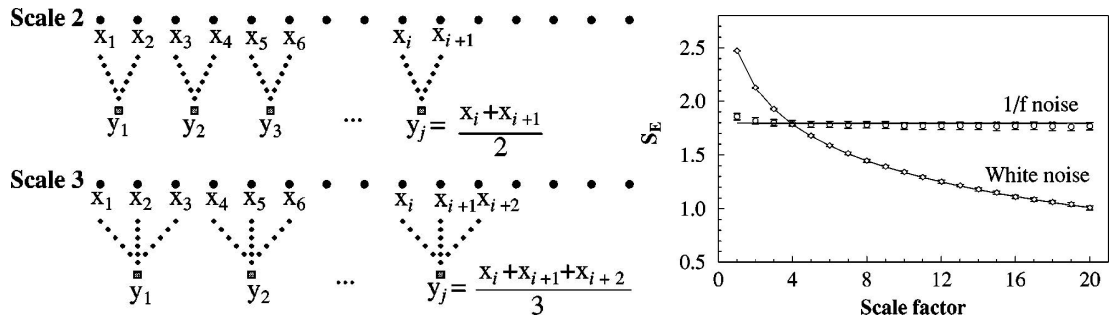


Figure 11: (left): illustration of scaling process from the original DTS; (right): MSE functions of white and  $1/f$  noise for multiple scales; from (Costa et al., 2005)



Diagnostic biomarkers try to predict the current state of a patient. Mathematically speaking, a biomarker is a numeric quantity, from which a prediction model can be formulated. But often in biomarker research, there is no underlying theoretic model which can be used to fit the empiric data. This is because biological systems are very complex due to the enormous number of factors that come into play for biosignal genesis. Therefore to exploit possible correlations between a marker and another quantity that reflects the physical state of the biological system, it is common to simply assume a linear or quadratic relation. This relation can be tested via a regression analysis. Of course this simplifying procedure is more than questionable from a scientific point of view, but given a certain biomarker that - for whatever reason - has a very high correlation with the variable of interest, and this result is even confirmed by other studies, then we have actually found what we were looking for. Thus the predictive value of a biomarker is its most important feature, despite a possible lack of a theoretical foundation.

### 7.1 PREDICTIVE STRENGTH VIA LINEAR REGRESSION

Lets denote the biomarker value of patient  $i$  by  $x_i$  and the corresponding MMSE values by  $y_i$  (the respective means shall be denoted by  $\bar{x}$  or  $\bar{y}$  respectively). Then the linear or quadratic model  $y(x)$  is obtained from the dataset via minimizing the sum of all prediction errors<sup>1</sup>. Given the resulting model one can first quantify the quality of this fit via calculation of the coefficient of determination  $R^2$ , where

$$R = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \Sigma(y_i - \bar{y})^2}} \quad (6)$$

.  $R^2$  describes how much of the observed data variance can be explained by the model, thus a high  $R^2$  is *necessary* for a good biomarker. Given a linear prediction model for a patient's MMSE value, the sensitivity of the model can be quantitatively described by the slope of the model, hence the magnitude  $a$  of the model  $y(x_i) = a \cdot x_i + b$  describes how sensitive the biomarker  $x$  is to small MMSE score differences  $\delta y$ . The quadratic model extends the model by allowing different sensitivity across different marker ranges, for example a marker following a quadratic MMSE prediction model might be very sensitive between patients with

<sup>1</sup> the prediction error or residuum is the difference between the value being predicted according to the model  $y(x_i)$  and the observed value  $y_i$ , therefore  $e_i = (y_i - y(x_i))$

generally low MMSE (i. e. moderate to severe AD patients), but at the same time not able to discriminate between patients with high MMSE scores (less severely diseased patients).

Speaking in terms of our biomarker model, a high  $R^2$  but a small slope is as useless as having a high  $\alpha$  along with a small  $R^2$ . In this thesis, I will therefore present the main results via comparison of different  $R^2$  values, since a high  $R^2$  is the basic requirement for a diagnostic biomarker. However, it is important to keep in mind that the quality and sensitivity of a biomarker are not solely described by this quantity.

## Part III

### RESULTS AND INTERPRETATION

This part contains the results and the interpretation for each selected method onto the PRODEM database, thus the [MSE](#) (chapter 8) and the [AMI](#) (chapter 9) method. Each chapter starts with a short review of previous findings concerning that method. Then the selection of biomarkers deduced from the measure together with the statistical results are presented. Afterwards, an in-depth analysis about the method is given by going through several key points which can explain the outcome of the observed results.

Finally, the last chapter ([10](#)) contains conclusions and final remarks about this study.



## MSE RESULTS

---

### 8.1 PREVIOUS RESEARCH RESULTS

(Escudero et al., 2006) looked at 5s epoch MSE at scales 1-12, observing different type of MSE profiles for the 2 patient groups (8AD vs 8 HC). At scales 1-9, AD patients have lower SaEn, but exhibit higher entropy for scales >10, hence there is a crossing point at around scale 10. To put this into numbers, the authors decided to compare the average slope between scales 6-12, and observed a significantly higher slope in AD patients (positive slope), than in HC patients, (negative slope) at the very frontal (Fp1,Fp2), and occipito-parietal electrodes (P3,P4,O1,O2).

(Park et al., 2007) compared 3 different groups (38HC vs 22Mild Cognitive Impairment (MCI) vs 26AD), and also observed different MSE profiles for the 3 different groups. Again AD patients had lower SaEn for scales 1-8, and exhibited higher SaEn for scales 10 and higher. Interestingly MCI patients showed significantly lower SaEn than HC for scales 2-4, and show a very similarly shaped MSE curve as HC patients, but shifted in scale. The authors did not use single electrode profiles for comparison, so differences reflect the averaged values from all electrodes.

(Mizuno et al., 2010) investigated MSE profiles 2 groups (18HC and 15AD patients with mild and moderate disease severity and observed similar MSE profiles as previous groups (AD have lower SaEn at small temporal scales, but higher SaEn at higher temporal scales beyond crossing point around scales 6-7). Significant differences at small scales (2-4) between all AD and HC were found in fronto-central region (F3,F4,Fz, C3), and all channels except F7,F8 showed significantly lower SaEn at scales 9-20. Additionally a strong (negative) correlation between MMSE and SaEn at large scales (11-20) was found in the central region (C3,C4,Fz).

(Labate et al., 2013) investigated MSE differences among 3 different groups (4HC vs 4MCI vs 4AD) in 3 different regions (frontal, occipital, mixed) for scales 1-5. However he had a very small patient cohort. The biggest observable differences between AD vs HC were found in the occipital region at scales 4-5, where SaEn was around 40-45% higher for HC than AD .

(Yang et al., 2013) investigated MSE profiles for 4 different groups (15HC 15 very mild AD 25moderate/severeAD), divided the MSE profile into 2 parts: *MSE-short* (scales 1-6) and *MSE-long* (scales 16-20), and categorized by their Clinical Dementia Rating (CDR) (HC vs CDR=0.5 vs CDR=1 vs CDR=2+). They observed a significant trend that for increasing CDR, *MSE-short* decreases and *MSE-long* increases, with highest significance in occipito-parietal region, but couldn't differentiate between HC and very-mild AD (CDR=0.5) group. *MSE-short* is significantly cor-

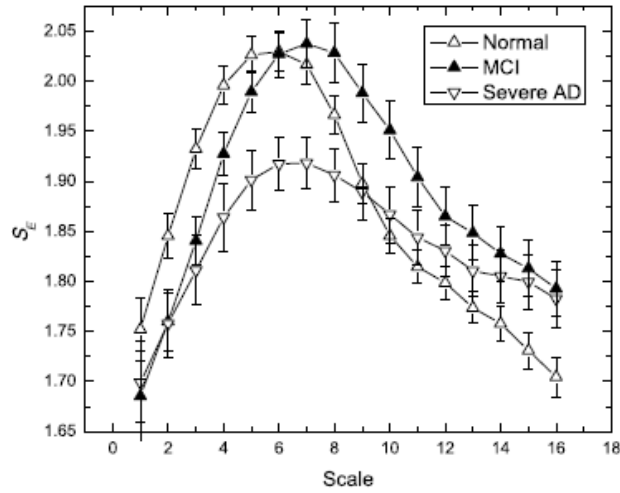


Figure 12: average MSE profile for 3 different groups (from (Park et al., 2007) ). All groups report the same characteristic shape of the MSE curve, as well as a crossing point between HC and AD. All groups reported the same trend: MSE of AD is lower for small scales, and higher for large scales. They also used  $f_{\text{sampling}} \approx 200 - 250\text{Hz}$ , thus the given scales are comparable to ours (256Hz)

related to MMSE at temporal (T5,T6), parietal (P3,P4) and occipital(O1,O2) electrodes, whereas MSE-long is significantly negatively correlated to MMSE at the same electrodes except T6,P4.

## 8.2 MAIN RESULTS

Since the MSE curve defines unpredictability across various scales, we should take into account the MSE values of as many scales as possible. In addition the previous results show that the MSE curves tends to change shape between different groups. Thus we introduce biomarkers that try to quantify information given by the shape of the function by a single numeric value. The final selection of biomarkers is the following:

### BIOMARKERS

1.  $AUC_{\text{full}}$  (area under the curve) takes into account the regularity of the signal across various scales, thus should represent the complexity.
2.  $AUC_{\text{crit}}$  (area under the curve of the MSE function until scale 8) Taking only acMSE values below scale 8 guarantees that no power-spectral influences come into play, and thus should reflect the complexity with less emphasis on power spectral properties (see section 8.3).

3.  $\text{slope}_{\text{crit}}$  (the fitted linear slope of the function at scale  $8 \pm 1$ ) This value is a very simple way to characterize the approximate function shape.
4.  $\text{MSE}_{\text{max}}$  (maximum MSE value) The maximum MSE value reflects the maximum irregularity of the signal. Since each EEG signal differs in frequency architecture, we can analyze each signal across various scales and then determine the unpredictability of the signal independent of scale and thus probably reducing the impact of individual differences.
5.  $\text{MSE}(\tau)$  One can also simply compare the entropy values of each scale between patients. This way we can discriminate between high  $R^2$  values that randomly pop up at a single scale (and thus might result from a non-physiological origin) and those  $R^2$  which are consistent across a certain range of scales.

Fig 13 and Fig 14 summarize the main results from statistical analysis of MSE functions for all patients and all electrodes, averaged over all epochs. Similar Figures were calculated for various electrode groupings but did not give novel insights and thus are not presented in this thesis. From Fig 13 (which represents the results from analyzing only signals from electrode P7) we can see that patients with low MMSE scores tend to have lower entropy (positive correlation) and hence their signals tend to be less complex than the ones of patients with high MMSE scores. This is in accordance with the previous research findings cited above. However, in contrast to previous results, the MSE function doesn't flip its characteristics for large scales. This is also true for all other channels (not shown here). Severe AD patients tend to have a lower entropy than mild AD across all scales as can be seen by the maintenance of the positive correlation across all scales. The reason for this lies in the fact that we restricted our signals to contain only spectral content below 15Hz (for a detailed explanation, see in-depth analysis below 8.3).

Although certain trends are visible, the predictive strength as measured by the coefficient of determination  $R^2$  is rather low and even fails to discriminate between patients with severe and mild form of AD. This can be clearly seen by visual inspection of the scatter plots (Fig 13), where several patients with a low MMSE score have a high biomarker values (in contradiction with the proposed model). Results from all other electrodes along with P7 are summarized in Fig 14. Here, each column represents the results of all scatter diagrams of a single electrode (like Fig 13), where color coding was used to visually represent the  $R^2$  values from the linear and quadratic prediction models. One can see that those regions with similar  $R^2$  values are neighbouring electrodes. The reason for this is the big effect of volume conduction, as discussed in part 1 of this thesis. The  $\text{AUC}_{\text{full}}$  and the  $\text{MSE}_{\text{max}}$  marker confirm the trends observed for other

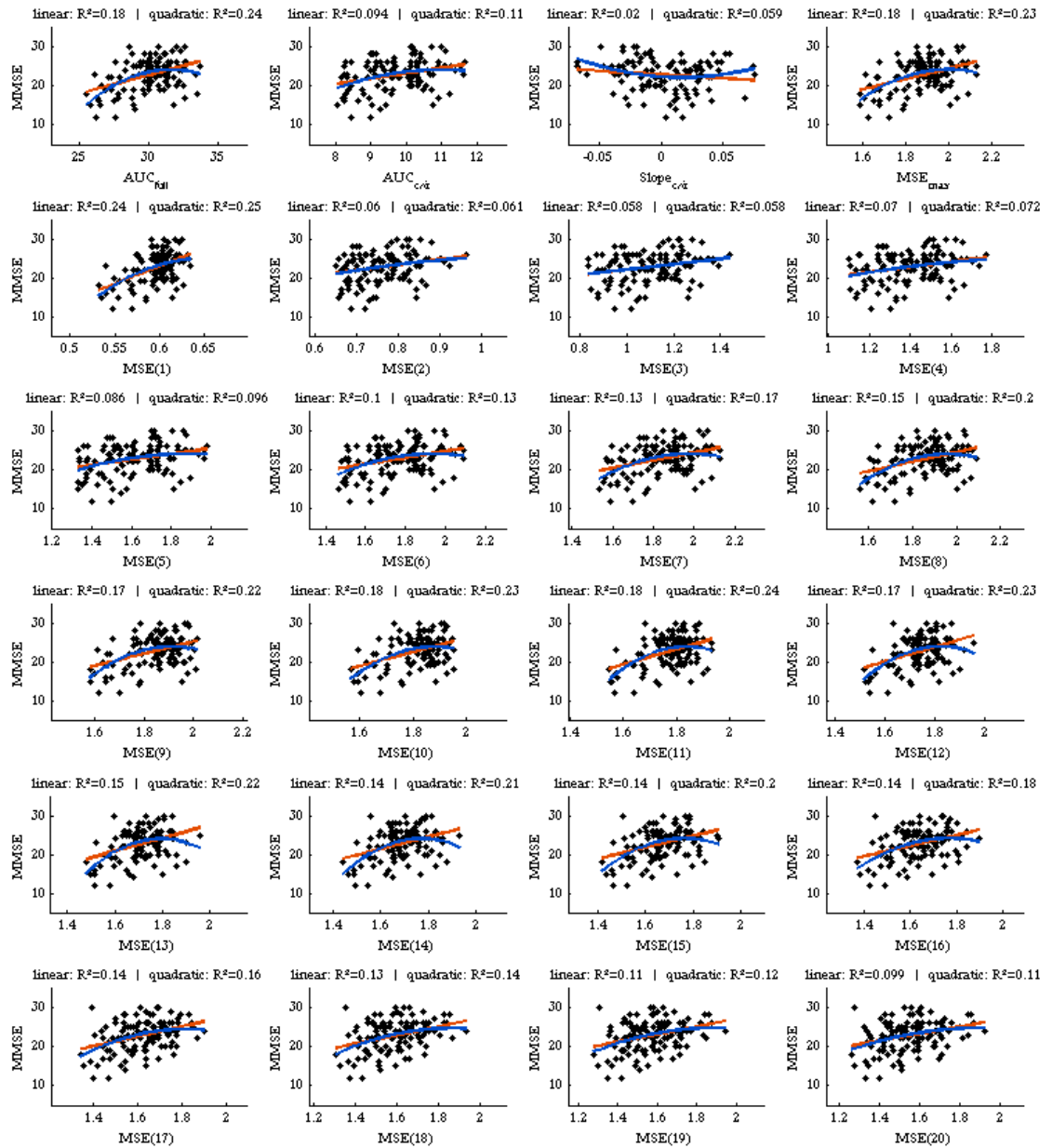


Figure 13: Scatter plots for MSE biomarkers vs MMSE and the linear and quadratic prediction model obtained by regression analysis; for a single channel (P7).



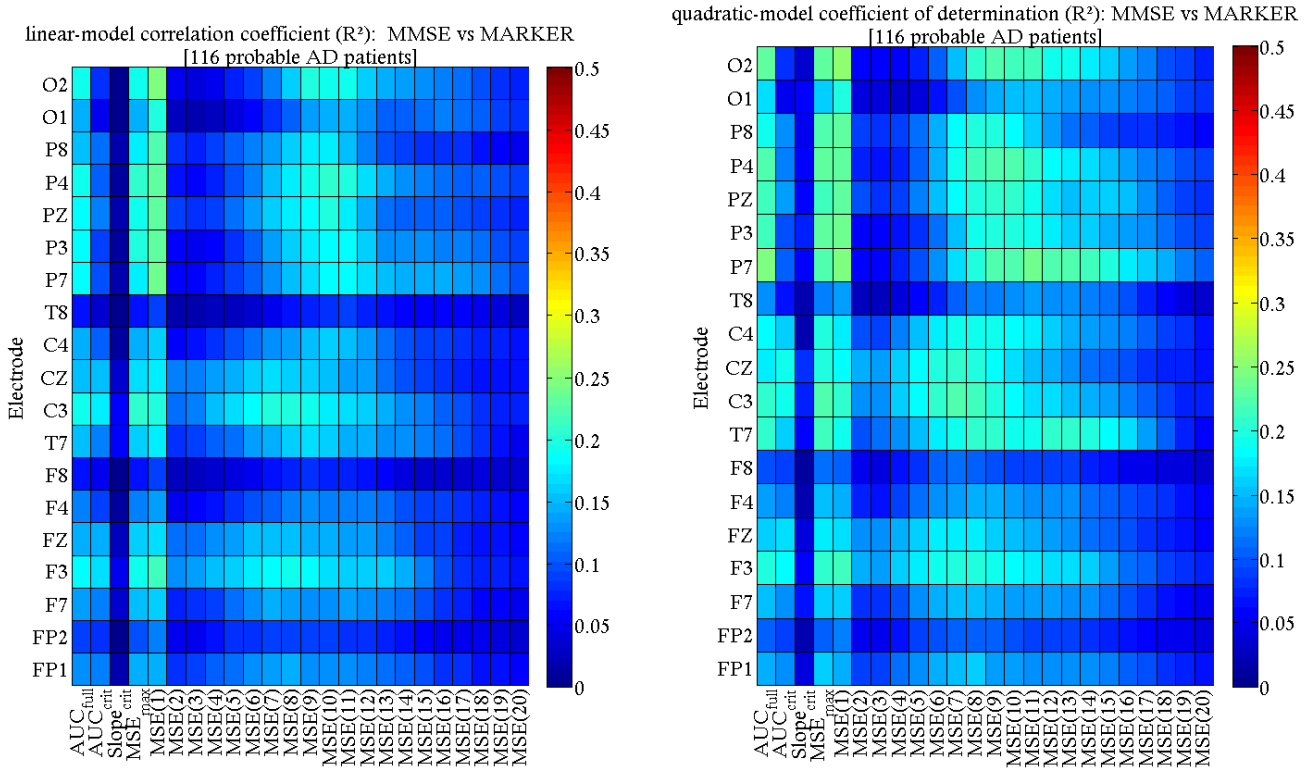


Figure 14: Predictive strength of MSE depicted biomarkers and MSE function measured by the coefficient of determination ( $R^2$ ): linear (left) and quadratic (right) fit to model a possible correlation between MMSE and MSE biomarkers for all scales and channels.

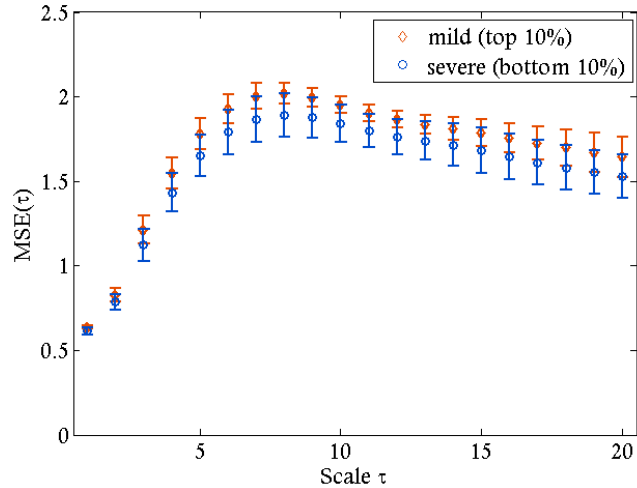


Figure 15: Comparison of the average MSE function (averaged over all epochs, all channels and each patient) between patients with lowest and highest MMSE respectively. The error bar reflects the standard deviation between the scores of individual patients with respect to the mean.

markers, where the  $AUC_{full}$  marker is believed to be the most robust marker containing most information from the MSE function. Interestingly, although the MSE(1) marker exhibits the highest correlations, the results from MSE(2) and MSE(3) cannot support this trend. Given a physiological basis for this observation, we would expect this trend to be stable across various scales since the marker values change according to a smooth transition depicted by the shape of the MSE curve. This transition can also be observed in Fig 13 where the shape of the scatter plots gradually change for increasing  $\tau$ . Even more interesting is the observed increase of  $R^2$  values for  $MSE(\tau > 8)$ . This is the scale where the patients high frequency spectral power content starts to be *washed-out* and thus individual differences in the EEG spectra for  $f < 15\text{Hz}$  start to come into play (for a detailed explanation, see in-depth analysis below 8.3).

Altogether, MSE confirms the trends from previous findings, but fails to discriminate well across the dataset.

### 8.3 IN-DEPTH ANALYSIS

**PARAMETERS** The MSE method, as presented in chapter 6, was applied to the study sample, such that the average MSE function was calculated for each electrode of each patient. The final choice of parameters  $m = 2$ , and  $r = 0.15$  was chosen according to the research literature, and its validity rechecked via consistency analysis on a test sample. The MSE was calculated until scale 20, then the predictive strength of the MSE function of each channel and selected biomarkers was quantified via the coefficient of determination  $R^2$ .

**UNDERSTANDING MSE AT HIGHER SCALE** The multiscale approach extends the complexity analysis by creating an additional dimension which in theory leaves a lot of room for improved classification and discrimination between different severity levels of AD. All groups reported similar MSE profiles for AD and HC patients, where AD patients usually have an MSE profile with a less steep increase, therefore a lower absolute maximum, and a more flat profile at high scales which leads to higher values at higher scales (see Fig 12). It is noteworthy, that despite different recording and preprocessing techniques, the groups achieved somewhat consistent results. Still, there was no consensus about how to interpret this observation.

There is, however a very important point to consider which explains this characteristic behavior. According to the upscaling procedure, when reaching higher scales, the new value is calculated by averaging over such a big number of points that one starts to lose the ability to properly represent higher frequencies whose oscillation period is in the same range as the number of points one is averaging over. This means upscaled versions of the original DTS start to *wash out* high frequency contributions from the EEG signal. To represent a frequency of  $f_0$  Hz, one needs a sampling frequency that is at least as high as  $f_{\text{sampling}} = 2f_0$  (Nyquist theorem). Going up the scale reduces the effective sampling frequency and therefore one also cuts down the frequency spectrum of the signal. For  $f_{\text{sampling}} = 256$  Hz, scales 3-5 already start to neglect  $\gamma$ -band contributions. Hence applying MSE to a non-bandpass filtered signal, the unpredictability of the signal automatically decreases at certain scale since the high-frequency content is neglected. Given the fact that there is a general trend for AD patients to have *slower* EEG spectra (i. e. more spectral power in low frequency band) than HCs, the characteristic flip between the MSE profiles of AD and HC becomes fully understandable. The drop in the *average* MSE in AD patients at higher scales can thus be attributed to a frequency spectrum that is on *average* shifted to lower frequencies. The tight relationship between complexity based measures and methods that directly analyze the frequency spectrum of EEG records is nicely reviewed in (Dauwels et al., 2011).

What does this mean for our analysis? First, we need to check for the maximum frequencies that are present in our signal in order to determine at which scales the frequencies of the original signal are *washed out*. As mentioned in part 2, since we are only interested in analyzing EEG contributions that originate from resting-state brain-activity, each EEG signal was bandpass filtered between 2 and 15Hz in addition to ECG and EOG artifact removal. We exclude any frequencies above 15Hz, thus most of the high frequency activity (commonly termed as  $\beta$  and  $\gamma$  activity) of the patient is neglected since they cannot be distinguished from muscular artifacts. A maximum frequency of 15Hz means that we do not only make sure that the original signal originates from cortical activations, we furthermore eliminate the influence of each patient's high-frequency spectrum in

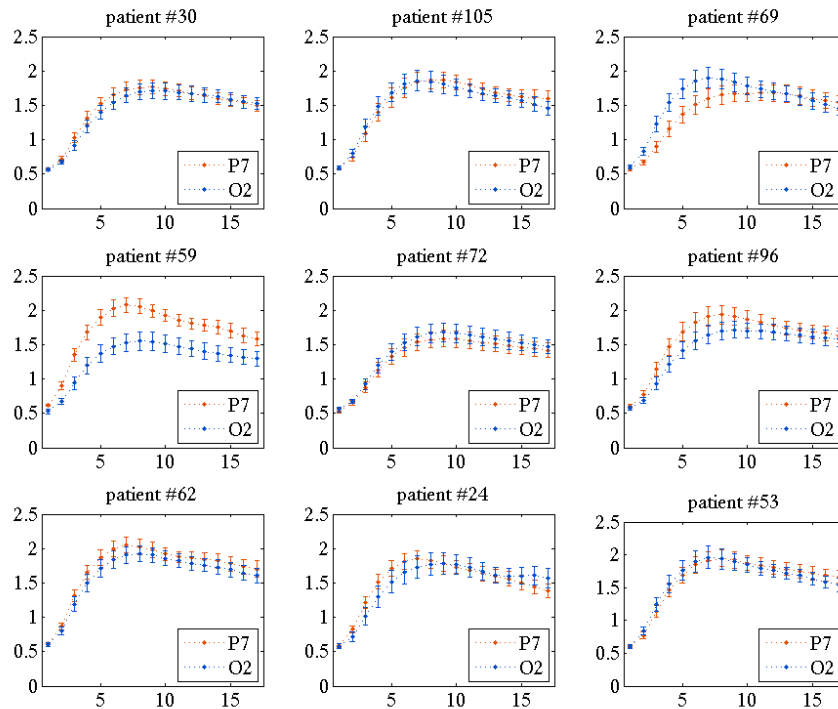


Figure 16: Examples for single-patient, single-channel MSE functions (averaged over all epochs) for 9 random patients and 2 random electrodes. The error bar reflects the corresponding standard deviation between individual *epochs*

the complexity analysis when looking at scales up to  $8^1$ . Let's look at a simple diagram to illustrate the keypoints mentioned above. We use 2-15 Hz signals, thus in contrast to the other groups, we expect a different behavior for large scales, since the high frequency content doesn't play any role here. Fig 15 compares the average MSE plots between the group with the lowest and highest 12 MMSE scores (12 reflects 10% of the dataset; chosen arbitrarily for demonstration purposes only). As one can immediately see, no crossing point is observed! This result confirms the above considerations which provides a nice interpretation of MSE at high scales.

**SINGLE-SUBJECT VARIANCES** Figure 16 shows how much (on average) MSE values can deviate between *electrodes* within a single subject, shown for 9 different patients. In addition the inter-epoch variability is shown in terms of standard deviation from the mean. This observation shows 2 things: first, the unpredictability or complexity of different epochs, even within a channel, change in a dynamic way resulting in a high *inter-epoch variance*. Second, *inter-channel* devia-

<sup>1</sup> given  $f_{\text{sampling}} = 256$ ,  $f_{\text{max}} = 128$  Hz according to the *Nyquist theorem*. Scale 1 can contain frequencies up to 128 Hz. Therefore scale  $\tau$  can only represent frequencies up to  $\frac{128}{\tau}$  Hz, and scale of 8 is the highest scale that still contains all of the frequency content of the EEG records (which is between 2 and 15 Hz)

tions tend to be of the same magnitude as *inter-subject* or *inter-group* differences. One can see that **MSE** exhibits high inter-electrode and inter-epoch changes that are of the same magnitude as inter-subject differences. Thus although **MSE** is able to reveal certain trends, the high variability between epochs and electrodes of a single subject make the method a non-robust measure and thus **MSE** fails to predict functional changes in the brain.



## AMI RESULTS

---

### 9.1 PREVIOUS RESEARCH RESULTS

(Jeong et al., 2001) investigated AMI differences between AD and HC as well as correlation to MMSE scores of patients. For comparison, they also calculated the rate of decrease and observed significantly lower values in AD patients than in HCs at all electrodes except O<sub>2</sub>, with best discrimination in the frontal and temporal region. They also found that the AMI rate of decrease was strongly correlated with MMSE scores of the patients in those 2 regions.

(Abásolo et al., 2008) showed that AMI has a strong negative correlation with ApEn and SaEn. They investigated the rate of decrease for AMI and observed a significantly smaller rate of decrease for severe AD patients in comparison with HCs, especially in the parietal (P<sub>3</sub>,P<sub>4</sub> showed on average 32-35% smaller rates) and occipital (O<sub>1</sub>,O<sub>2</sub>)-temporal (T<sub>5</sub>,T<sub>6</sub>) regions. Although the group used just histogram counting in order to determine the probabilities for AMI, the severity level of AD patients was large enough to observe meaningful differences.

(Garn et al., 2014) is the basis for the topic of this thesis. They investigated the correlation of MMSE scores to AMI, Shannon Entropy (ShEn) and Tsallis Entropy in 79 patients with *probable AD* (hence a subgroup of the PRODEM dataset). The predictive strength of these measures was expressed via coefficients of determination ( $R^2$ ) values from linear and quadratic regression models respectively. Certain values of the AMI function explained up 43% / 48%, ShEn up to 37% / 48% of the variations in MMSE scores, all at left temporal (T<sub>7</sub>) electrode site. The steepest slope of the linear regression was found for the same AMI value. Comparing to traditional slowing measures, complexity measures yielded higher coefficients of determination.

### 9.2 MAIN RESULTS

The *decrease* of AMI function reflects the unpredictability of the signal itself. The papers cited above all modeled the rate of decrease by fitting a linear function through the first datapoints of the AMI-function (some took a certain maximum time-delay; some fitted a line until the first minimum). However, it becomes clear from just looking at the AMI function that the decrease is far from being linear, which is why this estimation is not very precise and therefore the estimate itself inherits a certain degree of randomness. A bad model produces an additional error source, which is costly in the search of robust biomarkers. We should therefore rely on features that reflect the AMI curve in better way.

## BIOMARKERS

1. AUC Instead of the rate of decrease, one can as-well avoid the problem of defining an endpoint for the regression model by calculating the area under the curve (AUC) of the normalized [AMI](#). A small area reflects a strong decrease (high complexity), whereas a large area reflects a slowly decreasing function (low complexity). The AUC was calculated until time-shift 8 since time-shifts below this value do not contain contributions from dominant frequencies from the individual spectra (see in-depth analysis below [9.3](#)).
2.  $ROD_{exp}$  A slightly improved model for the rate of decrease is an exponential fit, via a linear regression of the  $\log(AMI(\tau))$  (fitted through time-delays of 0-4 datapoints)
3.  $ROD_{lin}$  Although a suboptimal model, we also calculated the rate of decrease using the classical method used in the cited literature for the sake of completeness (linear fit through time-delays of 0-4 datapoints)
4.  $nAMI(\tau)$  Of course, it is also possible to simply compare each value of the normalized [AMI](#) function between each patient. However, a single [AMI](#) value reflects the correlations for a distinct time shift, thus taking a single function value as a biomarker reduces the reproducibility of this measure since many random factors come into play and might slightly increase/decrease its value, leading to a bad estimation accuracy and high method-dependency. We still perform this calculation in order to compare the results to ([Garn et al., 2014](#)), where the high correlations were also found by using just a single [AMI](#) function value.

Fig [17](#) and Fig [18](#) summarize the main results from statistical analysis of [AMI](#) functions for all patients and all electrodes, averaged over all epochs. From Fig [17](#) (which represents the results from analyzing only signals from electrode P7) we can see that patients with low [MMSE](#) scores tend to have higher [AMI](#) values (negative correlation). Again this is in accordance with previous research results, hence signal complexity tends to decrease for increasing [AD](#) severity. When looking at [17](#) we can see that the overall behaviour between the depicted biomarkers  $AUC$ ,  $ROD_{exp}$  and  $ROD_{lin}$  is approximately the same, where the  $AUC$  marker is believed to be the most robust marker providing the best model for the [AMI](#) decrease. Compared to [MSE](#) biomarkers, [AMI](#) biomarkers perform better, hence have a higher  $R^2$  and a steeper linear model fit. However, although the same trends become visible, the predictive strength is too low. This becomes visible in Fig [18](#) where again the results for each channel are represented by the color coding the  $R^2$  of the various markers. Interestingly one can also observe



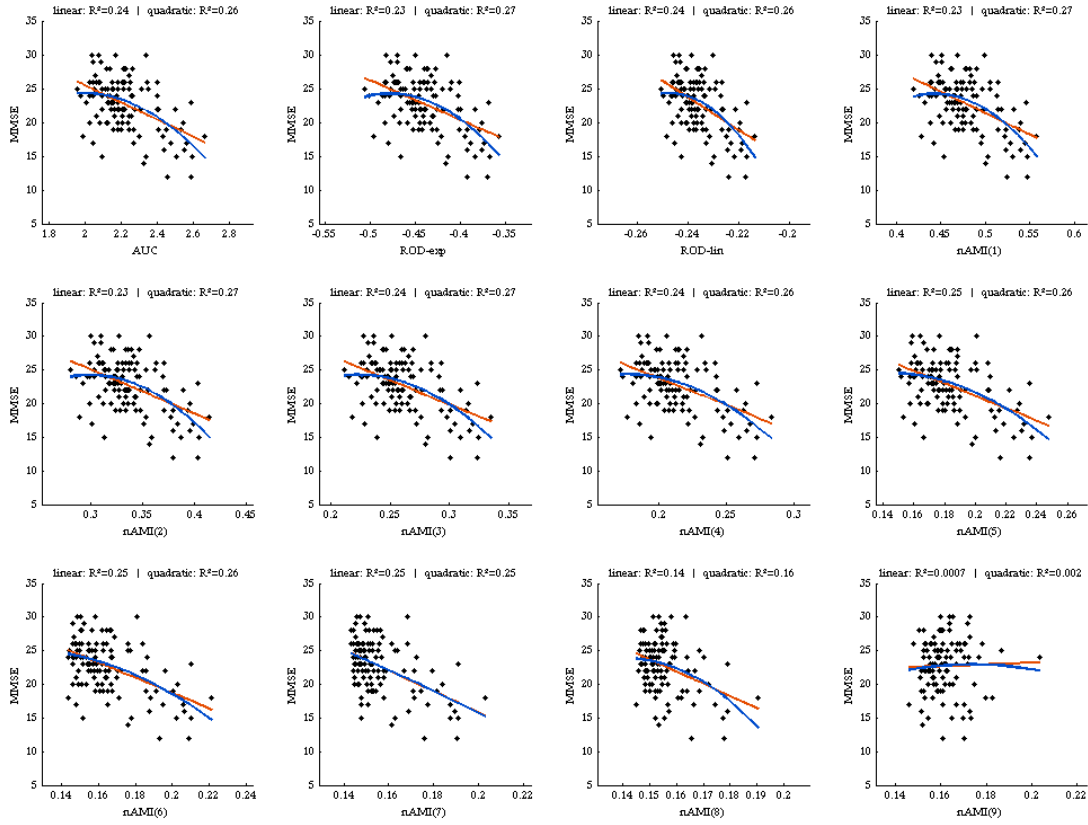


Figure 17: Scatter plots for AMI biomarkers vs MMSE and the linear and quadratic prediction model obtained by linear regression for a single channel (P7).

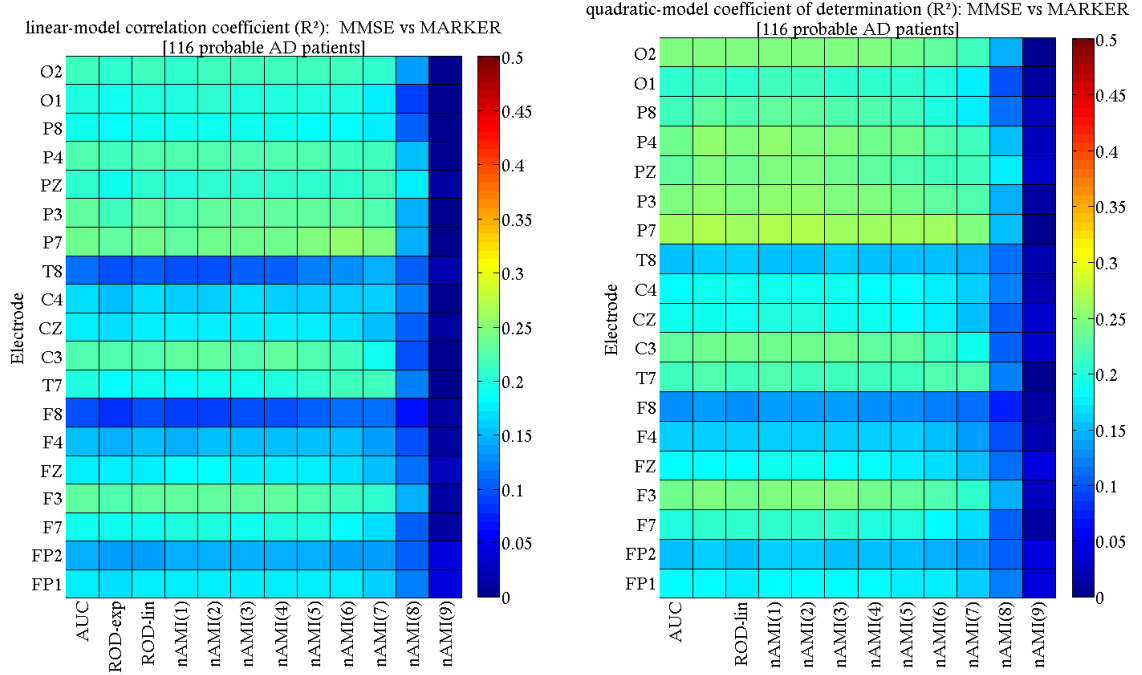


Figure 18: Predictive strength of AMI biomarkers measured by the coefficient of determination ( $R^2$ ): linear (left) and quadratic (right) fit to model a possible correlation between MMSE and biomarkers for all channels.

an abrupt drop in  $R^2$  for time-shifts higher than 7. The reason for this lies in the fact that the AMI shows significant increases for time-shifts that correspond to dominant frequencies within a patient’s epoch (discussed in detail in the in-depth analysis below 9.3).

To conclude, the promising results from (Garn et al., 2014) who applied the AMI to a subset of this database, could not be confirmed since the AMI, although revealing certain trends, is not specific enough to produce biomarkers with diagnostic value.

### 9.3 IN-DEPTH ANALYSIS

**PARAMETER SELECTION** Fig 19 shows 2 examples for the type of AMI functions from individual EEG epochs. The only tunable parameter in AMI calculation itself is the number of bins ( $N_{bin}$ ) which are used to model the probability distribution for occurrence of voltage amplitudes during an epoch. The functions are normalized since we are interested to compare the relative decrease of mutual information between patients and because the initial value of 0 time-delay can be additionally extracted and reflects the Shannon Entropy (ShEn) of the signal. One can observe how different  $N_{bin}$  values result in different level of detail probability distributions (small figures on right). Each value of the AMI function (left) is calculated directly from this 2d probability distributions. One can observe, that

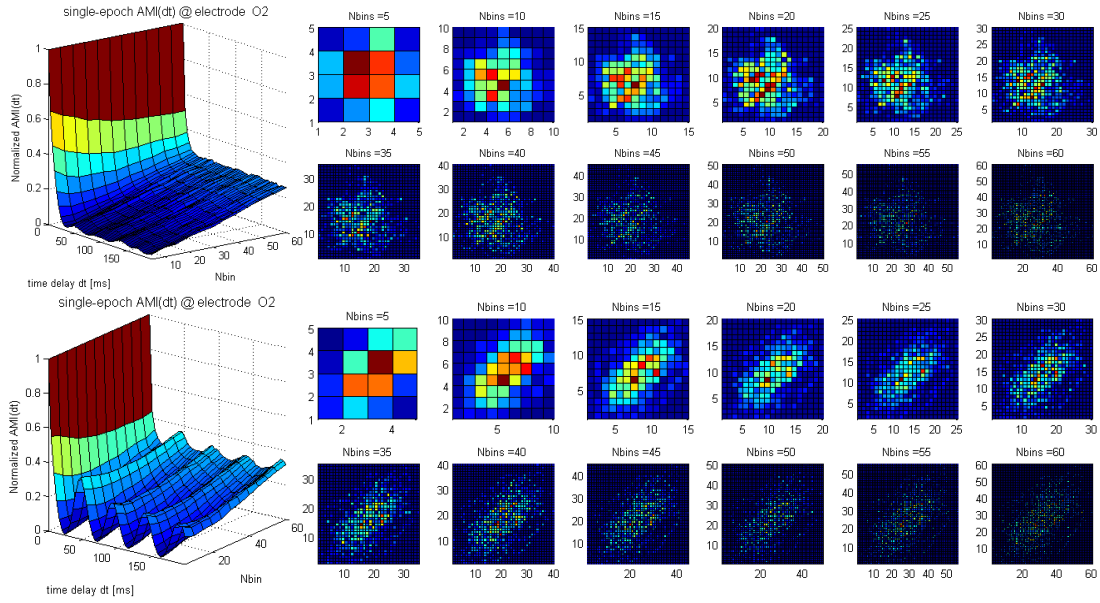


Figure 19: **AMI** curves for different  $N_{bin}$  calculated for a sample epoch from 2 randomly chosen patients of the PRODEM database. The figures on the right are the different joint-probability distributions that belong to the **AMI** values of  $dt=192ms$  (=maximum calculated time-delay from the left figure).

with increasing  $N_{bin}$ , the **MI** doesn't decrease as much for high  $N_{bin}$ , (a high *baseline AMI* for high  $N_{bin}$ ), but in turn seems to be more robust to oscillatory behavior. One can observe that for  $N_{bin}$  higher than around 40, too many bins are used for the discretization of the probability distribution which leads to a **AMI** function that decreases slow and therefore sensitivity to detect differences between patients diminishes. On the other hand,  $N_{bin}$  of 10 and less seem to oversimplify the real probability distribution and are therefore more sensitive to fluctuations. Consistent results were obtained for  $10 < N_{bin} < 35$ . The main results from the previous section were obtained by using  $N_{bin} = 30$ .

**AMI AND IAF** By scanning over all averaged **AMI** profiles, one can observe that every patient exhibits bumps at certain frequencies at different time-shifts but more importantly different magnitude. Those bumps could be considered as random fluctuations of the **AMI** measurement. However, for most of the patients, prominent bumps survived even after averaging the **AMI** function of each channels over all epochs. There were still noticeable differences between each channel though. A further analysis of the most prominent, first bump in the **AMI** curve revealed that this bump occurs at half of the period that corresponds to the Individual Alpha Frequency (**IAF**) from the PRODEM database. The PRODEM **IAF** was determined *independently*, that is from measurements of the occipital area and is not available for all patients. Interestingly there is a very high accordance between the **IAF** from the database and the peak position extracted from the **AMI**

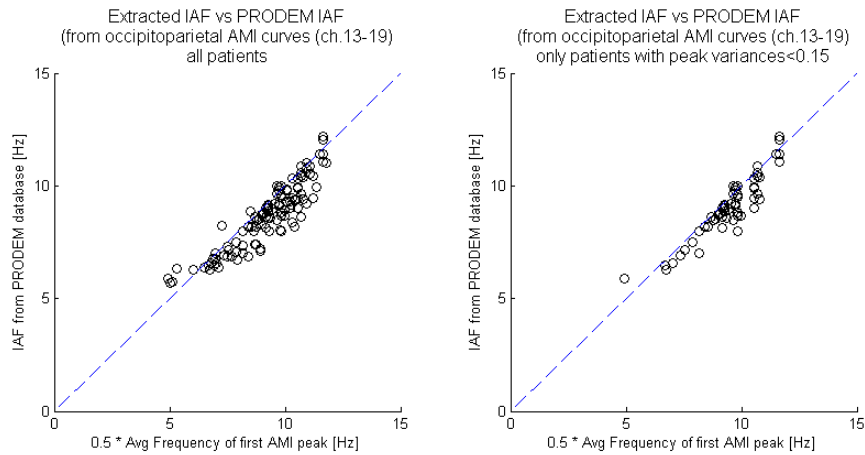


Figure 20: Scatter Plot of the AMI-bump frequency and the IAF from the PRODEM database. Notice that due to inter-channel-variability, the position of the first peak was determined by averaging over Channels 13-19. The right figure shows only data from patients with low inter-channel variability.

profiles of each patient, averaged over all electrodes of the occipito-parietal area (see Fig 20).

To explain this, we have to ask ourselves how an increase of AMI at a certain time-interval can actually arise. Let's assume the spectrum of a patient's EEG exhibits a prominent peak at some frequency. If we shift our signal by exactly the time that corresponds to the period of this frequency, we should obtain a higher similarity to the original signal than for other time shifts (before and after) since certain pairs of amplitudes occur more often and thus have a much higher estimated joint probability. Now the same holds true for half a period of this frequency, since (even though amplitudes are reversed), there is an increase of coincidences of certain amplitude pairs as well. As a consequence, if an EEG epoch is driven by a dominant frequency, the corresponding AMI curve will show bumps at all time-shifts that correspond to multiples of the halved frequency. The height of those bumps will be proportional to the amplitude in the power spectrum. Secondly, since every EEG exhibits more bumps than the peaks that arise due to IAF, other dominant frequencies will also lead to increases in AMI and produce bumps at multiples of its half-period. That is, the AMI contains information about not only one but all dominant frequencies within the EEG. Going back to our discovery, it now also becomes clear why the first peak is the IAF and no higher frequency, since all EEG epochs were bandpass filtered ( $2 < f < 15\text{Hz}$ ) and their frequency power spectrum restricts to around 2 to 15Hz (no hard-cut-off edges, so there remains some spectral content slightly below 2Hz and slightly above 15Hz). Therefore the frequency-spectrum related influence on the AMI curves is restricted to time shifts bigger than a critical time-shift that equals 2 times the maximum EEG frequency. In our case, the 30Hz limit corresponds to a critical time-shift of 33ms i. e. a timeshift of approx. 8 datapoints.

Interestingly, when comparing this with Fig 18, one can observe an abrupt drop of correlation values at this predicted time-shift.

**WHAT DO WE ACTUALLY MEASURE** If we take up the idea of the **AMI** being sensitive to physiological content such as the **IAF** at bigger timeshifts, we arrive at a point where we have to ask ourselves which physiological content is actually influencing the complexity measures we apply. The decrease of the **AMI** function reflects the unpredictability of a signal, but the measure itself is based on estimating correlations between a signal and the time-shifted version of itself. It is quite obvious that the **AMI** drops after shifting it by a few datapoints, but taking only a few datapoints into account means that we have actually analyzed the complexity on a very small timescale! A shift by 1 datapoint corresponds to 4ms (256Hz), 8 datapoints to 32ms (31Hz). Thus the **AMI** rate of decrease, but also measures like **SaEn** or **ShEn** look at the regularity on timescales that reflect diffuse and tiny changes in voltage activity, with frequencies above 60Hz. Activity at such small timescales probably reflects to some extent, the dynamics of many, many *microsources* switching on and off and thus enabling such fast changes. If this was a fact, then - according to the trend - **AD** patients tend to have a slower dynamics in the microsource activation level. However, taking into concern that noise plays a very big factor at such small timescales, it doesn't come as a surprise that no complexity-based biomarkers have made any impact up till now. Concerning this study, given the fact that all **EEG** signals were bandpass filtered ( $2 < f < 15\text{Hz}$ ), the observed correlations occur at those time-scales that we actually excluded due to several reasons in the first place.

The results of this investigation as well as previous complexity based **EEG** biomarker studies found correlations by looking at irregularities on a timescale that is of the same order as noise and muscular artifacts. Hence, even though this study used state-of-the-art signal preprocessing techniques, the end result makes it hard to hide the fact that these preprocessing techniques have been basically useless in this work and that other research groups were not aware of this fact.



## CONCLUSIONS

---

As we have seen, statistical correlations between the proposed biomarkers and the severity of AD represented by the patient's MMSE scores are not strong enough to provide evidence for a possible diagnostic value and even a basis for further investigations of this type. Considering the full EEG record as fingerprints of an underlying system, whose functional state can be classified according to the unpredictability or complexity of the waveforms, is just a very oversimplified model. In addition, most research groups analyzed the complexity of EEG signals at very small timescales (even as small as the sampling frequency). The increased level of noise and the lack of a profound underlying physiological model question the reasoning behind application of complexity-based markers at such scales. However, looking at complexity at bigger timescales where noise becomes less of a factor, the tight relationship to each patient's frequency architecture comes into play, and thus strong individual differences make it hard to obtain robust biomarkers. Our results confirmed that individual inter-electrode and inter-epoch deviations in the waveform complexity are of the same magnitude as the statistical deviations between different patients. This fact, was hardly reported in the cited literature, where the hunt for high correlation coefficients or low p-values seemed to enjoy a much higher reputation than the correct evaluation and interpretation of the methods being used.

Another crucial point is the experiment design. Event-related potentials, which are changes in cortical electrical activity as a response to a *task*, can provide much more specific and less noisy data than a simple continuous observation of the resting phase activity. The main reason is that an event related signals can be well identified from background and automatically separated from all other sources. Thus one can get rid of the noise across *all* frequencies ranges. In analogy to the Default Mode Network which emerged from fMRI studies, a theoretical model for resting state EEG activity would allow a more specific analysis even for resting-phase records and would be the first step in order to increase the diagnostic value of these types of records. A first approach to a more refined resting state analysis is given in B

Still, despite the shortcomings of this and previous investigations the EEG is a very powerful tool, with its main strength to record signals on a very fast timescale. At the same time the fast and complex dynamics make it difficult to interpret the physiological meaning of this biosignal. It is also the reason why the diagnostic impact of EEG is still comparatively small. The search for EEG biomarkers in a conservative way, that is without using a profound theoretical

model, will, in my opinion, not contribute much to the development of new diagnostic possibilities concerning brain-related diseases. The approach of this work can be compared with the attempt to analyze the waveform of a patient's QRS complex (from ECG) without knowing considering the physiology of heart contraction. Of course, a purely mathematical approach (e.g. frequency-content or wave-form-irregularity analysis) might even reveal certain trends, but its use is very limited due the fact that individual differences (starting with small anatomical differences) lead to very high variances between biosignals from different patients. Only after understanding the physiology of heart contraction the ECG becomes a powerful tool since it can then be used to detect a possible malfunctioning of the heart even on an *individual* basis.

After all there is still the fact that the brain is the most complex and thus least understood organ of the human body. Rough estimates suggest that the central nervous system of higher mammals is comprised of  $10^{11}$  neurons communicating with each other via  $10^{15}$  synapses. It doesn't come to a surprise that the functional aspects of global electric activity are far from being understood. Concerning the future of biomarkers for precise diagnosis of Alzheimer's Disease, one should probably start with the question about how large-scale network properties are changed by small-scale network changes (similar to [Palop and Mucke \(2010\)](#)). Just as the disease itself is defined on a chemical or cellular level, such a theory would provide a profound basis and result in a theoretical framework, which might be the foundation for future research and hence impact future generations.



## Part IV

### APPENDIX

The Appendix, in addition to the bibliography, contains two additional chapters. The first introduces the concept of the [CMI](#) method, which extends the concept of [AMI](#) and thus can be used in a variety of ways, but is strictly speaking not a method that investigates signal complexity. The second chapter presents a method to visualize the impact of synchronous events, termed *bursts*, which occur quite often during the resting phase records of patients, and thus this approach can be seen as a first step towards a more refined resting phase [EEG](#) analysis.



The **CMI** method, as described in the methods section, provides an interesting measure to exploit correlations between different electrodes within the **EEG** record. It provides numerous possibilities and thus deserves to be mentioned in this context, although strictly speaking cannot be seen as a complexity measure. However, during my work I also investigated the methodology of the **CMI**, and since not mentioned elsewhere, I want to briefly comment on several things that became clear and thus allow the interested reader to use this measure for future research.

#### A.1 AN INTUITIVE PICTURE OF CMI

Figure 21 shows examples for the type of **CMI** functions that arise from individual **EEG** epoch analysis. The only tunable parameter in **CMI** calculation itself is, exactly like in the **AMI** Analysis, the number of bins ( $N_{bin}$ ). However, instead of time-shifting the very same epoch of a signal, the epochs of other electrodes are time-shifted. Therefore, instead of capturing unpredictability or complexity, the change of mutual information over time between *different* epochs describes the temporal evolution of the nonlinear and linear correlations between the measurements from two different electrode locations. Again, the oscillatory waveform of each signal results from the summed electrical activity of numerous cortical regions beneath the skull that are active during the recorded time. Hence it makes sense that records from nearby electrodes contain contributions from the same sources and hence share **MI**. The reason for this is the volume conduction (see Fig 3). Therefore each **CMI** function intrinsically contains spatial information about cortical activity.

The **CMI** function of an electrode pair is obtained in the following way: when comparing the signals of 2 different electrodes, we can either shift the signal of the one or the other electrode in time. This is the equivalent of taking one of the two electrodes as a reference and shifting the signal of other forward and backward in time. Both functions (denoted by **CMI-in**, **CMI-out**) can therefore be synthesized into a single **CMI** function (full **CMI**). It is noteworthy that only few groups have used **CMI** as a discriminative measure in their analysis concerning **AD** detection (e.g. Jeong et al. (2001) and Wan et al. (2008)). Those works did not report about the properties of **CMI** functions. This chapter therefore presents novel features that can be extracted from the analysis of nonlinear correlations between electrode pairs using **CMI**.

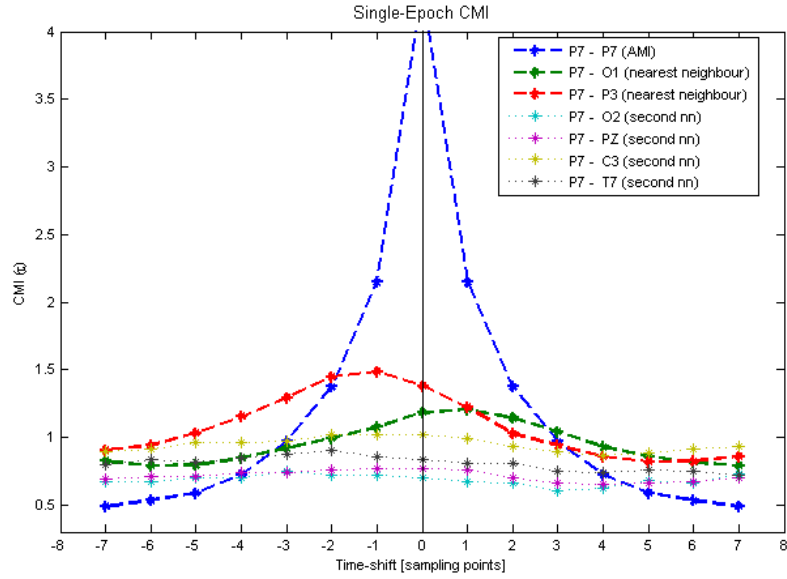


Figure 21: CMI function of a single epoch for various channel pairs

## A.2 MODELLING THE CMI FUNCTION

Figures 21 and 22 show examples for the behavior of CMI-in, CMI-out functions and their synthesized full CMI functions extracted from various neighboring channels. We can see in the examples that there is a trend in the CMI functions depending on the distance to the reference electrode.

1. Nearby channels, have very similar patterns with the reference channel, such that the MI shows a strong decrease for short time-delays, similar to a weakened AMI function. The high similarity in patterns arises due to the fact that both electrodes partly contain electrical activity from the same cortical sources and therefore share an increased MI compared to MI values at bigger time-shifts.
2. Channels that are further away do not show such a behavior in the CMI functions. There is no observable decrease for short time-delays, since the function stays approximately constant for all time-shifts.

Type 1 CMI signals contain an interesting feature since they show a peak in MI which is not symmetrically centered but rather displaced in time. When scanning through various single epoch CMI-function, one can observe that this displacement (which lies in the range of 0-15ms) changes from epoch to epoch. This can be understood through consideration of different cortical sources being active at sampling points. The time-delay also reveals a directionality of information flow, since it can be either positive or negative, and when comparing the delays between all electrodes, they could be mapped as a 2d or 3d gradient field. This map would then reveal the time-resolved spatial information about the shared

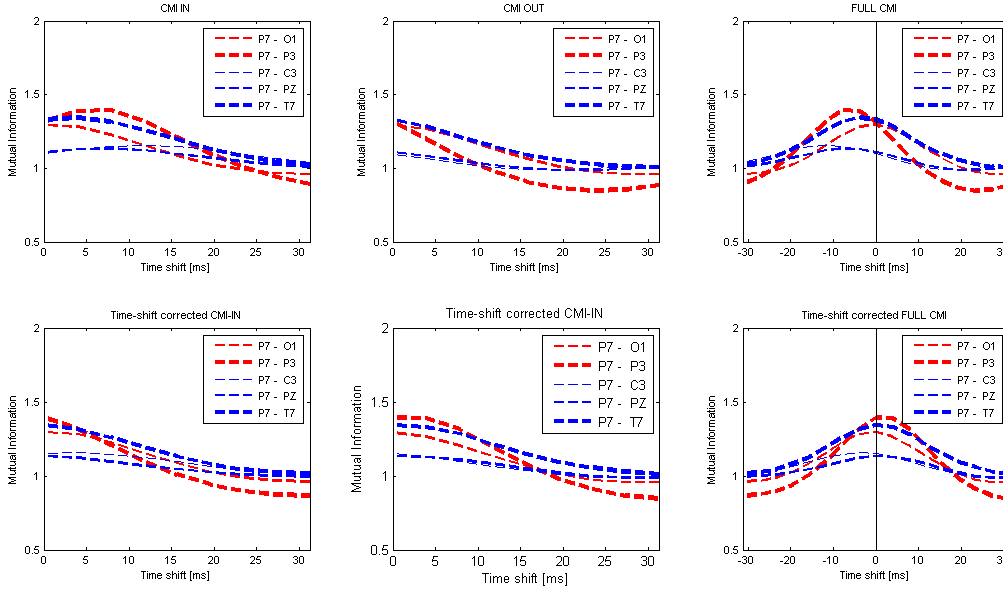


Figure 22: Illustration of CMI-in, CMI-out, full-CMI for the original signals (top) and time-shifted versions (bottom). Note that the time-shift correction is limited by the time-resolution (in our case 4ms) defined by the inverse sampling frequency.

sources of the EEG signals.

Type 2 CMI signals can give an estimation for the spread of volume conduction, thus if there is no peak in the CMI function one can assume that the signals to share no common cortical sources. Thus the electrodes of type 2 actually define the spatial extend of cortical sources and thus, in combination with type 1 signals give information about the location, time and spread of cortical sources.

This type of analysis, where the similarity between the signals of different electrodes can be tracked over time, composes a first step to reconstruct those cortical sources which produce the main contributions to the global measured electrical potential. A variety of spatial source reconstruction models already exist (Ramirez, 2008), and these methods also combine the spatial information with the time resolution to maximally exploit the level of information inherent in the full EEG record. This type of approach however goes beyond the framework of this thesis and thus is not further exploited and in the following section I will restrict myself to quantitative description of the CMI function.

### A.3 CMI BIOMARKERS

The following list is a suggestion of features that can be extracted from the CMI function and could be used as a biomarker.

## BIOMARKERS

1. BL Estimates the baseline CMI and therefore the average MI between 2 channels
2. tshift The estimated time-shift of the CMI curve [measured in sampling point units] - lies between -2,1, 0,1,2 due to limited time-resolution (4ms).
3. AUC area under the CMI-curve of increased MI in the critical range - calculated via Trapez-formula (as an approximation to the integral)
4. RODout – exp The estimated exponential decay factor to model the rate at which the CMI between the reference channel (fixed) and all other electrodes (time-shifted) decreases.
5. RODin – exp The estimated exponential decay factor to model the rate at which the CMI between all electrodes (fixed) and the reference channel (time-shifted) decreases.
6. cCMI(t) After extracting the time-shift, each CMI curve can be repositioned by (0-2) datapoints according to the time-shift value. The values for all different time-shifts can be used as markers as-well. Each time-shift corresponds to a certain frequency.

## A.4 CMI IN AD

For the sake of completeness, Fig 23 and 24 show the results obtained by using the above biomarkers to predict the MMSE score of AD patients for electrode P7 as a reference electrode.

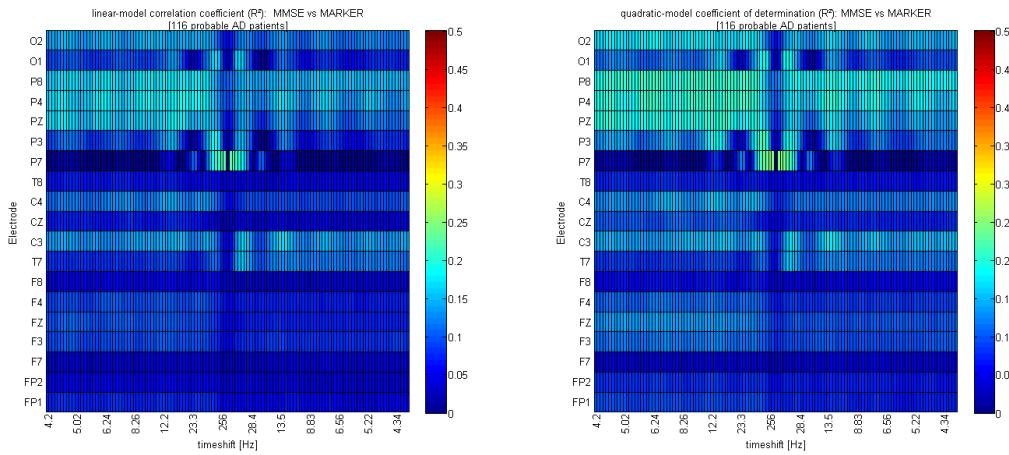


Figure 23: The predictive strength of corrected CMI function values measured by the coefficient of determination  $R^2$  for linear (left) and quadratic (right) regression models. This Figure represents the biomarkers obtained from the CMI function between channel P7 and all other channels for various time shifts, where each time shift can be translated into a frequency shift (as depicted by the xlabel unit). The row containing the P7 values corresponds to the AMI of P7.

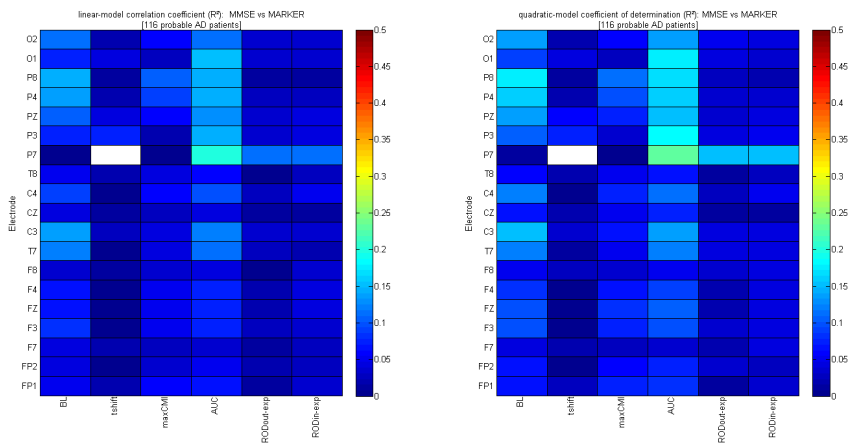


Figure 24: The predictive strength of CMI biomarker values (see list above) measured by the coefficient of determination  $R^2$  for linear (left) and quadratic (right) regression models. This Figure represents the biomarkers obtained from the CMI function between channel P7 and all other channels. The row containing the P7 values corresponds to the AMI of P7 for which no tshift can be calculated (hence the white spot)





## BURST ANALYSIS

---

### B.1 MOTIVATION

The main results of the complexity analysis made clear that the high variance between epochs within a single subject is the main reason for the shortcomings of the used methods. In addition, we saw that the occurrence of IAF in resting state EEG records of patients played a very influential role on the values of the AMI at higher time-shifts, which reflects a measures of signal (ir)regularity. Thus patients that show an increased number of epochs with dominant frequencies over several channels, (e. g. strong alpha activity), might be biased towards more regular signals than patients without. Therefore the idea behind this burst analysis was to extract and identify those epochs which showed increased synchronous, regular activity. This way one can refine the complexity based analysis by separating epochs into two different types: burst epochs, and non-burst epochs. This separation could eliminate the direct influence of a patients tendency for high synchronous activity and thus reduce the inter-subject variability since occurrence of synchronous activity is known to be highly depended on mental state of the subject (such as  $\theta$  and  $\delta$  activity changes during sleep stages, or  $\alpha$  activity modulation through directed attention)

### B.2 IMPLEMENTATION

To detect periods of high cortical synchronization, the following algorithm was implemented:

1. Calculate the spectrum of each epoch (4s epoch length; 3s overlap between epochs) and determine the biggest peaks in the spectra
2. Find bursts by extracting periods of repetitive occurrence of a peak at the same frequency in at least 5 successive epochs. Between each epoch the peak position can vary by max. 1 datapoint (0.25Hz; which is the resolution of the calculated spectra and is restricted by the sampling frequency. For each channel, create a list of start-time, duration, and peak frequency of all detected burst, called burst-matrix. After having calculated the spectrum of all epochs, this matrix is checked for periods of successive occurrence of peaks at the same frequency position. That is, if the spectrum has a peak during at least 5 epochs, and the peak

position does change by a maximum of 0.25 Hz, then the indices of those epochs, as well as the value of the occurring peaks are stored in the 'burst-matrix'.

### B.3 APPLICATION

To show an example for the implemented algorithm, Fig 25 shows how the burst matrix can be used to determine dominant individual frequency bands from the resting state EEG records of different patients. By comparing epoch peaks of all channels, one can determine a patients individual frequency bands by investigation of the burst density profile which is obtained by summing over all epochs as shown on the right in Fig 25. Scanning over all patients' profiles, one can see that different patients can exhibit very different burst behavior, and some patients do not contain clear burst periods in their record, whereas others do. This leads to the conclusion that the physiological content of EEG resting state records strongly varies between different patients, probably due to its high dependence on the mental state during this resting period. Thus any investigation that aims to analyze and compare resting state records between patients could probably achieve more consistent results by differentiating between epochs of different type.

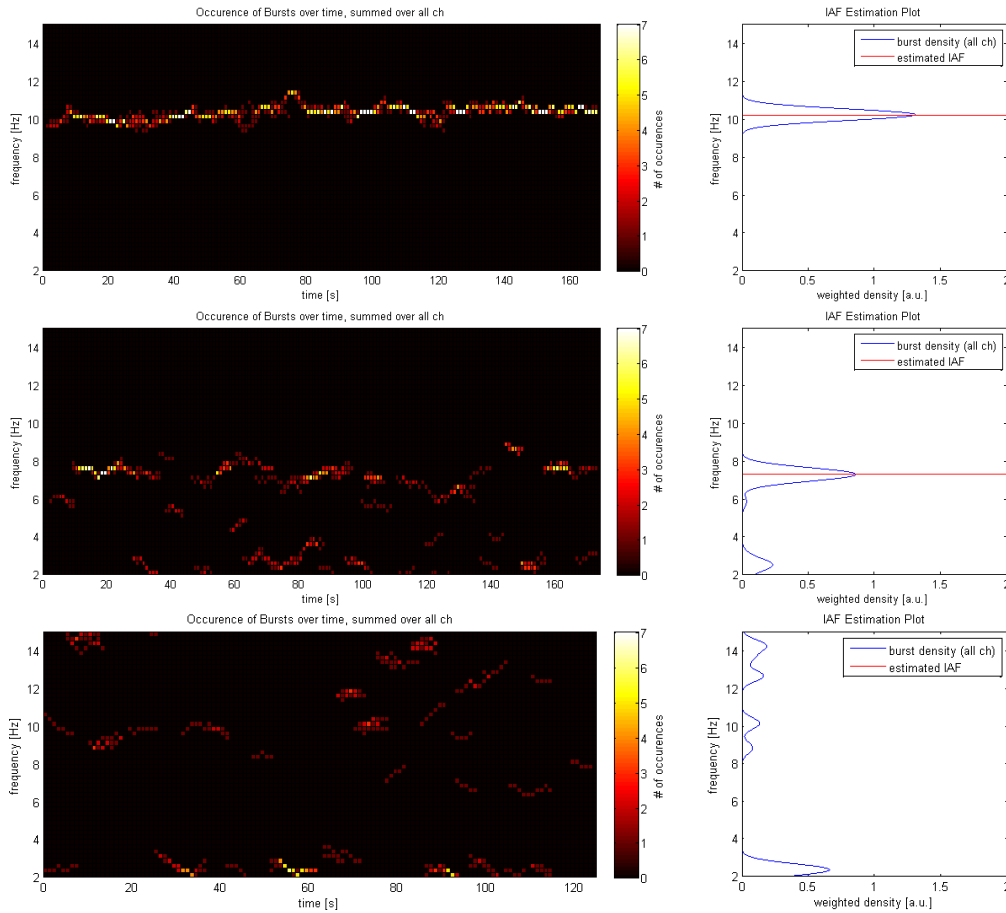


Figure 25: Visualization of the burst matrix (left) and determination of the IAF via the summation over all epochs. The plots reflect the burst matrices for 3 different patients (top-middle-bottom). The color code reflects the number of channels with the same peak frequency, thus brighter values indicate high synchronous activity, while dark spots are regionally restricted bursts. One can see that the profiles differ strongly, with a clear indication of a single frequency band (top), two distinct frequency bands (middle) and occurrence of multiple dominant frequencies (bottom) where no IAF estimation was possible.



## BIBLIOGRAPHY

---

- Abásolo, D., Escudero, J., Hornero, R., Gómez, C., and Espino, P. (2008). Approximate entropy and auto mutual information analysis of the electroencephalogram in alzheimer's disease patients. *Medical & biological engineering & computing*, 46(10):1019–1028. (Cited on page 45.)
- Acar, Z. A. and Makeig, S. (2010). Neuroelectromagnetic forward head modeling toolbox. *Journal of neuroscience methods*, 190(2):258–270. (Cited on page 9.)
- Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford University Press. (Cited on page 6.)
- Costa, M., Goldberger, A., and Peng, C.-K. (2005). Multiscale entropy analysis of biological signals. *Phys. Rev. E*, 71(2). (Cited on pages v, vii, 29, and 30.)
- Dauwels, J., Srinivasan, K., Ramasubba Reddy, M., Musha, T., Vialatte, F.-B., Latchoumane, C., Jeong, J., and Cichocki, A. (2011). Slowing and loss of complexity in alzheimer's eeg: two sides of the same coin? *International journal of Alzheimer's disease*, 2011. (Cited on page 41.)
- Dauwels, J., Vialatte, F., and Cichocki, A. (2010). Diagnosis of alzheimer's disease from eeg signals: where are we standing? *Current Alzheimer Research*, 7(6):487–505. (Cited on pages v, vii, and 13.)
- Delbeuck, X., Van der Linden, M., and Collette, F. (2003). Alzheimer's disease as a disconnection syndrome? *Neuropsychology review*, 13(2):79–92. (Cited on page 13.)
- Escudero, J., Abásolo, D., Hornero, R., Espino, P., and López, M. (2006). Analysis of electroencephalograms in alzheimer's disease patients with multiscale entropy. *Physiological measurement*, 27(11):1091. (Cited on page 35.)
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3):189–198. (Cited on pages v, vii, 13, and 20.)
- Garn, H., Waser, M., Deistler, M., Benke, T., Dal-Bianco, P., Ransmayr, G., Schmidt, H., Sanin, G., Santer, P., Caravias, G., et al. (2014). Electroencephalographic complexity markers explain neuropsychological test scores in alzheimer's disease. In *Biomedical and Health Informatics (BHI), 2014 IEEE-EMBS International Conference on*, pages 496–499. IEEE. (Cited on pages 13, 14, 45, 46, and 48.)

- Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and clinical neurophysiology*, 10:371–375. (Cited on page 17.)
- Jeong, J. (2004). Eeg dynamics in patients with alzheimer’s disease. *Clinical neurophysiology*, 115(7):1490–1505. (Cited on pages v, vii, and 13.)
- Jeong, J., Gore, J. C., and Peterson, B. S. (2001). Mutual information analysis of the eeg in patients with alzheimer’s disease. *Clinical Neurophysiology*, 112(5):827–835. (Cited on pages 25, 45, and 57.)
- Labate, D., La Foresta, F., Morabito, G., Palamara, I., and Morabito, F. C. (2013). Entropic measures of eeg complexity in alzheimer’s disease through a multivariate multiscale approach. *Sensors Journal, IEEE*, 13(9):3284–3292. (Cited on page 35.)
- McBride, J. C., Zhao, X., Munro, N. B., Smith, C. D., Jicha, G. A., Hively, L., Broster, L. S., Schmitt, F. A., Kryscio, R. J., and Jiang, Y. (2014). Spectral and complexity analysis of scalp eeg characteristics for mild cognitive impairment and early alzheimer’s disease. *Computer methods and programs in biomedicine*, 114(2):153–163. (Cited on page 13.)
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., et al. (2011). The diagnosis of dementia due to alzheimer’s disease: Recommendations from the national institute on aging-alzheimer’s association workgroups on diagnostic guidelines for alzheimer’s disease. *Alzheimer’s & Dementia*, 7(3):263–269. (Cited on pages 12 and 20.)
- Mizuno, T., Takahashi, T., Cho, R. Y., Kikuchi, M., Murata, T., Takahashi, K., and Wada, Y. (2010). Assessment of eeg dynamical complexity in alzheimer’s disease using multiscale entropy. *Clinical Neurophysiology*, 121(9):1438–1446. (Cited on page 35.)
- Palop, J. J. and Mucke, L. (2010). Amyloid-[beta]-induced neuronal dysfunction in alzheimer’s disease: from synapses toward neural networks. *Nature neuroscience*, 13(7):812–818. (Cited on pages 11 and 54.)
- Park, J.-H., Kim, S., Kim, C.-H., Cichocki, A., and Kim, K. (2007). Multiscale entropy analysis of eeg from patients under different pathological conditions. *Fractals*, 15(04):399–404. (Cited on pages 35 and 36.)
- Ramirez, R. R. (2008). Source localization. 3(11):1733. (Cited on page 59.)
- Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656. (Cited on pages v, vii, and 22.)

- Stam, C. J. (2005). Nonlinear dynamical analysis of eeg and meg: review of an emerging field. *Clinical Neurophysiology*, 116(10):2266–2301. (Cited on pages v, vii, and 21.)
- Takahashi, T. (2013). Complexity of spontaneous brain activity in mental disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 45:258–266. (Cited on page 13.)
- Tombaugh, T. N. and McIntyre, N. J. (1992). The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society*. (Cited on page 14.)
- Wan, B., Ming, D., Qi, H., Xue, Z., Yin, Y., Zhou, Z., and Cheng, L. (2008). Linear and nonlinear quantitative eeg analysis. *Engineering in Medicine and Biology Magazine, IEEE*, 27(5):58–63. (Cited on page 57.)
- Waser, M. and Garn, H. (2013). Removing cardiac interference from the electroencephalogram using a modified pan-tompkins algorithm and linear regression. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 2028–2031. IEEE. (Cited on page 19.)
- Wu, S.-D., Wu, C.-W., Lee, K.-Y., and Lin, S.-G. (2013). Modified multiscale entropy for short-term time series analysis. *Physica A: Statistical Mechanics and its Applications*, 392(23):5865–5873. (Cited on page 30.)
- Yang, A. C., Wang, S.-J., Lai, K.-L., Tsai, C.-F., Yang, C.-H., Hwang, J.-P., Lo, M.-T., Huang, N. E., Peng, C.-K., and Fuh, J.-L. (2013). Cognitive and neuropsychiatric correlates of eeg dynamic complexity in patients with alzheimer’s disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 47:52–61. (Cited on page 35.)