Development of Differential Connectivity Graph for Characterization of Brain Regions Involved in Epilepsy

Ladan Amini

Development of Differential Connectivity Graph for Characterization of Brain Regions Involved in Epilepsy

Ladan Amini

Submitted in partial fulfillment of the requirements
for the joint degree of Doctor of Philosophy in
Signal Processing and Telecommunications

in the

University of Grenoble, Grenoble, France
Grenoble Images Signal Parole et Automatique Laboratory (GIPSA-lab)

and

University of Tehran, Tehran, Iran
College of Engineering, School of Electrical and Computer Engineering
Control and Intelligent Processing Center of Excellence (CIPCE)

under the joint supervision of

Prof. Christian Jutten and **Prof. Hamid Soltanian-Zadeh**

and co-supervision of

Dr. Sophie Achard and **Dr. Gholam Ali Hossein-Zadeh**

December 21, 2010

UNIVERSITE DE GRENOBLE
INSTITUT POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque:

THÈSE EN COTUTELLE INTERNATIONALE

pour obtenir le grade de

DOCTEUR DE L'Université de Grenoble

délivré par l'Institut polytechnique de Grenoble

et

de L'Université de Téhéran

Spécialité : Signal, Image, Parole, Télécoms

préparée aux laboratoires **Grenoble Images Signal Parole et Automatique (GIPSA-lab)**

dans le cadre de **l'École Doctorale**

Électronique, Électrotechnique, Automatique et Traitement du Signal

et au laboratoire **Control and Intelligent Processing Center of Excellence (CIPCE)**

présentée et soutenue publiquement le 21 decembre 2010 par **Ladan AMINI**

Développement de Graphe de Connectivité Différentiel pour Caractérisation des Régions Cérébrales Impliquées dans l'Épilepsie

DIRECTEURS DE THÈSE : Christian JUTTEN et Hamid SOLTANIAN-ZADEH

ET CO-ENCADRANTS DE THÈSE : Sophie ACHARD et Gholam Ali HOSSEIN-ZADEH

JURY

M. Jean-Philippe LACHAUX, Président
M. Fabrice WENDLING, Rapporteur
M. Dimitri VAN DE VILLE, Rapporteur
M. Christian JUTTEN, Co-directeur de thèse
M. Hamid SOLTANIAN-ZADEH, Co-directeur de thèse
Mme. Sophie ACHARD, Co-encadrrant
M. Philippe KAHANE, Examinateur
M. Mohammad Bagher SHAMSOLLAHI, Examinateur
M. Gholam Ali HOSSEIN-ZADEH, Invité
M. Laurent VERCUEIL, Invité

Dedicated to Mohammad and Ali,

to my parents, my sister, and my brother

Acknowledgments

I would like to thank Prof. Hamid SOLTANIAN-ZADEH for his serious and kind supports and helps. I owe to him for learning the alphabet of research. Hamid kindly provided me the possibility to gain teaching and research experiences. He greatly supported me to start our collaboration with Prof. Christian JUTTEN.

I would like to thank Christian for his heartfelt and honest helps and supports in this thesis. He spent lots of time for solving the methodological problems of the thesis, and for our collaboration with hospital in order to obtain the data. I never felt any hierarchy while I was working with Christian. Without helps and supports of Christian finishing this thesis was impossible. I owe to Christian for learning how to do the research first morally then technically.

I would like to thank Dr. Sophie ACHARD. Sophie and I worked together like two friends. I enjoyed a lot working with Sophie. She spent a great time on this thesis and teaching me the basics of connectivity analysis. She was always very optimistic for solving the problems.

I would like to sincerely appreciate the exceptional help and supports of Prof. Mahmood KAMAREI for our co-tutelle collaboration between University of Tehran and University of Grenoble. Without his honest helps, my collaboration with Christian was not possible.

I would like to thank the reviewers of this thesis, Prof. Fabrice WENDLING and Prof. Dimitri VAN DE VILLE for the evaluation of this work and their fruitful comments. I would like to thank the president of the jury members, Dr. Jean-Philippe LACHAUX.

I would like to thank Prof. Philippe KAHANE and his colleague, Dr. Lorella MINOTTI for providing us the intracerebral EEG recordings and the related useful information about this data. Indeed the data has an essential importance in the thesis. The clinical key points of Philippe has opened many insights in our work.

I would like to thank Dr. Laurent VERCEUIL. His sympathy and kindness was always very encouraging. Laurent has done the crucial step of our method for labeling the data. His helpful clinical interpretations was always leading for our work.

I would like to thank Dr. Gholam Ali HOSSEIN-ZADEH for his kind, and friendly helps.

I would like to thank Dr. Mohammad Bagher SHAMSOLLAHI for his invaluable and helpful comments on this work.

I would like to thank Dr. Olivier DAVID for his great collaboration about the data, efficient discussions, and providing me other related facilities which without these helps this thesis could not be proceeded.

I would like to thank Mrs Hengameh DASTMALCHI for all her efficient and very friendly helps and supports during my study in University of Tehran. In the grace of her kind helps I could do my research work without any anxiety about the official works.

I would like to thank Jean-Marc SACHE, Lucia BOUFFARD-TOCAT, Isabelle CIEREN, and Nadine CHRISTIN for their exceptional helps during my thesis.

I would like to thank my husband, Mohammad for his great helps. Our discussions were very leading in this thesis. I owe to my son, Ali for his honest helps. Ali had always some solutions for my problems which was making me very happy. Always their love was encouraging for me during this thesis. I would also like to thank my parents, my sister, my brother, my niece, and my nephew for all their encouragements, and honest supports.

I would like to thank my friends Reza, Cédric, Léonardo, Zahra, Agatha, Samareh, Mohammad, Ronald, Rosie, Sandra, Damien, and Guanghan for their invaluable helps.

Résumé

Les patients pharmaco-résistants sont des candidats pour la chirurgie de l'épilepsie. Le but de cette chirurgie est d'enlever les zones à l'origine de la crise (SOZ) sans créer de nouveaux déficits neurologiques. Pour localiser les SOZs, une des meilleures approches consiste à analyser des électroencéphalogrammes intracérébraux (iEEG). Toutefois, l'enregistrement des crises, qui sont des événements rares et critiques, est compliqué contrairement à l'enregistrement de décharges épileptiques intercritiques (IED), qui sont généralement très fréquentes et anodines. La prévision des SOZs, par estimation des régions à l'origine des IEDs, est donc une alternative très intéressante, et la question de savoir si l'estimation des régions IED peut être utile pour prédire les SOZs, a été au coeur de plusieurs études. Malgré des résultats intéressants, la question reste ouverte, notamment en raison du manque de fiabilité des résultats fournis par ces méthodes.

L'objectif de cette thèse est de proposer une méthode robuste d'estimation des régions à l'origine des IEDs (notées ℓIED) par analyse d'enregistrements intracérébraux iEEG. Le point essentiel de cette nouvelle méthode repose sur la détermination d'un *graphe de connectivité différentiel* (DCG), qui ne conserve que les noeuds (électrodes) associées aux signaux iEEG qui changent de façon significative selon la présence ou l'absence d'IEDs. En fait, on construit plusieurs DCGs, chacun étant caractéristique d'une échelle obtenue après transformée en ondelettes. La fiabilitié statistiques des DCGs est obtenue à l'aide des tests de permutation. L'étape suivante consiste à mesurer les quantités d'information émise par chaque noeud, et d'associer à chaque connexion (arête) du graphe une orientation qui indique le transfert d'information du noeud source vers le noeud cible. Pour celà, nous avons introduit une nouvelle mesure nommée *Local Information* (LI), que nous avons comparée à des mesures classiques de graphes, et qui permet de définir de façon robuste les noeuds sources pour les graphes de chaque échelle. Les ℓIEDs sont finalement estimées selon une méthode d'optimisation multi-objectifs (de type Pareto, peu utilisée dans la communauté signal-image) construite à partir des valeurs des LI des DCG dans les différentes bandes de fréquences.

La méthode proposée a été validée sur cinq patients épileptiques, qui ont subi une chirurgie d'exérèse et sont déclarés guéris. L'estimation des régions ℓIED a été comparée avec les SOZs détectées visuellement par l'épileptologue et celles détectées automatiquement par une méthode utilisant une stimulation destinée à provoquer des crises. La comparaison révèle des résultats congruents entre les SOZs et les régions ℓIED estimées. Ainsi, cette approche fournit des ℓIED qui devraient être des indications précieuses pour l'évaluation préopératoire en chirurgie de l'épilepsie.

چکیده

بیماران مبتلا به صرع کانونی و مقاوم در برابر دارو برای عمل جراحی صرع[۱] توصیه می شوند. هدف از این عمل جراحی حذف مناطق شروع تشنج (SOZ[۲]) بدون ایجاد اختلال عصبی جدید است. بهترین راه شناسایی SOZ ها ضبط تشنج از موج نگاری داخل مغزی (iEEG[۳]) می باشد. به هر حال ضبط تشنج دشوار است، در مقابل ضبط دشارژهای الکتریکی در فواصل زمانی بین تشنج ها (IED[٤]) ساده تر است. بنابراین پیش بینی SOZ از طریق تخمین مناطق IED بسیار ارزشمند است. مطالعات زیادی انجام شده است که آیا تخمین مناطق IED می تواند برای پیش بینی SOZ و در نهایت برای ارزیابی های قبل از جراحی مفید باشد یا خیر؟ با وجود نتایج امیدوار کننده این مطالعات، همچنان برای حل این مشکل جواب قطعی پیدا نشده و هنوز این موضوع تحت بررسی است. مشکل عمده در همه این مطالعات به قابلیت اطمینان نتایج به دست آمده بر میگردد.

هدف از این پایان نامه تخمین مناطق IED پیشرو (ℓIED[٥]) از تجزیه و تحلیل داده iEEG در فواصل زمانی IED به کمک روش گراف اتصال و مطالعه رابطه بین این مناطق وSOZ می باشد. در این پایان نامه یک روش قابل اعتماد تحلیل گراف به نام روش تجزیه و تحلیل گراف اتصال دیفرانسیلی (DCG[٦]) ارائه می کنیم. این گراف برای تجزیه و تحلیل سیگنالهای iEEG بین تشنجی به منظور تخمین مناطق ℓIED بدون استفاده از سیگنالهای تشنج پیشنهاد می شود.

در روش ارائه شده کوپلینگ های زمان–فرکانس چند متغیره بین جفت سیگنالهای کانالهای iEEG در فواصل زمانی IED و غیر IED تخمین زده می شوند. تست چندگانه ای بر اساس جایگشت[۷] روی این کوپلینگ ها اعمال می شود تا اتصالات دیفرانسیلی بین حالتهای IED و غیر IED مغز شناسایی شود. این اتصالات دیفرانسیلی، گراف تحلیل دیفرانسیلی (DCG) را می سازند. آنگاه جهت یال های DCG تخمین زده می شوند. سپس DCG جهت دار شده (dDCG[۸]) با استفاده از روش اندازه گیری گراف پیشنهادی برای حصول گره های منبع تحلیل می شود. در نهایت نواحی ℓIED با استفاده از روش بهینه سازی چند منظوره[۹] روی مقادیر اندازه گیری شده گره های گراف های dDCG در باندهای فرکانسی مختلف تخمین زده می شوند.

روش ارائه شده روی داده ضبط شده از پنج بیمار مبتلا به صرع اعمال شده است. این بیماران تحت عمل جراحی صرع قرار گرفته و پس از عمل جراحی بهبود یافته وعاری از تشنج گشته اند. مناطق ℓIED تخمین زده شده با SOZ هایی که به طور بصری توسط متخصصین اپیلپسی تشخیص داده شده و همچنین با SOZ های شناسایی شده با استفاده از داده iEEG بر مبنای تشنج القایی، مقایسه می شوند. مقایسه نتایج، هماهنگی بین مناطق ℓIED تخمین زده شده و SOZ ها را به خوبی نشان می دهد. عاری بودن بیماران از تشنج و شامل بودن مناطق ℓIED تخمین زده شده در مناطق خارج شده از مغز حین جراحی، می تواند نشان دهنده قابل اعتماد بودن روش پیشنهادی برای تخمین مناطق ℓIED در ارزیابی های قبل از جراحی باشد.

---

[۱] Resective surgery
[۲] Seizure onset zone
[۳] Intracereberal electroencephalogram
[٤] Interictal epileptiform discharges
[٥] Leading IED
[٦] Differential connectivity graph
[۷] Multiple testing methods based on permutation
[۸] Directed DCG
[۹] Multi-objective optimization methods

Abstract

Drug-resistant epileptic patients suffering from focal epilepsy are recommended for epilepsy surgery. The aim of this surgery is to remove the seizure onset zones (SOZ) without creating new neurological deficits. To identify the SOZs, the best way is to record the seizures from intracerebral electroencephalogram (iEEG) recordings. However recording seizures is complicated contrary to the recording of interictal epileptiform discharges (IED). Therefore prediction of SOZ by estimating the IED regions is very valuable. There are several studies wondering if the estimation of the IED regions can be useful to predict the SOZ and eventually for presurgery evaluations. Although the encouraging results of these studies, the problem is still an open issue. The main problem of the previous studies is the reliability of the results.

The aim of this thesis is to estimate the leading IED (ℓIED) regions from interictal analysis of iEEG recordings through connectivity graph. The main originality of the proposed method refers to a new reliable graph analysis method called *differential connectivity graph* (DCG). This graph is designed to identify the significant discriminated connections between IED and non-IED brain states. The statistical reliability of DCG is obtained by using permutation-based multiple testing. In the proposed method, multiple DCGs associated with different frequency bands are constructed. Each DCG includes both source and sink nodes involved in IED events. To identify the source nodes related to ℓIED regions, the directions of the edges of DCG are estimated and a new measure called *local information* (LI) is proposed to measure the emittance contribution of each node. To estimate the ℓIED regions from the LI values related to multiple directed DCGs, a multi-objective optimization method is used.

The proposed method is applied on five epileptic patients. These patients underwent resective surgery and they are seizure-free after the surgery. Estimated ℓIED regions are compared with SOZ detected visually by the epileptologist and SOZ detected by a method using induced ictal iEEG. The comparison reveals congruent results between estimated ℓIED regions and SOZs. Being all of the patients seizure free and inclusion of estimated ℓIED regions in the removed regions during surgery shows the reliability of estimated ℓIED regions for presurgery evaluations.

Contents

List of Figures

List of Tables

Abbreviations

AP	Action Potential
CCF	Cross Correlation Function
DCG	Differential Connectivity Graph
dDCG	directed Differential Connectivity Graph
digraph	directed graph
DWT	Discrete Wavelet Transform
ECoG	Electrocorticography
EEG	Electroencephalogram
eSOZ	Seizure Onset Zone detected by using induced ictal period by electrical stimuli
EZ	Epileptogenic Zone
FDR	False Discovery Rate
fMRI	Functional Magnetic Resonance Imaging
FWER	Familywise Error Rate
IED	Interictal Epileptiform Discharges
iEEG	intracerebral Electroencephalogram
LI	local information
IZ	Irritative Zone
ℓIED	leading IED
MI	mutual information
MHT	Multiple Hypothesis Testing
MODWT	Maximal Overlap Discrete Wavelet Transform
MMCC	Maximal MODWT cross-correlation
MR	Magnetic Resonance
MSC	Magnitude Squared Coherence
Non-IED	Non Interictal Epileptiform Discharges
PSP	Post Synaptic Potential
SOZ	Seizure Onset Zone
SP	Statistical multi-objective optimization method
T1	Task one
T2	Task two
vSOZ	Visually detected Seizure Onset Zone by epileptologist

Chapter 1

Introduction

We are interested in processing the connectivity graph of epileptic patients from electroencephalogram (EEG) recordings for the identification of leading interictal epileptiform discharges (IED) regions [1, 8–13]. EEG recordings have high temporal resolution which is a great advantage to follow transient epileptic events in different frequency bands. We started our preliminary analysis using scalp EEG recordings (appendix A, page 123) since they are non-invasive and simple [14]. However scalp EEG has very low spatial resolution and to identify the interictal brain regions, one can consider two possibilities: 1) using the inverse problem methods or 2) using invasive intracerebral EEG (iEEG) recordings. We had access to iEEG recordings through neurology department of Grenoble university hospital (CHUG), Grenoble, France, so we continued our research work with iEEG recordings [15].

Available methods for iEEG analysis to estimate leading IED regions through connectivity graph have several problems. These problems can be clustered in three levels as follows.

1. The classic method to measure and interpret the functional connectivity [16, 17] is to measure the interdependencies between the pairs of time series to provide a coupling strength matrix, then an edge is involved in the graph if its related coupling strength is greater than the threshold value [18, 19]. Such graphs are threshold dependent and the threshold can be different for different time intervals. The graph inference methods based on multiple testing [20, 21] are preferred to threshold-based methods. However, usually in these methods certain prior assumptions for the distribution of the test statistics are assumed that are not necessarily true. Multiple testing methods based on resampling like permutation and bootstrap methods allow powerful significance test by assuming less assumptions about the underlying data. However, graph inference methods based on bootstrap cannot be convenient for long time period processing of high-density electrode array recordings due to expensive computation load and processing time [11]. Consequently, a confident graph inference method which can be practically applied for long and high-density electrode arrays of invasive EEG recordings is demanding in the available graph analysis methods. Furthermore the methods for identification of epileptic regions through connectivity graph, use only the interictal periods of EEG recordings [9, 10, 12]. Such graphs are very complicated to be interpreted since other non-interested connections such as the connections of

background activity are also included.

2. Assuming the leading IED regions (sources) be more important than propagated IED regions (sinks) [22,23], the directions of the relationships between brain regions is necessary. Once the directed relationship between brain regions are estimated, the directed graph is characterized to extract the information. To estimate the leading IED regions, a graph measure is needed to quantify the source and sink nodes of the graph. Classic graph measures like global efficiency, local efficiency, and node degree do not include the necessary information for identification of source and sink nodes.

3. Since IEDs can appear in specific frequency band, frequency decomposition of these signals is useful. Obtaining the graph related to each frequency band and its characterization provide multi-dimensional graph measure values. Analysis of multi-dimensional measure values from high-density electrode array of iEEG recordings encounters huge multi-criteria data processing. A data mining tool is needed to extract the estimated leading IED regions from these measure values. In the available studies [9, 12, 24], at each frequency band the nodes of related graph which have the graph measure values greater than a threshold were selected as the IED regions. Therefore a set of IED regions for each frequency band were proposed, while the preference between these frequency bands is unknown. This method or more generally single-objective optimization method, provide a unique solution for each objective function. A method is needed to take into account all of the objective functions simultaneously assuming that the preference between different objective functions is unknown.

In this work, firstly we explain how we can obtain a reliable functional connectivity graph from high-density electrode array of iEEG recordings. Moreover, we explain how to make easier the interpretation of complex IED related graphs. Secondly we explain how we can measure the emittance contribution of each node to the rest of directed graph for source and sink node identification. Thirdly, we explain how a set of solutions can be obtained using multi-dimensional graph measure values without knowing the preference between different dimensions.

Rest of the thesis is organized as follows. The structure of the thesis is summarized in Figure 1.1.

1. Part I: the state of the art is explained including three chapters to describe the necessary background:
   (a) Chapter 2: epilepsy and EEG measurement are explained in this chapter.
   (b) Chapter 3: this chapter is about graph theory including: brain connectivity, definition of graph, graph types, graph measures and coupling measures.
   (c) Chapter 4: in this chapter we review the IED identification methods through connectivity graph.

2. Part II: proposed iEEG analysis method is explained in this part including three chapters:

(a) Chapter 5: this chapter is about the basic ideas of differential connectivity graph (DCG).

(b) Chapter 6: in this chapter, we explain different steps of DCG calculation. Moreover the estimation of causal relationship of the edges of DCG is explained in this chapter.

(c) Chapter 7: in this chapter we explain how to characterize the dDCG and how to estimate the leading IED regions from multi-dimensional graph measure values.

3. Part III: this part is devoted to experimental results including three chapters:

(a) Chapter 8: data and parameters are explained in this chapter.

(b) Chapter 9: the results of dDCG are reported and the choice of parameters and reliability of DCG are discussed.

(c) Chapter 10: the estimated leading IED regions based on different graph measures are compared. Moreover the estimated leading IED regions are compared with the SOZ detected by other methods. The reliability of the results and the choice of parameters are discussed as well.

4. Chapter 11: the concluding remarks, prospectives, and the list of related publications are brought in this chapter.

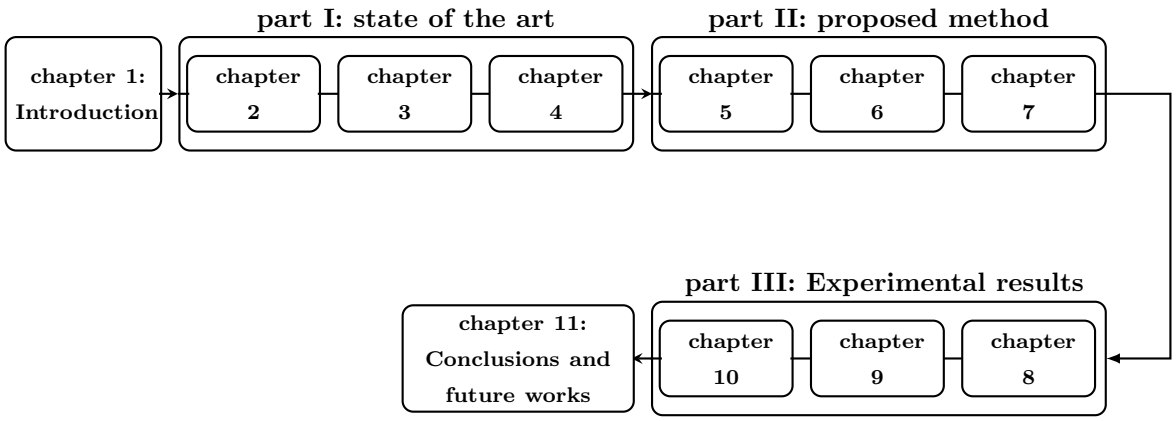


Figure 1.1: A block representation of the thesis structure.

Part I

State of the art

Chapter 2

Epilepsy and EEG

2.1 Introduction

This chapter includes two main sections: 1) epilepsy, and 2) EEG measurement. In section 2.2 (epilepsy), we explain about the definitions of 1) interictal and ictal time periods, 2) IEDs, and 3) different dysfunctional epileptic regions like epileptogenic zone and seizure onset zone. Section 2.3 is devoted to EEG measurements including scalp and invasive EEG recordings.

2.2 Epilepsy

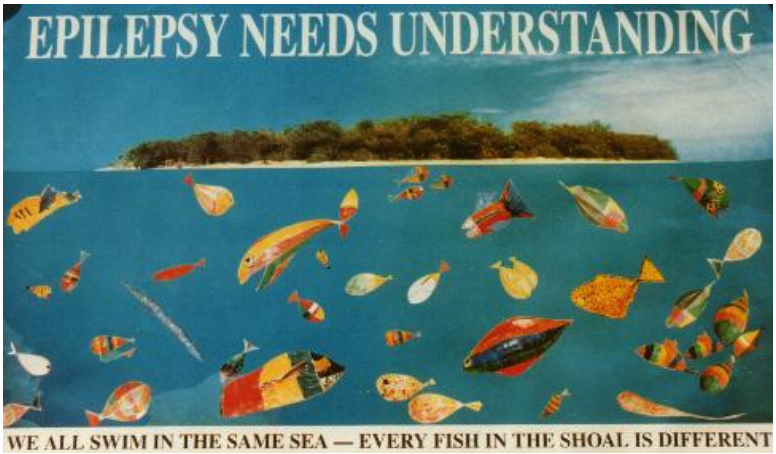


Figure 2.1: "Epilepsy needs understanding. We all swim in the same sea - every fish in the shoal is different." [2]

2.2.1 What is epilepsy?

Epilepsy is defined by the recurrence of epileptic seizures responsible for transient increase in hypersynchronous electrical activity within relatively large neuronal networks. This may induce disruption of normal brain functioning thus causes ictal semiology including motor symptoms, loss of consciousness, hallucinations, etc.

2.2.2 How can we recognize epilepsy?

Patients with epilepsy are usually considered to have two main brain states:

1. *The ictal state:* the ictal state has short duration (few seconds) [25–27] and it is directly related to epileptic seizures.

2. *The interictal state:* this state is represented the vast majority of time in the interval between two seizures.

During interictal state the interictal discharges and during seizures, the seizure patterns can be recognized from EEG recordings.

2.2.3 Interictal epileptiform discharges (IED)

The International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN) defines interictal discharges as a subcategory of "epileptiform pattern," in turn defined as "distinctive waves or complexes, distinguished from background activity, and resembling those recorded in a proportion of human subjects suffering from epileptic disorders" [28]. Sharp waves, spikes, spike-wave complexes (also called spike-and-slow-wave complexes), and polyspike-wave complexes (also called multiple-spike-and-slow-wave-complexes) are morphological divisions of IEDs. In practical terms, the certainty with which these entities can be characterized from physiologic or nonspecific sharp transients are more important than the morphological distinctions. IEDs may appear in isolation or in brief bursts; bursts longer than a few seconds characterize electrographic seizures rather than interictal discharges. The following descriptions are in use [29]:

- *Sharp wave*: distinct from background, transient activity, with duration of 70-200 milliseconds (msec).

- *Spike*: the pattern of spikes is the same as sharp wave, but with duration of 20 to less than 70 msec.

- *Spike-and-slow-wave complex*: the pattern includes a spike followed by a slow wave (in general the slow wave has higher amplitude than the spike).

- *Multiple spike-and-slow-wave complex*: the pattern is the same as spike-and-slow-wave complex, but with 2 or more spikes related to one or more slow waves.

2.2.4 Classification of epileptic seizures

Seizures can be generally divided into two groups, partial, and generalized. Incidence of seizures by type is illustrated in Figure 2.2.

- *Partial*: partial seizures occur in a segment of one hemisphere of the brain, but can sometimes secondarily generalize. Partial seizures last only a few seconds, and are categorized into three sub-classes:

  - *Simple*: the most focal of the partial seizures. Consciousness is not affected.
  - *Complex*: the most usual seizure class that is not restricted to one seizure focus. Consciousness affected to some extent or is lost.
  - *Secondary generalized*: a simple or complex seizure spreads to all areas of the brain.

- *Generalized*: generalized seizures contain all areas of the brain, i.e. there is no anatomical localization. Consequently, their effect on the patient is more serious than partial seizures. Generalized seizures are classified into three sub-classes:

  - *Tonic-clonic*: most patients loose consciousness without warning, but some of them experience a forewarning or "aura" before a tonic-clonic seizure. The muscles suddenly contract, and thus the patient falls and lies rigidly for 10-30 seconds (sec), throughout the first tonic phase. When the patient inhales, a high-pitched sound called stridor may be emitted, if the throat or larynx is involved. As the seizure enters the clonic phase, spasms related to alternate relaxation and contraction of muscles arise for 30-60 sec. The bowel or bladder control of patient may be lost. The complete seizure usually lasts a total of 2-3 minutes (min), and next the patient may remain unconscious for some minutes. The patient is usually confused and extremely tired, on regaining consciousness. After a tonic-clonic seizure, a serious headache like migraine happens.
  - *Absence*: absence seizures may include a short arrest of physical movement and loss of attention during a short duration (<10-45 sec). They may even be undetected by others. Mild clonic motions (usually movement of the eyelids), automatisms (unconscious acts such as opening and closing a door), decreased postural tone (such as head nodding), increased postural tone (arching of the back), and autonomic phenomena are some general manifestations. Absence seizures and loss of consciousness may arise as often as 50-100 times a day.
  - *Other seizures*:
    * *Atonic (akinetic) seizure*: affects loss of muscle tone. It may involve a single body part or the whole body.
    * *Myoclonic seizures*: include a series of brief jerky contractions of specific muscle groups, like the face or trunk.
    * *Simple tonic or clonic seizures*: in tonic seizures, the muscles contract and consciousness is changed for about 10 sec, but the seizures do not develop the clonic or jerking phase. Clonic seizures, that are rare, arise initially in young children, who experience spasms of the muscles but not their tonic rigidity.

2.2.5 Surgical treatment

Where antiepileptic drug treatments fail for some of patients with partial or focal epilepsy, the epilepsy surgery may cure [30]. This surgical procedure is aimed at removing the brain tissue

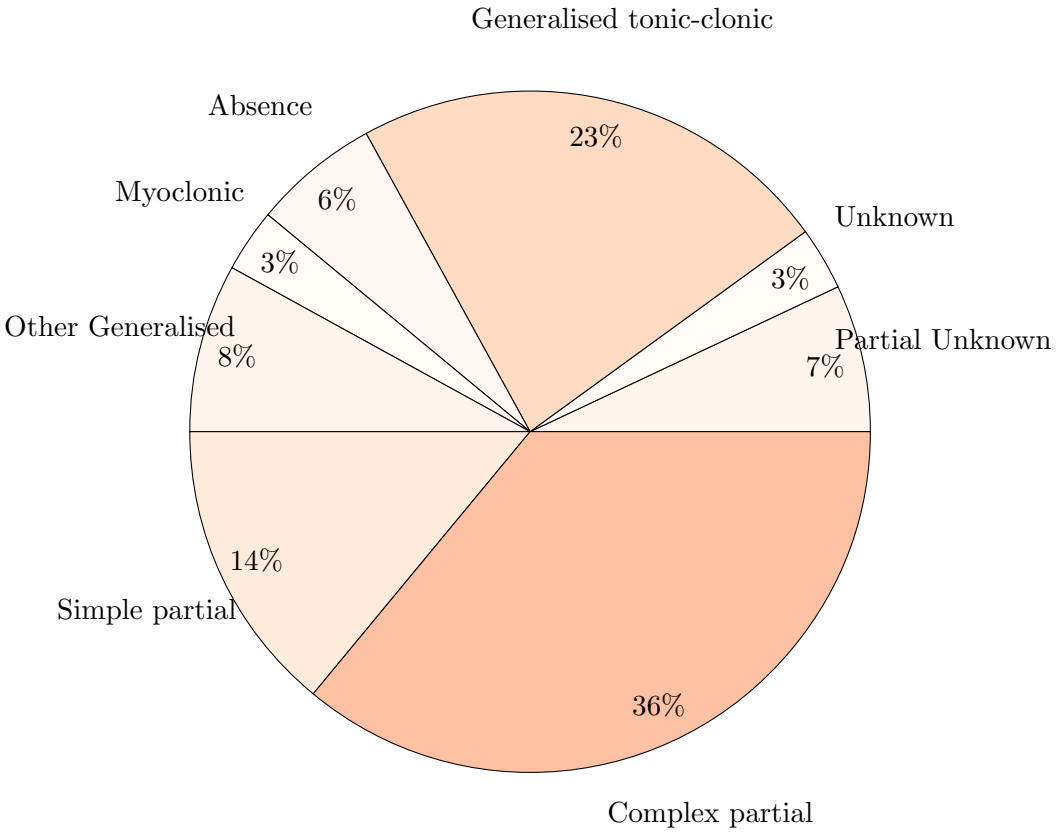


Figure 2.2: Proportion of incident of epilepsy in Rochester, Minnesota, USA, between 1935 and 1984 [3].

causing epileptic seizure without creating new neurological deficits. Despite the improvement of resective surgical outcome in treatment of epileptic seizures in the past decade, the failure of surgical outcome is still 10% to 40% in temporal lobe epilepsy and 34% to 55% in extratemporal lobe epilepsy [31]. These statistics show that although valuable improvements, still pre-surgery evaluations need more developements.

2.2.6 Dysfunctional epileptic regions (EZ, SOZ and IZ) and their relationships

Several dysfunctional regions have been identified for epileptic patients, based on neurophysiological and surgical considerations [14, 30]. Here, we define three major important ones as follows:

- *Epileptogenic zone (EZ)*: the area of the cortex that is essential for the generation of epileptic seizures. In other words, the region which its removal is both necessary and sufficient to abolish seizures.
- *Seizure onset zone (SOZ)*: the region where the first electrophysiological changes is detected at ictal onset.
- *Irritative zone (IZ) or IED regions*: the site responsible for IEDs generation.

Relationships between SOZ, EZ and IZ have been the subject of numerous studies and debates [14]. The extent of SOZ may be greater or smaller than EZ: if EZ is smaller than SOZ, partial resection of SOZ may lead to seizure free outcome, because the rest of SOZ cannot generate seizures, or equivalently because the essential region for seizure generating is removed. In other words, if we were able to find and disconnect EZ, other regions of the SOZ would not be able to generate seizure any more. Contrarily, if EZ is greater than SOZ, even totally removing SOZ will not result in seizure-freedom. However currently the extent and location of EZ cannot be identified before surgery. More precisely if the patient is seizure-free after surgery, then epileptologists deduce that the removed region included the EZ. Therefore practically one must infer the location and extent of EZ indirectly from location and extent of other dysfunctional areas like SOZ and IZ [14]. The best way to delineate SOZ is to record seizures by using implanted intracerebral depth electrodes providing intracerebral EEG (iEEG) [1, 30, 32–35]. However recording ictal EEG is difficult since seizures are not frequent in most of the patients. Therefore the number of seizures are limited and obtaining statistically reliable results from studying seizure time intervals is quite difficult [22]. Consequently, SOZ prediction through interictal analysis can be valuable.

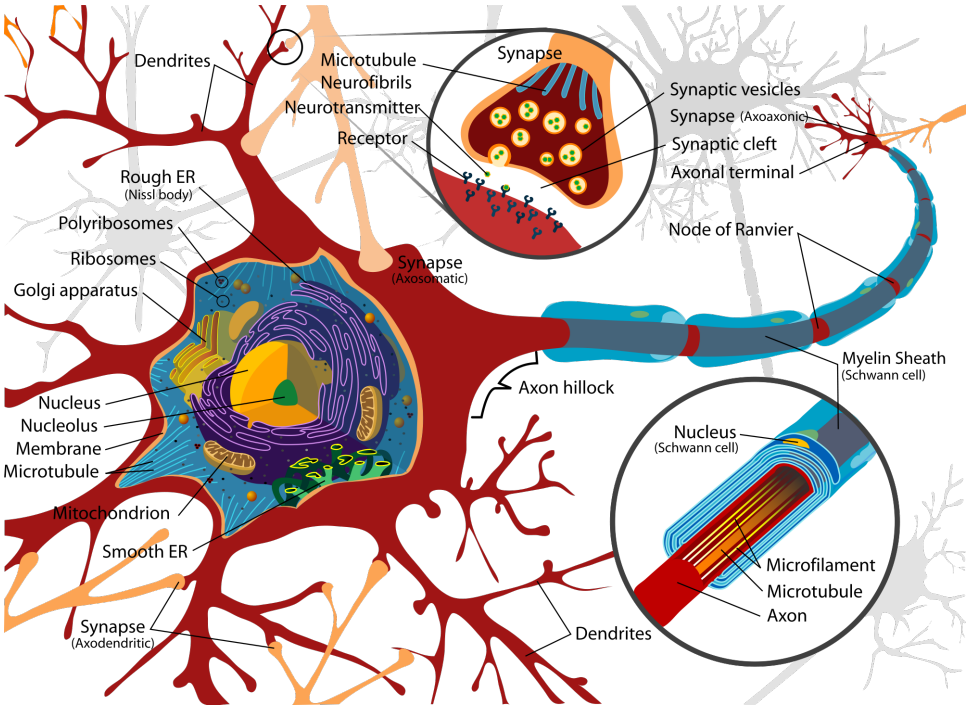


Figure 2.3: The anatomy of a typical neuron [4].

2.3 EEG measurement

2.3.1 What does EEG record?

The schematic diagram of a neuron including its axon, dendrites and synapse is depicted in Figure 2.3 (taken from [4]). To explain what EEG measures, we start with some general definitions as follows:

- *Synapse:* the area where two neurons signal to each other. Synapse permits the neurons to communicate with each other and also to connect and control other systems like muscles or glands [36].
- *Neurotransmitters:* the chemical molecules which transmit the signals from a source neuron to a target one across the synapse.
- *Post-synaptic neuron:* the target neuron or the neuron on the opposite side of synapse.
- *Action potential (AP):* the short-lasting alternation of the membrane potential is called action potential or nerve impulse. Voltage alternation is due to ion movement.
- *Post-synaptic potential (PSP):* temporary changes of membrane potential of post-synaptic neuron. The membrane potential decreases when the neuronal activity is inhibited, conversely when the membrane potential rises over certain threshold, the neuronal activity is excitatory. A PSP is defined as excitatory when it increases the probability of the occurrence of AP, while inhibitory PSP decreases the probability of a future AP.

- *Resting state:* resting state is the condition where there is no perturbation to the neuron opposed to neuronal activity state which neuron receives the information or transmits it to another neuron. The resting state membrane potential difference between inside the membrane and the surrounding media is about 70 mv.

The AP or the nerve impulse travels along the axon, when it reaches the synapse, it provokes the neurotransmitters to be released in the synapse to bind to chemical receptor molecules located in the membrane of post-synaptic neuron. This causes the electrical current in the dendrites and therefore depolarizations or hyperpolarization in the membrane at the axon hillock (Figure 2.3) of the post-synaptic neuron, i.e. increase or decrease of post-synaptic membrane potential or PSPs. The sum of these PSPs may provide an AP which propagates down the axon of post-synaptic neuron. Now this AP can provoke another AP in another neuron of the neuronal network. This is how the information is transmitted in the neuronal networks. In EEG recordings, we measure the PSP [37].

Figure 2.4: Scalp EEG recording, CHUG, Grenoble, France.

2.3.2 Scalp EEG

The IED recordings from scalp EEG were first reported by Hans Berger in 1929 [38]. Still the scalp EEG is the most common diagnosis tool for epilepsy.

Scalp EEG reflects the accumulation of the synchronous activity of thousands to millions of

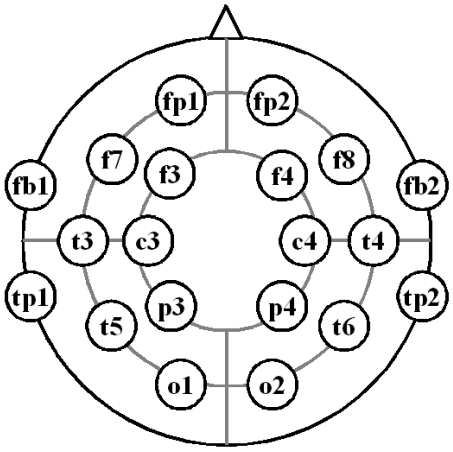


Figure 2.5: A typical scalp EEG electrodes' placement in 10/20 system.

neurons oriented radially to the scalp. Scalp EEG measurement is sensitive to the activation of these neurons in the cortex, while the activations of the neurons in deep regions of gray matter cannot be reflected in this measurement. Besides, the potentials associated with vectors tangential to the surface can be ignored by scalp EEG recordings. The EEG recordings from scalp are attenuated and can be distorted because of meninges, skull and scalp. Eventually EEG localization by scalp EEG is very complicated [39]. However scalp EEG (Figure 2.4) is noninvasive, cheap, and simple. Figure 2.4 demonstrates the scalp EEG recording in epilepsy department of CHUG. A schematic diagram of scalp EEG placement with electrode labels is demonstrated in Figure 2.5.

2.3.3 Invasive EEG

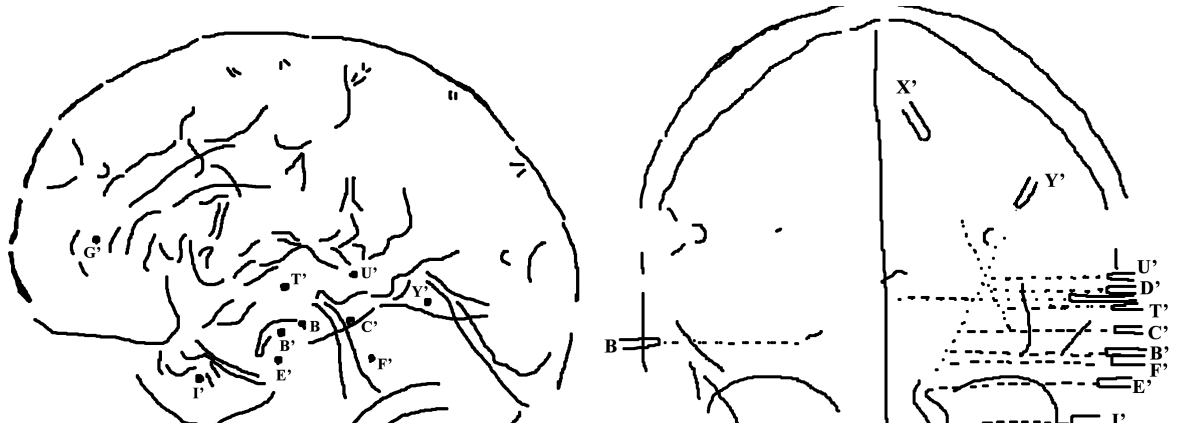
Foester and Altenburger published the first report on measuring electrical activity directly from human cerebral cortex [40] in 1935 and electrocorticography (ECoG) recordings became crucial in surgical treatment of temporal lobe epilepsy in early 1950s. In 1950s the chronic intracranial electrodes were also suggested permitting the ictal activity recording [41] . Scalp EEG and ECoG were now commonly used for IZ and SOZ detection for presurgical studies in addition to clinical knowledge and ictal symptoms that had been previously used.

The ECoG can be recorded using subdural grids or strips located directly on the lateral surface of the brain [42]. There are other clinical routines for recording invasive EEG using multi-lead depth electrodes [15]. The depth electrode implantations (Figure 8.1) are usually referred to as stereotactic EEG due to stereotactic technique developed by Talairach and Bancaud for the electrodes localization [14, 43–46]. In this manuscript, such implantations are referred to intracerebral EEG (iEEG) recordings. Multilead depth electrodes are semi-flexible and shaped as narrow needles. The implantation of these electrodes needs only small burr holes opposed to larger craniotomy of subdural grid implantations. The subdural grid provides cortical coverage and cortical maps of gyral activity, while depth multi-lead electrodes have the potential to

record the activity of both *sulci* and *gyri* and from beneath the cortical surface to deep cortical structures. However both recordings have sufficient spatial resolution to localize neural activity at the gyral level as good as functional magnetic resonance imaging (fMRI).

Simultaneous video recording of clinical manifestations with continuous EEG recording is used frequently in presurgery evaluations [47–50]. During video-EEG monitoring sufficient number of IEDs during awake and sleep stages can be recorded. Seizures are less frequent and mostly long recordings is required to record a few seizures.

A typical segment of simultaneous iEEG and video recordings is demonstrated in Figure 2.7. In this Figure the labels from 3 to 4 indicate some examples of IED time interval, while the labels from 7 to 8, indicate examples of non-IED time intervals. Please note that all of IED time intervals are not labeled in this Figure. A picture of the video-iEEG recording center of Neurology department of CHUG is shown in Figure 2.8.



(a) Sagittal view (b) Coronal view

Figure 2.6: Implantation scheme of the iEEG electrodes (a) Sagittal view; (b) Coronal view.

2.3.4 EEG and fMRI

Invasive EEG and fMRI

Although newer technologies like fMRI, video-iEEG monitoring still remains gold standard in the definition of IZ and SOZ [14]. fMRI measures blood flow changes, i.e. relative decrease in interictal blood flow with respect to ictal state. This recording allows non-invasive evaluation of all areas of the brain with the same accuracy. However, fMRI recording is short, therefore it is limited to the patients who have very frequent seizures and also with retained consciousness and without motor symptoms. Movements and related artifacts prevent meaningful fMRI recordings. The spatial resolution of invasive EEG recordings and fMRI is comparable, while the temporal resolution of invasive EEG recordings of about msec is not comparable with the temporal resolution of fMRI that is about 1-4 sec.

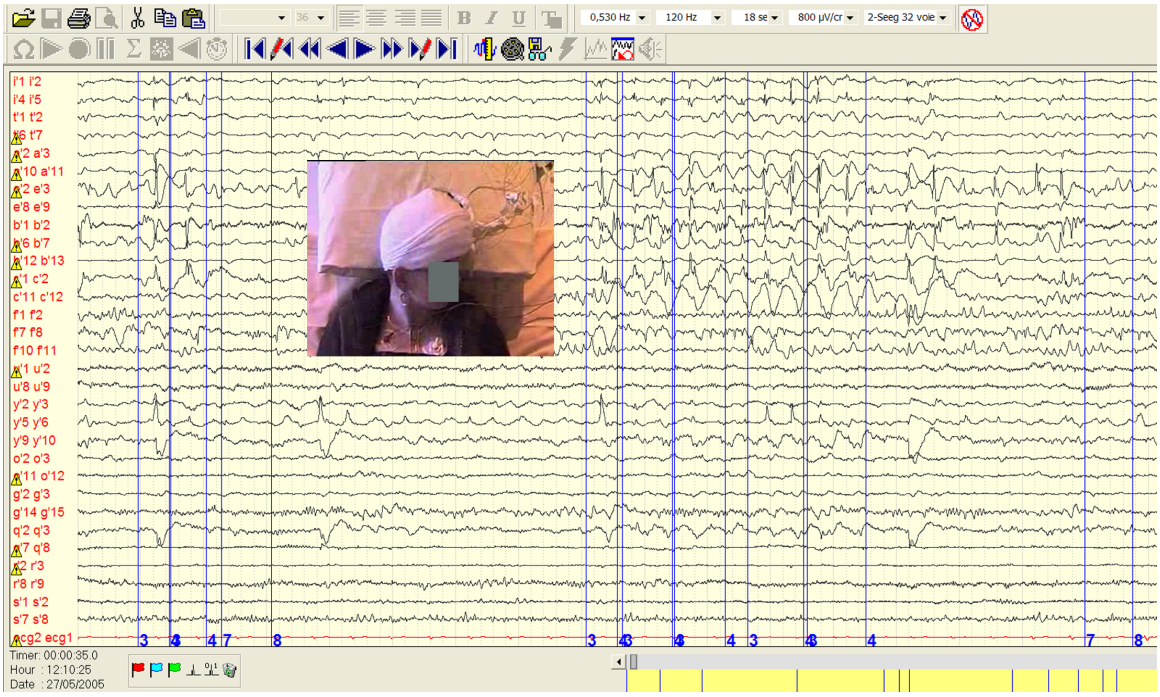


Figure 2.7: A typical segment of iEEG recordings. Some typical IED and non-IED time intervals are labeled. Labels 3, 4, 7, 8 show onset and offset of IED and non-IED time intervals, respectively.

Figure 2.8: Video-iEEG recording center in Neurology department, CHUG, Grenoble, France.

Figure 2.9: Simultaneous scalp EEG and fMRI recording, CHUG, Grenoble, France.

Simultaneous scalp EEG and fMRI

Integrating high spatial resolution of fMRI and high temporal resolution of EEG seems interesting in simultaneous scalp EEG and fMRI analysis. Figure 2.9 demonstrates the simultaneous EEG and fMRI recording in CHUG. However EEG recorded inside a magnetic resonance (MR) scanner is interfered with very strong artifacts including: 1) MR artifacts, 2) RF pulse artifacts, and 3) balistocardiogram (BCG) artifacts.

The MR artifact is caused by the switching of the magnetic filed gradients used in the image acquisition, which induces unwanted voltages on the EEG recordings. The RF pulse artifact is due to the time varying electromagnetic field pulses (RF pulses) used for excitation in MRI recordings. Although, the RF pulse has a very high frequency (in the order of several megahertz); but it is nonlinearly rectified to low frequencies (below 100Hz), within the EEG bandwidth. BCG artifact is caused by the micro motions of head EEG leads and wires within the static magnetic field. These motions are related to the pulsatile blood-flow in the head.

Although there are efficient methods for removing these artifacts (appendix B, page 129), it is difficult to know how much of EEG information is removed during artifact removal methods. However the non-contaminated EEG should be clean enough that IED time intervals can be detected. IEDs are spontaneous events and cannot be controlled, therefore to analyze fMRI recorded simultaneously with EEG, the IED time intervals are required to be detected from non-contaminated EEG.

2.4 Conclusion

Drug resistant epeileptic patients are recommended to undergo the resective surgery. Practically the aim of resective surgery is to remove the SOZ. The best way of SOZ detection is ictal analysis of iEEG recordings. However seizures are infrequent in most of the patients and thus obtaining statistically reliable results from a few seizures is complicated. Therefore SOZ prediction through interictal iEEG analysis remains valuable.

Chapter 3

Graph theory

3.1 Introduction

In this chapter first we explain the necessary background of graph theory including the definition of graph, graph types, and graph measures used in this work. Next different coupling measures to measure the functional connectivity are explained.

3.2 Brain connectivity

Brain connectivity [16, 17] refers to three main topics: anatomical connectivity, functional connectivity and effective connectivity.

- **Anatomical connectivity:** this connectivity is the set of structural links between neural units at a given time. The neural units refer to individual neurons, neuronal populations, or anatomically segregated brain regions. The anatomical connections are relatively static in short time durations (seconds to minutes), while these connections can be dynamic in longer time durations (hours to days).

- **Functional connectivity:** functional connectivity is time dependent, i.e. the temporal couplings or interdependencies between the activities recorded from spatially remote brain regions associated with neural units.

- **Effective connectivity:** effective connectivity is based on the effect of one neural unit over another. Effective connectivity can be estimated using model-based methods (like structural equation modeling [51], dynamic causal modeling [52–56], or Granger causality [57]) or model-free methods by estimating the time causality between neural events recorded from different brain regions [58–63].

Anatomical and functional connectivity are not independent according to the theory of brain plasticity, where thinking, learning, and acting modify both of these connectivities. For example learning in adulthood is not only through the changes of the strength of the connections, but also through adding or eliminating connections, or adding cells. Furthermore, effective connectivity

can be derived from functional connectivity [16, 17, 64, 65] based on the causal information between the interacting regions using model-free methods. Consequently, functional connectivity has the most important role among the three concepts of brain connectivity.

In this work we focus on functional connectivity and the effective connectivity is derived from functional connectivity using model-free methods.

The interpretation of the measured functional connectivity can be illustrated by a graph. In the following section, we will explain the graphs and some of their related concepts in graph theory.

3.3 Graph

A graph is a set pair of vertices (or nodes or points or dots) and edges (or connections or lines or links), which edges connect the vertices. Graph theory [66–72] helps us to understand the behavior and properties of a variety of systems like internet, World-Wide Web, social networks and brain neural networks. Systems in the form of networks or mathematically graphs abound in the world [73].

Reviews in [17,73–80] permit us to know about the graph theory and related problems. In the following, we explain the types of graphs, some general definitions for directed graphs, widely used usual graph measures and coupling measures.

3.4 Types of graphs

There are different types of graphs which some of them are explained in appendix C (page 141). In this work we focus on a particular directed graph (digraph) in which each edge is permitted to have a unique direction, i.e. between nodes i and j only one direction exist. Such graphs include neither multiple edges nor loops. These particular digraphs are called oriented graphs. However we simply call them digraphs in the following.

3.5 General definitions for the digraphs

Here we explain about the general definitions for digraphs which further we use to define the graph measures.

3.5.1 Adjacency matrix

The functional connectivity structure of a digraph can be represented by the adjacency matrix indicating whether two nodes communicate or not and in which direction.

Let's assume digraph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ and $\mathbf{E}$ are the set of vertices and edges. We describe the digraph $\mathbf{G}$ with an adjacency matrix denoted as $\mathbf{A_G} = [a_{ij}] \in \{0, 1\}^{N \times N}$, where N is the number of nodes. The adjacency matrix of digraphs (oriented graphs) has no symmetric

pair of directed edges (each edge is permitted to have a unique direction) and has a diagonal of zeros (no loops are allowed). If there is an edge from node i to node j, then $a_{ij} = 1$ and $a_{ji} = 0$, otherwise if there is no edge, then $a_{ij} = a_{ji} = 0$. The maximum possible number of directed edges for such digraphs is equal to $N_c = (N^2 - N)/2$ ($\sum_{ij} a_{ij} = (N^2 - N)/2$).

3.5.2 Path and shortest path length

A path from node i to node j (or outgoing path of i, or incoming path of j) is an ordered sequence of non-repeated edges and nodes connecting node i to node j [17]. The shortest path length between nodes i and j is the minimum number of edges traversed to get from i to j.

We assume the shortest path length matrix denoted as $\mathbf{L_G} = [l_{ij}]$, where l_{ij} is the shortest path length between nodes i and j. If any path does not exist from i to j, then $l_{ij} = \infty$.

3.5.3 Size and order

The number of nodes (N or the length of set $\mathbf{V}$) and number of edges (length of set $\mathbf{E}$) are defined as the size and order of a digraph, respectively.

3.5.4 Subdigraph

Subdigraph $\mathbf{G}_{i^-} = (\mathbf{V}_i, \mathbf{E}_i)$ is the digraph of the first-order outgoing neighbors of node i excluding node i, i.e. $\{j \neq i \in \mathbf{V} |\ a_{ij} = 1\}$. When node i is included, the subdigraph is denoted as $\mathbf{G}_{i^+}$.

3.6 A sample digraph

Figure 3.1 demonstrates a simple example of digraph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ (left) and subdigraph $\mathbf{G}_{7^-} = (\mathbf{V}_7, \mathbf{E}_7)$ (right). The size of digraph $\mathbf{G}$ is equal to 7 or equivalently $\mathbf{V} = \{1, 2, 3, 4, 5, 6, 7\}$. The order of digraph $\mathbf{G}$ is equal to 12, i.e. $\mathbf{E}$ includes 12 directed edges among $N_c = 21$ possible edges. The adjacency ($\mathbf{A_G}$), and shortest path length ($\mathbf{L_G}$) matrices of digraph $\mathbf{G}$ are as follows:

$$\mathbf{A_G} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}, \mathbf{L_G} = \begin{pmatrix} 3 & 1 & 2 & 3 & 4 & 1 & 2 \\ \infty & \infty & 1 & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty \\ 3 & 3 & 1 & 3 & 1 & 2 & 2 \\ 2 & 2 & 2 & 2 & 3 & 1 & 1 \\ 2 & 2 & 2 & 2 & 3 & 3 & 1 \\ 1 & 1 & 1 & 1 & 2 & 2 & 3 \end{pmatrix}. \tag{3.1}$$

The shortest path length matrix ($\mathbf{L}_{\mathbf{G}_{7^-}}$) of subdigraph $\mathbf{G}_{7^-} = (\mathbf{V}_7, \mathbf{E}_7)$, $\mathbf{V}_7 = \{1, 2, 3, 4\}$ is as follows:

$$\mathbf{L}_{\mathbf{G}_{7^-}} = \begin{pmatrix} \infty & 1 & 2 & \infty \\ \infty & \infty & 1 & \infty \\ \infty & \infty & \infty & \infty \\ \infty & \infty & 1 & \infty \end{pmatrix} \tag{3.2}$$

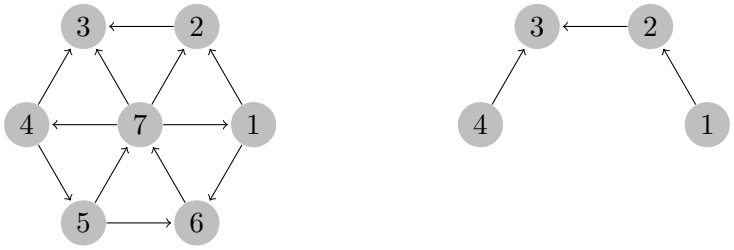


Figure 3.1: A sample digraph $\mathbf{G}$ (left) and the subdigraph of node 7, $\mathbf{G}_{7^-}$ (right).

3.7 Graph measures

There are different graph measures like clustering coefficient, path length (appendix D, page 143), betweenness, centrality, etc. Here we focus on node degree, global and local efficiencies.

3.7.1 Indegree and outdegree

The outdegree $k_{out}[i]$ of node i is the number of outgoing edges incident the node [77]:

$$k_{out}[i] = \sum_{j \in \mathbf{V}} a_{ij} \tag{3.3}$$

Accordingly, indegree $k_{in}[i]$ of node i is related to its incoming edges:

$$k_{in}[i] = \sum_{j \in \mathbf{V}} a_{ji} \tag{3.4}$$

where a_{ij} are the entries of adjacency matrix $\mathbf{A}_{\mathbf{G}}$. The total degree of node i, $k_{tot}[i]$, can be defined as the difference between $k_{out}[i]$ and $k_{in}[i]$: $k_{tot}[i] = k_{out}[i] - k_{in}[i] = \sum_{j \in \mathbf{V}} (a_{ij} - a_{ji})$.

The total degree ($\mathbf{k}_{tot} = [k_{tot}[1], \ldots, k_{tot}[7]]^T$) for the digraph $\mathbf{G}$ depicted in Figure 3.1 is equal to the vector $\mathbf{k}_{tot} = [1, -1, -3, 1, 1, -1, 2]^T$.

3.7.2 Global efficiency

If we define the efficiency of the connection between node pair (i, j) as $e_{ij} = 1/l_{ij}$ [81], then the global efficiency of the node $i \in \mathbf{V}$ is defined as:

$$E_{glob}[i] = \frac{1}{N-1} \sum_{j \in \mathbf{V}, j \neq i} \frac{1}{l_{ij}} \tag{3.5}$$

and the global efficiency of the whole digraph $\mathbf{G}$ is defined [81] as:

$$E_{glob}[\mathbf{G}] = \frac{1}{N} \sum_{i \in \mathbf{V}} E_{glob}[i]. \tag{3.6}$$

$E_{glob}[i]$ and $E_{glob}[\mathbf{G}]$ are in the range $[0, 1]$. $E_{glob}[i]$ is equal to zero for a node without any outgoing path and equal to one if $l_{ij} = 1$, $\forall j \neq i$, $j \in \mathbf{V}$, i.e. if from node i we can reach each node of the digraph with one single edge.

We calculate global efficiency (3.5)-(3.6) for the example depicted in Figure 3.1. The global efficiency ($\mathbf{E}_{glob} = [E_{glob}[1], \ldots, E_{glob}[7]]^T$) for the seven nodes of digraph $\mathbf{G}$ and the global efficiency of digraph $\mathbf{G}$ are equal to $\mathbf{E}_{glob} = [0.6, 0.17, 0, 0.61, 0.67, 0.56, 0.83]^T$ and $E_{glob}(\mathbf{G}) = 0.5$, respectively. Global efficiency of node 7 is high since from this node we can reach the rest of digraph with one or two edges (the seventh row of matrix $\mathbf{L_G}$).

3.7.3 Local efficiency

The $E_{glob}[\mathbf{G}]$ can also be evaluated for the subdigraph $\mathbf{G}_{i^-}$ to measure the local properties. The local efficiency of node i can be defined as the evaluation of (3.6) for the subdigraph $\mathbf{G}_{i^-}$ as following:

$$E_{loc}[i] = E_{glob}[\mathbf{G}_{i^-}]. \tag{3.7}$$

$E_{loc}[i]$ shows the efficiency of the connections between the first-order outgoing neighbors of i when i is removed. Equivalently, local efficiency measures the "resilience" of digraph to the damage of node removal, i.e. if we remove a node, how efficient its first-order outgoing neighbors can communicate. $E_{loc}[i]$ is in the range $[0, 1]$. $E_{loc}[i]$ is equal to zero if the first-order outgoing neighbors of node i are not connected with each other, i.e. $\mathbf{G}_{i^-}$ is disconnected. $E_{loc}[i]$ is also equal to zero if node i is not connected to any other node or if node i does not have any first-order outgoing neighbor.

Local efficiency (3.7) is calculated for the example depicted in Figure 3.1. The local efficiency ($\mathbf{E}_{loc} = [E_{loc}[1], \ldots, E_{loc}[7]]^T$) for the nodes of digraph $\mathbf{G}$ is equal to $\mathbf{E}_{loc} = [0, 0, 0, 0, 0.5, 0, 0.3]^T$. For calculation of local efficiency of node 7, the subdigraph $\mathbf{G}_{7^-} = (\mathbf{V}_7, \mathbf{E}_7)$ depicted in Figure 3.1 (right) is used. $E_{loc}[5]$ is higher than $E_{loc}[7]$ because the neighbors of node 7 (nodes 1, 2, 3, and 4) do not communicate as well as neighbors of node 5 (nodes 5, 6, and 7 make a triangle).

3.8 Coupling measure

A graph can be inferred using different coupling measures between signal pairs providing a N by N matrix for N signals. The different methods of graph inference are explained in section 4.2.2, page 35. In this section, we explain some of widely used non-directional and directional coupling measures [82–84].

3.8.1 Non-directional couplings

Here we explain the following non-directional coupling measures:

1. Cross-correlation
2. Wavelet cross-correlation
3. Coherence
4. Wavelet coherence and wavelet phase Synchrony
5. Hilbert phase Synchrony
6. Mutual information.

Cross-correlation

We assume N-dimensional observations $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$, where $\mathbf{x}_i = [x_i[1], \ldots, x_i[T]]^T$, the ith column of matrix $\mathbf{X} \in \mathbb{R}^{T \times N}$, contains T samples of the signal recorded from the ith channel, and N is the number of channels. The cross-correlation coefficient [85–88] between channel pair (i, j) is defined as:

$$\hat{\rho}\left(\mathbf{x}_i, \mathbf{x}_j, \tau\right) = \frac{\widehat{cov}(x_i[k], x_j[k-\tau])}{\sqrt{\widehat{var}(x_i[k])\widehat{var}(x_j[k-\tau])}} \tag{3.8}$$

where $\widehat{cov}$ and $\widehat{var}$ [85] are estimated by empirical average and τ is the time delay. The cross-correlation is in the range $[-1, 1]$. A perfectly linearly correlated signal pair has the maximum cross-correlation equal to 1 if the signals are in phase. Accordingly, perfectly correlated signal pair which signals are in opposition has the minimum cross-correlation equal to -1.

Wavelet cross-correlation

Let's assume $\mathbf{d}_i^{j_w}$ be the wavelet coefficients of $\mathbf{x}_i$ at level j_w. Wavelet coefficients of level j_w are associated with frequency interval $[\frac{f_s}{2^{j_w+1}}, \frac{f_s}{2^{j_w}}]$, where f_s is the sampling rate (Hz). The wavelet cross-correlation [20, 82, 89] between $\mathbf{d}_i^{j_w}$ and $\mathbf{d}_j^{j_w}$ at level j_w is estimated as:

$$\hat{\rho}\left(\mathbf{d}_i^{j_w}, \mathbf{d}_j^{j_w}, \tau\right) = \frac{\widehat{cov}(d_i^{j_w}[k], d_j^{j_w}[k-\tau])}{\sqrt{\widehat{var}(d_i^{j_w}[k])\widehat{var}(d_j^{j_w}[k-\tau])}} \tag{3.9}$$

Values of the wavelet cross-correlation coefficients are in $[-1, 1]$. The cross-correlation can also be calculated for narrow band pass filtered signals [83].

Coherence

The coherence [90–94] between $\mathbf{x}_i$ and $\mathbf{x}_j$ at frequency f can be estimated as below:

$$\hat{\varrho}\left(\mathbf{x}_i, \mathbf{x}_j, f\right) = \frac{\left|\widehat{W}_{\mathbf{x}_i\mathbf{x}_j}[f]\right|}{\sqrt{\widehat{W}_{\mathbf{x}_i\mathbf{x}_i}[f] \cdot \widehat{W}_{\mathbf{x}_j\mathbf{x}_j}[f]}} \tag{3.10}$$

where $\widehat{W}_{\mathbf{x}_i\mathbf{x}_j}[f]$ is the estimated cross-spectra density function between $\mathbf{x}_i$ and $\mathbf{x}_j$ or Fourier transform of their cross-correlation function. $|\cdot|$ refers to magnitude. $\hat{\varrho}\left(\mathbf{x}_i, \mathbf{x}_j, f\right)$ is in the range $[0, 1]$. If the coherence is close to one for all of the frequencies, then there is a linear relationship between $\mathbf{x}_i$ and $\mathbf{x}_j$. Coherence is roughly equal to the correlation coefficient between the frequency components of $\mathbf{x}_i$ and $\mathbf{x}_j$ at frequency f [91, 95, 96].

Fourier-based coherence (3.10) is not suitable for the analysis of non-stationary signals, while wavelet coherence can detect transient linear couplings [96]. In the following we explain wavelet coherence, wavelet phase synchrony, and Hilbert phase synchrony.

Wavelet coherence and wavelet phase Synchrony

The instantaneous complex phase vector of signals $\mathbf{x}_i$ and $\mathbf{x}_j$ in the frequency range associated with level j_w can be estimated [97] as:

$$\widehat{\boldsymbol{\varphi}}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right) = \frac{\mathbf{d}_i^{*j_w} \cdot \mathbf{d}_j^{j_w}}{\left|\mathbf{d}_i^{j_w}\right| \cdot \left|\mathbf{d}_j^{j_w}\right|} \tag{3.11}$$

where $*$ denotes complex conjugate. Many complex wavelets may be chosen like Morlet or Hilbert. To provide less noisy complex phase measure [97], a smoothing can be applied as follows:

$$\widehat{\varphi_s}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right) = \frac{\langle \mathbf{d}_i^{*j_w} \cdot \mathbf{d}_j^{j_w} \rangle}{\sqrt{\langle \left|\mathbf{d}_i^{j_w}\right|^2 \rangle \cdot \langle \left|\mathbf{d}_j^{j_w}\right|^2 \rangle}} \tag{3.12}$$

where $\langle\rangle$ denotes averaging over window length of Δk, which Δk should be selected properly. Lachaux et al. explained about the selection of Δk in [96]. The wavelet coherence [96, 97] can be defined as:

$$\widehat{\varrho_w}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right) = \left|\widehat{\varphi_s}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right)\right| \tag{3.13}$$

The magnitude squared coherence (MSC) is defined as the squared of (3.13).

A dynamic study of the wavelet coherence between signal pairs is proposed in [97]. For this purpose $\widehat{\varphi_s}$ was calculated for sliding widows of length Δk where k is the sample index. For a window starting at time sample k_1, $\widehat{\varphi_s}$ was calculated between $\mathbf{d}_i^{*j_w}[(k+k_1) : (k+k_1+\Delta k)]$ and $\mathbf{d}_j^{j_w}[(k+k_1) : (k+k_1+\Delta k)]$ which its value substituted with the complex phase value of sample k_1 giving a smoothed complex phase vector denoted as $\widehat{\boldsymbol{\varphi_s}}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right)$. $\widehat{\boldsymbol{\varphi_s}}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right)$ includes complex phase values in time and in specific frequency range for the pair signal $\mathbf{x}_i$ and $\mathbf{x}_j$. To

study the dynamic relationships, the dynamic wavelet coherence can be considered as:

$$\widehat{\boldsymbol{\varrho}_w}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right) = \left|\widehat{\boldsymbol{\varphi}_s}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right)\right|. \tag{3.14}$$

Following to dynamic wavelet coherence definition, we explain different definitions of phase difference. Once the phase difference is obtained, different synchronization indices [98–100] like phase locking value (3.17) [94, 101] can be calculated to characterize the phase relationship between underlying signal pair.

In [97] the phase difference between signal pairs $\mathbf{x}_i$ and $\mathbf{x}_j$ in the frequency range related to level j_w was defined as:

$$\widehat{\Delta\phi_s}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right) = \arg(\widehat{\varphi_s}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right)). \tag{3.15}$$

The phase locking or phase synchrony interval was defined [97] as time period which has the following conditions: 1) when $\left|\widehat{\Delta\phi_s}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right)\right|$ is smaller than a threshold or the variation of phase difference is small. 2) when the wavelet coherence (3.13) is greater than a threshold. The two thresholds are required to be set properly.

In [94], the phase difference was defined differently as:

$$\widehat{\Delta\boldsymbol{\phi}}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right) = \arg(\widehat{\boldsymbol{\varphi}}(\mathbf{x}_i, \mathbf{x}_j, j_w)). \tag{3.16}$$

The phase synchrony or phase locking value [94] was defined as:

$$\widehat{PL}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right) = \left|\langle e^{j\widehat{\Delta\boldsymbol{\phi}}(\mathbf{x}_i, \mathbf{x}_j, j_w)}\rangle\right| = \left|\langle\widehat{\boldsymbol{\varphi}}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right)\rangle\right| \tag{3.17}$$

where averaging is over time period $\Delta k'$. $\widehat{PL}\left(\mathbf{x}_i, \mathbf{x}_j, j_w\right)$ measures the variability of the phase difference in time period $\Delta k'$. If the variation of phase difference is small during time period $\Delta k'$, the phase synchrony is great and close to one, otherwise close to zero. A bootstrap based test was used in [94] to test if the phase locking value is equal to one. This method is preferred to simple thresholding method. However bootstrap method is time consuming and for certain applications is not convenient.

The phase difference can also be estimated based on Hilbert transform under certain conditions [102] as explained below.

Hilbert phase Synchrony [103]

If we model real signal $\mathbf{x}_i$ as:

$$x_i[k] = A_i[k]\cos\phi_i[k] \tag{3.18}$$

then

$$x_i[k] + iH(x_i[k]) = A_i[k]\cos\phi_i[k] + iH(A_i[k]\cos\phi_i[k]) \tag{3.19}$$

where $H(\cdot)$ denotes the Hilbert transform. Using the Bedrosian theorem [104, 105], it can be deduced that:

$$H(A_i[k]\cos\phi_i[k]) = A_i[k]H(\cos\phi_i[k]) \tag{3.20}$$

if the spectral supports of $A_i[k]$ and $\cos\phi_i[k]$ do not overlap. For a narrow-band signal $\mathbf{x}_i$ which its spectral support is limited in a narrow range centered on the main frequency of its spectra (not necessarily single frequency signal), the following condition was verified [106]:

$$A_i[k]\cos\phi_i[k] + iA_i[k]H(\cos\phi_i[k]) = A_i[k]\cos\phi_i[k] + iA_i[k]\sin\phi_i[k], \tag{3.21}$$

which leads to the following model [102, 106–108] as:

$$x_i[k] + iH(x_i[k]) = A_i[k]\exp(i\phi_i[k]) \tag{3.22}$$

For a narrow-band signal, the relative variation of amplitude can be assumed to be much slower than the variation of the phase which implies the following condition [109]:

$$\left|\frac{d\phi_i[k]}{dk}\right| \gg \left|\frac{1}{A_i[k]}\frac{dA_i[k]}{dk}\right|. \tag{3.23}$$

The real signal $\mathbf{x}_i$ can be modeled by the pair $(\mathbf{A}_i, \boldsymbol{\phi}_i)$ using Hilbert transform (3.22) if the condition (3.23) is satisfied. The better this condition being satisfied the more likely the pair $(\mathbf{A}_i, \boldsymbol{\phi}_i)$ obtained with the Hilbert transform gives an accurate model for real signal $\mathbf{x}_i$ [102].

For the narrow band real signals which provide condition (3.23), the phase difference based on Hilbert transform can be defined as:

$$\widehat{\Delta\boldsymbol{\phi}_H}(\mathbf{x}_i, \mathbf{x}_j) = \boldsymbol{\phi}_j - \boldsymbol{\phi}_i. \tag{3.24}$$

Comparison of cross-correlation, coherence and phase synchrony

Coherence is bounded between 0 and 1, therefore it cannot distinguish between correlated and anti-correlated signal pairs opposed to cross correlation ($[-1, 1]$) which can differentiate between correlated and anti-correlated relationships. The maximum of wavelet cross-correlation coefficient reflects the same concept of wavelet coherence. However complex wavelet is used in wavelet coherence while in wavelet cross-correlation the wavelet is real. Although using different mother wavelets, still the linear relationship between wavelet coherence and maximum wavelet cross-correlation is conserved, but the value of each measure is different.

It is shown in [94] that wavelet coherence (3.13) increases in both following conditions: 1) if the cross-correlation between envelopes of the signal pairs increases and 2) if the phase difference between signal pairs has small variations in time. Therefore coherence does not specifically characterize the phase-relationships, whereas the phase synchrony (3.17) is only sensitive to the phase difference variations [94].

Mutual information (MI)

Let's assume time series $\mathbf{x}_i = [x_i[1] \ \dots \ x_i[T]]^T$ and $\mathbf{x}_j = [x_j[1] \ \dots \ x_j[T]]^T$ as the observations of two random variables X_i and X_j. The MI [110] between $\mathbf{x}_i$ and $\mathbf{x}_j$ is estimated as:

$$MI(\mathbf{x}_i, \mathbf{x}_j) = \sum_{mn} \hat{p}_{ij}(m,n) log \frac{\hat{p}_{ij}(m,n)}{\hat{p}_i(m)\hat{p}_j(n)} \tag{3.25}$$

where $\hat{p}_i(m)$, $\hat{p}_j(n)$, and $\hat{p}_{ij}(m,n)$ are marginal and joint estimated probabilities of $\mathbf{x}_i$ and $\mathbf{x}_j$. Previous coupling measures except phase synchrony are designed to measure the linear relationship between signal pairs, while using MI, the nonlinear relationships between signal pairs can also be measured. However for Gaussian signals, MI can be approximated as a nonlinear function of correlation coefficient.

3.8.2 Directional couplings

Using directional coupling measures like Granger causality [111,112] the drive-response relationship between signal pairs can be estimated. The causal relationship can also be estimated by time delay estimation between signal pairs observed at the nodes. Here we briefly review some of time delay estimation methods [58–61] and Granger causality [111,112] which are widely used in this domain.

Time delay estimation

Direction of a connection can be estimated from the estimation of the time delay between the signal pairs observed at the two ends of a connection [61]. There are several studies on the estimation of the time delay. Here we focus on two methods:

1. *Time delay estimation based on phase difference [61]*
2. *Time delay estimation based on cross-correlation [58–60].*

Here we review briefly these two methods.

1) Time delay estimation based on phase difference: The phase difference between signal pairs is in the range $[-\pi, +\pi]$. Each phase difference value can be added to $\pm 2k\pi$ which leads to different time delays assuming the following linear relationship for single frequency signals:

$$\Delta\phi = \frac{2\pi f}{f_s}\Delta\tau \tag{3.26}$$

where f, $\Delta\phi$, and $\Delta\tau$ refer to frequency, phase difference, and time delay, respectively. There is an ambiguity in choosing the correct time delay. This ambiguity is worse for higher frequencies. When the possible time delays are smaller than the maximum physiological time delay, even by using physiological constraint, we cannot reject any of them. Gotman [61] studied the phase difference in terms of frequency using cross-spectra density function (based on Fourier transform). To solve this ambiguity of time delay estimation, he proposed considering the slope of

$\Delta\phi$ over a frequency interval instead of considering single $\Delta\phi$ at each frequency bin. By choosing narrow frequency intervals (Δf) or wide time periods, $\Delta\phi$ may associate to time delays with large differences in their values. Hence the greater time delays can safely be rejected according to physiological constraints like the physiological time delay between different brain regions recorded from separate iEEG recordings. Moreover by choosing the slope of $\Delta\phi$, the difference between the choices of time delays does not depend on the frequencies used in frequency interval measurement and instead depends on only the length of the frequency interval.

Gotman [61] used the estimated time delays to interpret the causal relationships between signal pairs from iEEG recordings of epileptic patients. Ktonas et al. [113] followed this method and studied the statistical errors of this time delay estimation.

This method of time delay estimation may not be applied on phase difference methods based on wavelet transform. In wavelet transform the phase difference can be calculated for each frequency interval, i.e. a single phase difference is obtained for each frequency interval. In other words, we do not know the values of phase difference at the two single frequencies of a frequency interval to calculate the slope of $\Delta\phi$. Furthermore the variances of wavelet coefficients of different frequency bands are not generally the same. Wavelet transform has much more interests than Fourier transform which one may prefer to stick with wavelet and using another measure to calculate the time delay.

2) Time delay estimation based on cross-correlation: The estimation of time delay between signal pairs using cross-correlation was proposed in [58–60,91]. In most of these studies the following linear model was assumed:

$$\begin{aligned} x_1[k] &= s[k] + n_1[k] \\ x_2[k] &= s[k + \Delta\tau] + n_2[k] \end{aligned} \tag{3.27}$$

where $s[k]$, $n_1[k]$ and $n_2[k]$ are the realizations of uncorrelated Gaussian random process and $\Delta\tau$ is the time delay. In these studies, the cross-correlation is calculated for various time lags and the time delay was estimated as the time lag in which the absolute maximum of cross-correlation occurs. A simple simulation is demonstrated in Figure 3.2 to show how the cross-correlation estimates the time difference between two events in a signal pair.

Granger causality

An intuitive definition of causality between two temporal signals was introduced by Wiener [111] and formalized by Granger [112]. The Granger causality concept for two processes X_1 and X_2 is as follows.

If the knowledge of the past of both X_1 and X_2 reduces the variance of the prediction error of X_2 more than using only the knowledge of the past of X_2 then X_1 causes X_2. The linear Granger causality test [114] and its extended measures like direct coherence [57], and partial directed coherence [115] are based on a multivariate coupled autoregressive model. In linear Granger causality [114], the causality test is based on the variance of the prediction error in the

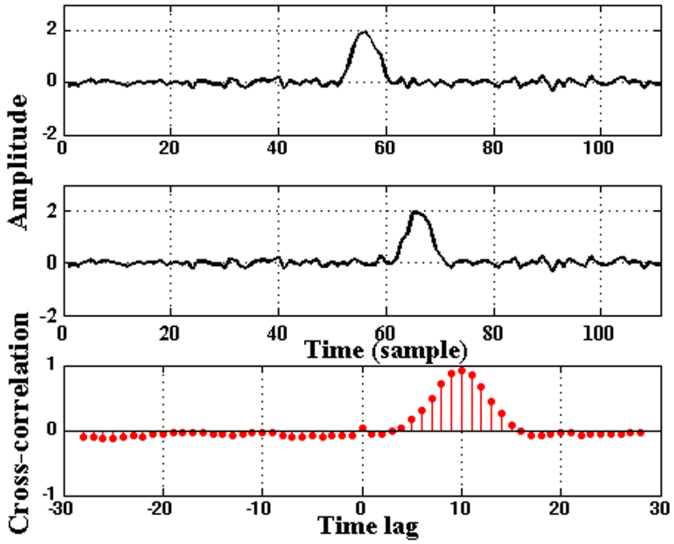


Figure 3.2: Time delay estimation based on cross-correlation coefficient. Time delay is the time lag which the absolute cross-correlation coefficient maximizes.

multivariate coupled autoregressive model. The direct coherence [57] shares the same model of linear Granger causality except an extra noise is added to the bivariate autoregressive model allowing noises to be uncorrelated. In the same perspective, the partial directed coherence [115] is based on Fourier transform of multivariate autoregressive model used in linear Granger causality.

The partial nonlinear Granger causality [116] uses the same concept of linear Granger causality in the context of probability, i.e. X_1 is not a nonlinear Granger cause of X_2 if the probability distribution of the future of X_2 does not change by adding the knowledge of the past of X_1.

Gourévitch et al. [117] studied different measures of Granger causality and extended linear Granger causality and nonlinear partial Granger causality to multivariate case. In [117], different Granger causality measures are compared using simulated and real scalp EEG recordings, concluding the following remarks:

1. Direct coherence [57] and partial direct coherence [115] measures do not give similar results and should be carefully interpreted. These two measures have good results for linear linkages, while they do not provide reliable results to some simple nonlinear linkages. If nonlinearity is suspected, partial nonlinear Granger causality [116] can only be used as a complementary method to for example partial directed coherence measure.

2. Although partial nonlinear Granger causality can provide good results for complex systems, it is highly dependent on the selected parameters.

There are other directed measures based on nonlinear methods like estimating the time delay using nonlinear correlation coefficient [83, 118, 119], generalized synchrony [120–124], transfer entropy [125], structural equation modeling [51], and dynamic causal modeling [52–56], which we do not go through these measures.

3.9 Conclusion

In this chapter, we explained about the directed graph, the characterization of its nodes using different graph measures, and different non-directional and directional coupling measures used for graph inference.

Among non-directional coupling measures, cross-correlation and coherence can measure the linear relationship between signal pairs, while phase synchrony and mutual information are useful when we suspect non-linear relationships between signal pairs. Several studies indicate no preference of nonlinear coupling measures over linear ones during interictal period of EEG analysis [8,10,100,126,127]. Therefore, in this work we use maximum wavelet correlation coefficient to measure the functional connectivity. This measure is calculated between signal pairs during IED and non-IED time intervals (for details refer to section 6.2.2, page 56). Once the relationship between different brain regions are inferred, we estimate the directions of these relationships. To this end, the time causality between IEDs recorded from iEEG channels is estimated using the same measure (maximum wavelet correlation coefficient) between signal pairs during the whole selected data for processing (see section 6.3.2, page 61).

Chapter 4

Identification of leading IED regions from invasive EEG

The aim of surgery of drug resistant focal epileptic patients is to remove the EZ for a seizure free outcome [14, 30]. However EZ cannot be determined before surgery by the current recordings and techniques. Therefore, practically we have to estimate the EZ by the estimation of other dysfunctional brain regions like SOZ or IZ. Usually SOZs are estimated using ictal state analysis from iEEG for pre-surgical evaluations [1,33,34]. Recording ictal EEG is time-consuming and expensive, because seizures are infrequent phenomena in most of the patients, thus EEG recording period may be prolonged (over few weeks) till a seizure occurs. This long recording period may increase the risk of complications like brain hemorrhage and infection [128–131]. Eventually the number of seizures are limited and each seizure period is short. Thus obtaining statistically reliable results from studying seizure time intervals is quite difficult [22]. Consequently, suggestion of SOZ through interictal analysis remains valuable.

In this chapter, the state of the art in identification of IZ or IED regions based on interictal discharge analysis from invasive EEG recordings is reviewed. Due to the rather old history of the EEG analysis of epileptic patients and the rich literature in this field, we focus on the works specifically developed for the problem of interest.

There are several studies wondering if quantitative analysis of interictal discharges can be used to guide the resection surgical decision [8–10, 12, 13, 22–24, 132–134]. There are two main methods:

1. *Univariate methods*, i.e. the interactions between the brain regions are not considered.
2. *Multivariate methods*, i.e. the couplings between signal pairs related to different brain regions are investigated.

4.1 Univariate methods

In [132], Alarcon et al. reported a complex and definite correlation between interictal and ictal activities according to their univariate analysis from ECoG recordings. They used univariate

indices to measure the latency, amplitude, and frequency of spikes from fast Fourier transform (FFT) of the ECoG signal during ictal and interictal stages and studied the relationship between resection of interictal discharge regions and post-surgical outcome.

The main problem of SOZ detection through interictal analysis is the propagation of interictal discharges [22, 135]. IEDs may start from a complex relationship between separate zones, following the propagation of neuronal activity through specific neural pathways. By studying leading epileptic regions, where interictal discharges demonstrate earlier peaks, Alarcon et al. reported that these regions may have higher epileptogenicity potentials rather than the sites with secondary propagated activity [22]. This may suggest the unnecessity of resecting all of IED regions to control the seizures, while removing leading regions may be helpful to guide a successful surgery. Following pioneer works of [22, 132], univariate measures such as earliest spikes, highest average amplitude, and shortest spike duration were studied for localization of IED regions [23]. In [23] a greater overlap between the averaged earliest spikes localizations and the site of seizure origin comparing to highest averaged spike amplitude localizations was reported.

There are other studies based on the estimation of the occurrence time of the interictal discharges to identify the leading IED regions [133, 136, 137]. In [133] the occurrence time of spikes from ECoG recordings was used to determine an activation map of different ECoG channels. To estimate the occurrence time of spikes, they tested four different measures including the occurrence time of peak amplitude of the temporal ECoG recordings, the peak of the first, second and Laplacian derivatives of the temporal ECoG recordings. They concluded that the time occurrence of the peak amplitude of the ECoG recordings is the best measure to determine the occurrence time of spikes. Based on the estimated time occurrence of spikes, the leading IED regions were identified as the regions which have the shortest latency. They concluded a strong correlation between leading IED regions and SOZ.

These studies confirmed that identification of regions which originally originate the interictal discharges can be valuable for the estimation of SOZ and presurgical evaluations.

4.2 Multivariate methods

Following to univariate methods, multivariate methods are used in this domain to add the information of the inter-relationships between different IED regions. Each EEG channel can be considered as a node and the relationship between nodes can be identified by edges. Here, we try to summarize the multivariate methods based on connectivity graph in the domain of studying the relationship between IED regions and SOZ [8–10, 12, 13, 24]. The schematic block diagram of these methods is depicted in Figure 4.1. The following steps are considered.

1. Step 1: *Coupling measure*
2. Step 2: *Graph inference:* this step includes two sub-steps:
   (a) *thresholding*

(b) *multiple testing*

3. Step 3: *IED identification*

In the following we explain these steps.

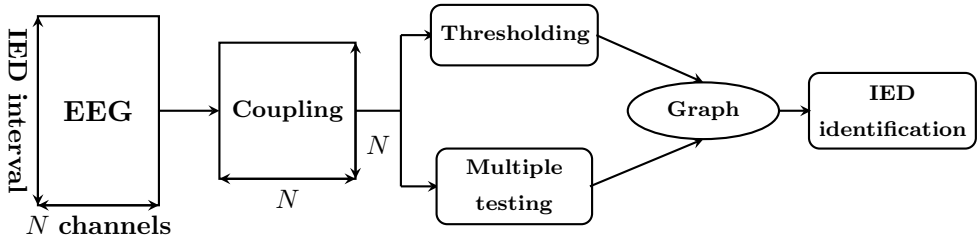


Figure 4.1: The schematic block diagram of multivariate methods in studying the relationship between the interictal brain regions and seizure onset zones through connectivity graph.

4.2.1 Step 1: Coupling measure

In multivariate methods, first interictal couplings between multivariate N-dimensional preprocessed EEG recordings are calculated based on linear, non-linear, or directed interdependencies giving a N by N coupling matrix. The preprocessing can include filtering, FFT, or wavelet transform. The coupling matrix, is a square matrix containing the coupling measures calculated between N channels. Some of widely used linear and nonlinear coupling measures [82–84] are explained in section 3.8, page 24.

4.2.2 Step 2: Graph inference

For inferring the graph from the coupling measures (Figure 4.1) or to define the associations between EEG channels (nodes), we explain two mostly used methods as follows.

(a) *Thresholding*

(b) *Multiple testing*

Thresholding

The classic method compares the coupling strength with a threshold value [18, 19, 138–140]. If the coupling strength is greater than the threshold, an edge is inserted between underlying nodes.

There are several studies following this procedure using different threshold selection strategies. One strategy is to choose a threshold which provides a fixed number of edges [138]. Another strategy is to apply a variety of thresholds to obtain a collection of graphs and extract interested feature(s) from these graphs. In these methods either the threshold was selected as the threshold in which the extremum of feature values occurs or the extracted features in terms of different

thresholds were studied [138, 141]. Kramer et al. [11, 19] tested latter procedure and found it time consuming and unsatisfactory.

Indeed threshold based graph inference has several problems. First the inferred graph is crucially dependent on the value of the threshold and the threshold may vary for different time intervals of processing the data. For example in dynamic study of graphs which appearing or disappearing of edges is important, the graphs of different time intervals are compared. Comparing these graphs which were inferred with different thresholds may not provide correct comparison. The same problem exists for comparing the graphs of different patients which their related thresholds may not be the same. Second in such inferred graphs the rate of falsely detected edges is unknown and there is no measure of graph uncertainty [11]. Therefore graph inference methods based on significance tests seem more interesting. We explain these methods in the next section.

Multiple testing

The necessary background about multiple testing is explained in appendix F, page 147. In graph inference methods based on multiple testing [20, 21, 142, 143], usually the following procedure is examined:

1. The statistical hypothesis and the test statistic in terms of the selected coupling measure are defined e.g. scaled maximum cross-correlation coefficient to its variance as the test statistic.

2. In the classic methods, a predefined distribution is assumed for the test statistic. For example normal distribution is assumed for the scaled maximum cross-correlation coefficient test statistic.

3. Upon this assumption the raw p-values are calculated for all of the tests or all of node pairs (N_c: the number of possible connections, see section 3.5, page 20).

4. The raw p-values are adjusted for multiple testing e.g. based on controlling familywise error rate (FWER).

5. The adjusted p-values are compared with α_{fw} (appendix F, page 147). The multiple testing including null hypothesis, i.e. H_0^n: "No coupling" versus alternative hypothesis, i.e. H_1^n: "Coupling" is done for all of the node pairs, $n = 1, \ldots, N_c$. For the test in which its adjusted p-value is smaller than α_{fw}, the null hypothesis is rejected and an edge is inserted between the related node pairs.

In the graph inference methods based on multiple testing, the expected rate of falsely detected edges can be declared if the p-values (raw and adjusted) are correctly estimated. Furthermore a fixed α_{fw} can be set for different time interval graphs, contrary to the threshold used in threshold-based graph inference methods. However if the assumption for the distribution of the test statistic (the second step of the above procedure) is not correct, then the raw p-values

cannot be estimated correctly. Hence incorrect p-values may not provide a correct significance test [11]. Besides for non-stationary invasive EEG recordings, the estimation of the test-statistic distribution is complicated. Therefore we need a method which requires less assumptions about the data. Resampling methods like permutations provide a suitable choice for this purpose as they need less assumptions about the underlying data allowing reliable significance test [144,145].

In several contexts graph inference methods based on multiple testing were used [11, 12, 21, 142,143]. Kramer et al. [11] compared a bootstrap-based multiple testing with two other multiple testing methods:

1. Multiple testing based on the predefined statistical distribution of the test statistic. In the following we refer to this method as naive method. The normal distribution was assumed for the test statistic of scaled maximum cross-correlation to its variance.

2. A new test-statistic in terms of maximum cross-correlation was proposed which was declared as a better approximation of normal distribution. This new test statistic was the normalized maximum of the Fisher transform of the cross-correlation.

Here we first explain the bootstrap based method used in [11] and then we explain about the comparison of this method with the above multiple testing methods reported in [11]. For the bootstrap based method, the temporal signals were resampled with replacement, preserving the spectral characteristics of the original signals. For each signal, the resampling was repeated N_b times to construct N_b resampled signals in the surrogate data. Therefore, for each original signal pair (i, j) there were N_b resampled signal pairs. For these N_b signal pairs, the test statistics were calculated providing N_b test statistics which formed the test-statistc distribution of signal pair (i, j). Using this distribution and the test statistic of original signal pair, the p-value of signal pair (i, j) was calculated. For N channels which gives $N_c = N \times (N-1)/2$ signal pairs, $N_b \times N_c$ bootstrap distributions were calculated.

To compare the naive method based on normal distribution, and the new test-statistic with the bootstrap-based method, the simulated and ECoG data were used in [11]. The new test-statistic and bootstrap method were reported to provide much better results than naive method. However their new test-statistic can be used only for maximum cross-correlation which the analytical expressions could be derived.

The bootstrap method used in [11] is very time consuming and the processing time is dependent on the length of selected data for processing. For one second of ECoG processing of $N = 97$ channels, the bootstrap-based method took 90 minutes and for 30 seconds, 45 hours (on a 2 GHz Core Duo processor). This method is not suitable for long length data processing which is crucial for reliable statistical results.

4.2.3 Step 3: IED identification

So far the graph was inferred from the coupling measure matrix and now the lED regions should be identified from the graph analysis. IED regions can be identified based on quantifying the nodes of the graph using graph measures. Some of classic graph measures quantifying the

structure of graph were reviewed in section 3.7, page 22. Other graph measures for the purpose of IED region identification are suggested. Ortega et al. [24] quantified each node i as the sum of coupling measure values between node i and its first neighbors in the grid of ECoG recording denoted as synchronization power of node i. The IED regions were identified as the cluster of nodes whose synchronization powers were greater than a threshold. Three coupling measures: cross-correlation coefficient, phase locking value based on Hilbert phase synchrony and mutual information were used in this study.

Another survey of similar strategy was proposed in [12]. Each node was quantified as the sum of its outgoing directed coupling measure values denoted as causal source activity. The nodes which have causal source activities greater than a threshold were chosen. The direct transfer function [146, 147] was used as the directed coupling measure. In the same strategy, Monto et al. [9] studied the detrended fluctuation analysis exponent of autocorrelation function (a self-similarity measure) of each ECoG channel. The nodes which have greater values were selected.

The problem of such studies is that they are threshold based. Furthermore in case of studying different frequency bands, for each frequency band a separate set of IED regions is obtained while the preference between different frequency bands is usually unknown. For the case of considering different frequency bands, one can consider different classic methods like clustering [148] (appendix A, page 123) for identification of IED regions from graph measures related to different frequency bands. Such methods and their problems are discussed in section 7.3, page 69.

4.3 Conclusion

We explained the problems of available multivariate methods on identification of IED regions through connectivity graphs. Here we summarize these problems and briefly present our proposed solutions.

4.3.1 Problems

1. Available methods analyze either IED or non-IED time intervals separately. The interpretation of such graphs is complicated as they include background activity and related volume conduction effect connections as well as interested connections related to IEDs.

2. Most of the graph inference methods used do not provide powerful significance tests.

3. Usual graph measures do not include enough information for quantification of graphs for leading IED regions identification. To identify the leading IED nodes, we need to quantify the strength of source and sink nodes.

4. Current methods for IED region identification provide a set of IED regions for each individual frequency band, while the preference between different frequency bands is unknown.

4.3.2 Proposed solutions

Toward solving the above problems, we proposed the following ideas, which are the main contributions of our work:

1. Proposing a new graph termed as differential connectivity graph (DCG) using both IED and non-IED time intervals. DCG includes significant discriminated connections between IED and non-IED states.

2. Application of permutation-based multiple testing [145] for a statistically confident graph inference in the domain studied.

3. Introducing a graph measure for characterizing the local information of the nodes of digraphs using lagged mutual information.

4. Proposing a multiple graph analysis to extract the information from the graph measure values of multiple graphs using multi-objective optimization method [149, 150].

In the following part, first the basic ideas of DCG and IED regions identification are explained in Chapter 5. Next, contributions 1 and 2 are described in Chapter 6 and finally Chapter 7 includes the explanations of contributions 3 and 4.

Part II

Proposed method

Chapter 5

Basic idea of differential connectivity and identification of leading epileptic regions

5.1 Basic idea of differential connectivity graph

Transient neuronal assemblies are referred to large group of distributed neurons becoming abruptly associated or disassociated [151]. Specific neuronal assemblies can emerge the functional elements of brain activity that execute the basic operations of information processing [152]. The construction of a neuronal assemble is along with an increase of total potential (excitatory/inhibitory PSP), while its destruction is described by the decrease of total potential [153]. These oscillations of total potential can be manifested in the EEG recordings in which temporary stable local microstates in the brain operations can lead EEG segments or piecewise "explicit" state EEG time intervals [37, 65, 154, 155]. An EEG segment may refer to e.g. alpha activity, interictal epileptiform activities, etc. The transition from one EEG segment to another, reflects the changes in the neuronal assembly state and/or in the neuronal assembly itself [65]. These transitions may change the functional brain connectivity. In this research work, assuming that functional connectivity changes during the transitions between the two explicit states, a new graph is proposed to demonstrate the differences of the connectivity between the two states. We proposed a *differential connectivity graph (DCG)* which is designed to detect significantly these discriminated connections. Although the method is general, in this thesis we focus on the application of epileptic patients and the two IED and non-IED states. In the following we explain the related problems in the DCG construction.

5.2 Problem statement

Identification of discriminated connections from two separated IED and non-IED connectivity graphs is challenging. Let's infer the graph related to one IED time interval with simplest method which is thresholding the coupling matrix [18]. For example for an IED time interval of length

2693 samples ($f_s = 512$ Hz) in the frequency range 4-8 Hz, the graph is demonstrated in Figure 5.1(a) from $N = 111$ iEEG bipolar channels of a focal epileptic patient. The graph of a non-IED time interval of length 4969 samples ($f_s = 512$ Hz) for the same patient, frequency range and threshold value is depicted in Figure 5.1(b). The threshold is set equal to 0.3. Different values of threshold will change the number of edges in these graphs [11] as it can be seen in comparing Figures 5.2 (threshold=0.5) and 5.1 (threshold=0.3). However these graphs are very complicated to be interpreted. Furthermore each IED related graph has some variability (in the set of connections) in time [13, 19]. This is true for non-IED related graphs as well. To demonstrate these time variability which is more in non-IED graphs since they reflect the background activity, the IED and non-IED graphs of two other time intervals of the same patient and frequency band are shown in Figure 5.3(a) and 5.3(b), respectively.

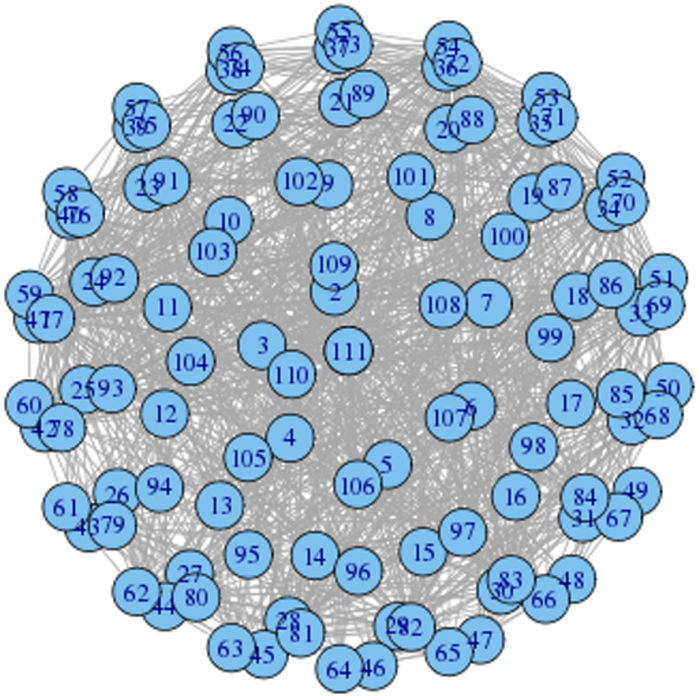
So if we construct the IED and non-IED related graphs of large number of IED and non-IED time intervals, we obtain two collections of graphs (IED and non-IED) with some time variability in each collection. The aim is to extract the most reliable connections which significantly change between these two collections of graphs. In other words the aim is to find the couplings with low variance (in time) in each collection and large mean difference between two collections. The variance of the couplings for different time intervals depends on the estimation error of the coupling of each time interval and also depends on the non-stationarity of the couplings among different time intervals. Controlling the confidence interval of the couplings for each time interval is challenging. Because the data is non-stationary so modeling the distribution of the test statistic of each time interval is complicated. Available methods for calculating the confidence interval of each coupling e.g. wavelet cross-correlation [89] are very conservative and propose very long time intervals for a reasonable confidence interval for low frequency bands which is not in accordance with real practical problems.

A solution can be using resampling based multiple testing methods [145] which provide powerful significance test. In these methods, the distributions of the test statistics are estimated under null hypothesis assuming less assumptions about the data. Using permutation test, the most reliable connections differing between IED and non-IED couplings can be detected while controlling the error rate (FWER or FDR).

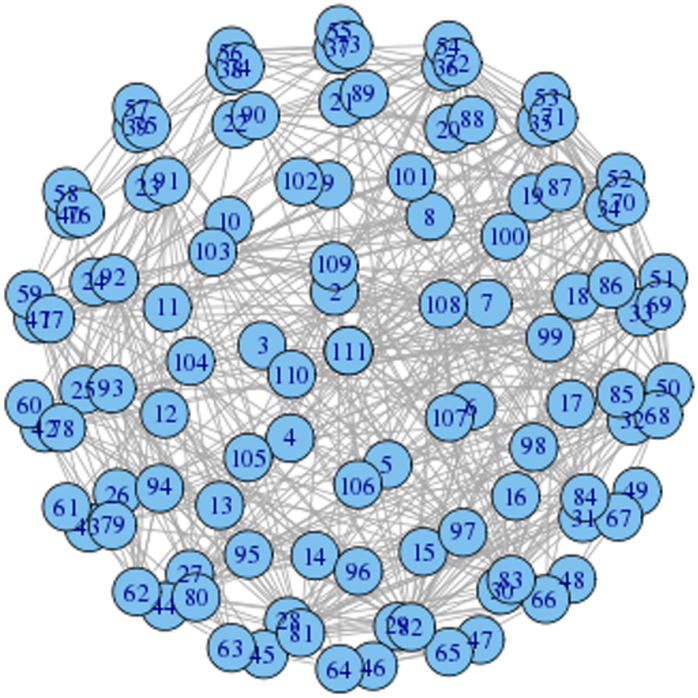
Here we summarize how DCG solves the above problems.

1. To identify the discriminated connections between two brain states instead of comparing two collections of graphs, we focus on the connections whose coupling measures change significantly between different IED and non-IED time intervals.

2. Instead of analyzing each time interval separately which gives a complicated graph, DCG searches for statistically significant connections among large number of IED and non-IED time intervals giving sparser graph which can be interpreted more easily. The DCG of the same patient and the same frequency band of the graphs of Figures 5.1-5.3 is depicted in Figure 5.4. In part (a) the same sphere layout as Figures 5.1-5.3 is used for a better comparison. This graph is depicted in part (b) in a different layout for better representation excluding disconnected nodes.

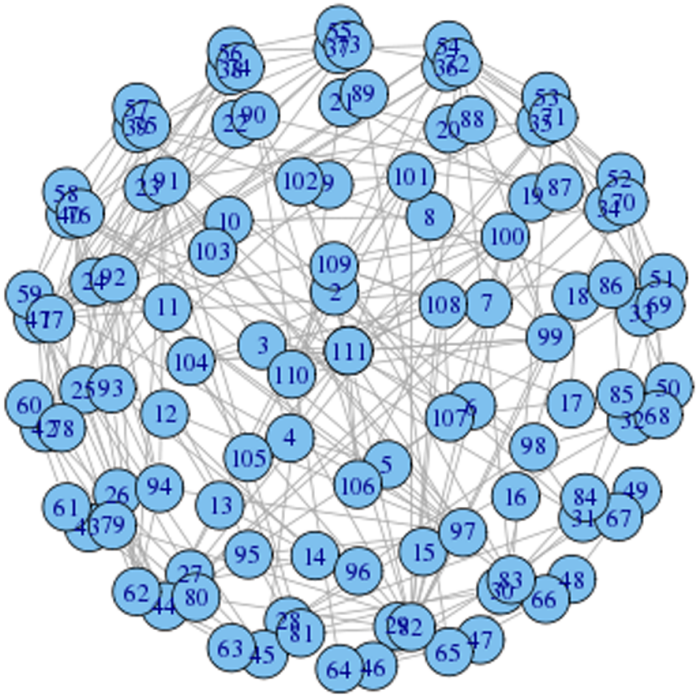


(a) IED

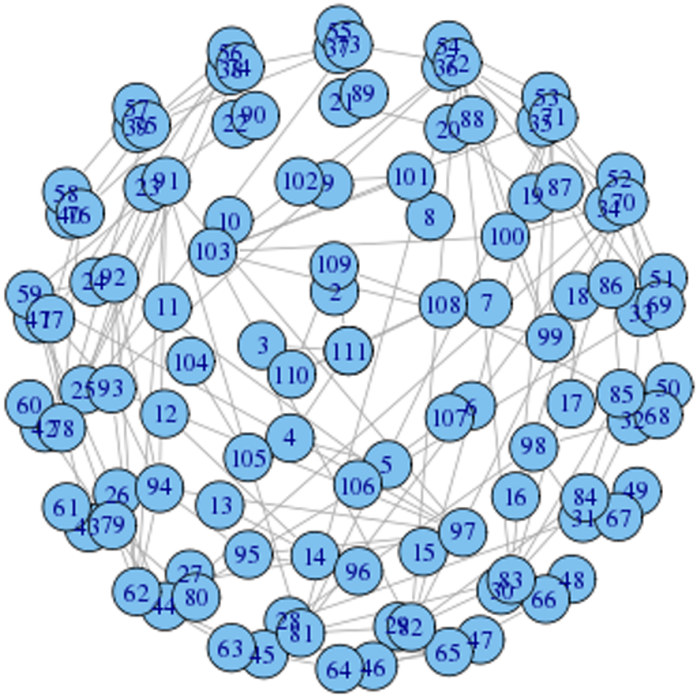


(b) Non-IED

Figure 5.1: Graph related to one (a) IED time interval of length 2693 samples and (b) non-IED time interval of length 4969 samples in 4-8 HZ from $N = 111$ iEEG bipolar channels of a focal epileptic patient. The sampling rate, f_s is equal to 512 Hz. The graph is inferred by simple thresholding for threshold equal to 0.3.

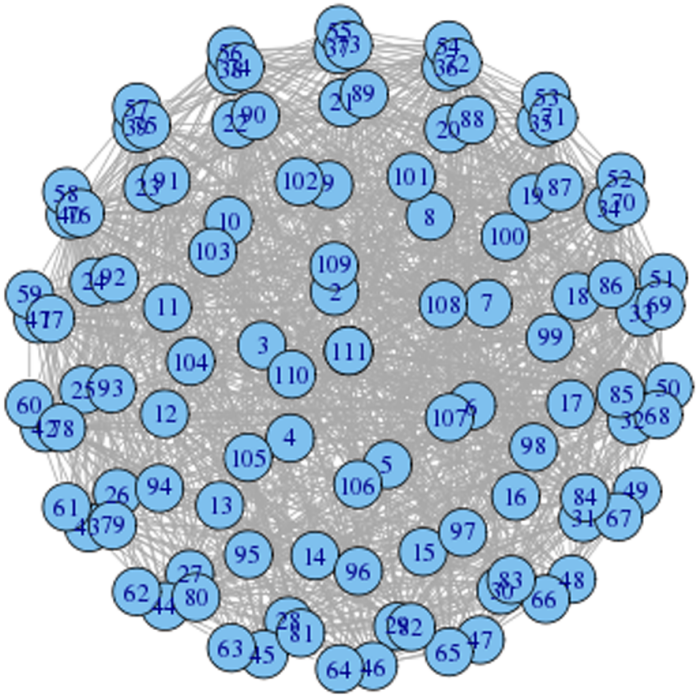


(a) IED

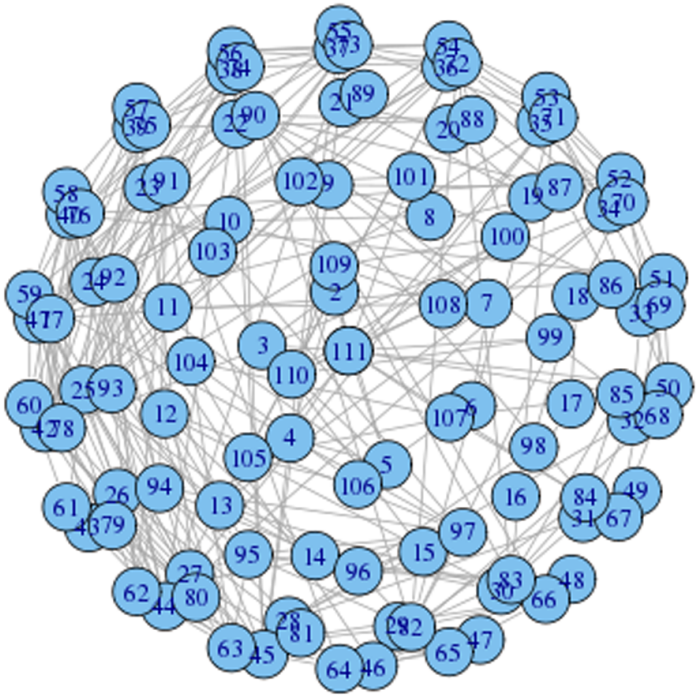


(b) Non-IED

Figure 5.2: The graphs of the same (a) IED and (b) non-IED time intervals used in Figure 5.1 for a different threshold value equal to 0.5.



(a) IED



(b) Non-IED

Figure 5.3: Graph related to one (a) IED time interval of length 2311 samples and (b) non-IED time interval of length 7023 samples in 4-8 HZ from $N = 111$ iEEG bipolar channels of a focal epileptic patient. The sampling rate f_s, is equal to 512 Hz. The graph is inferred by simple thresholding for threshold equal to 0.3.

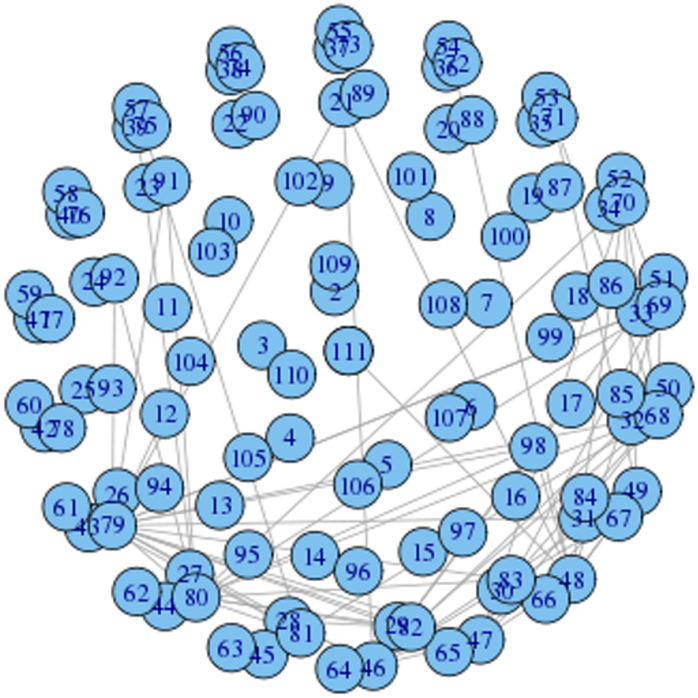
3. By using permutation-based multiple testing [145], the distributions of the test statistics are estimated under null hypothesis assuming less assumptions about the data with the controlled error rate.

5.3 Schematic diagram of the main blocks of the proposed iEEG analysis method

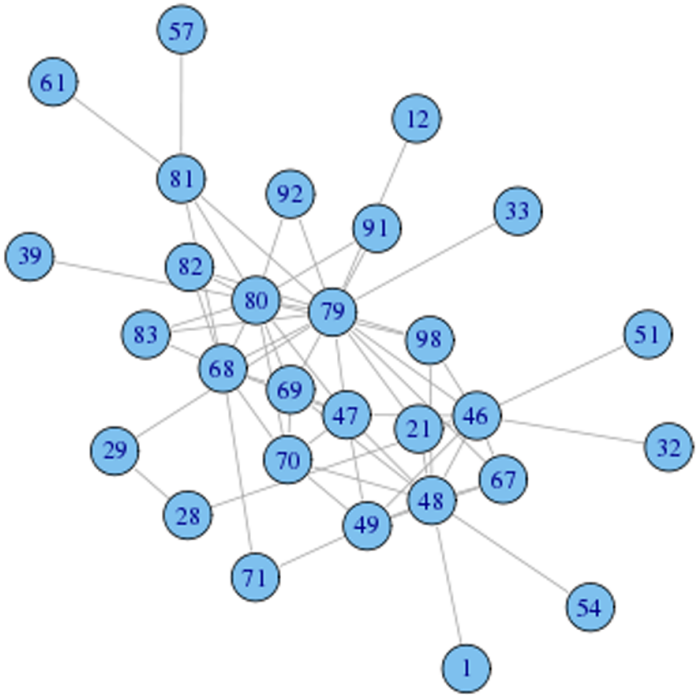
So far the basic idea of DCG is explained. Once the DCG is obtained, different graph measures can be applied to characterize the nodes of the graph for the purpose of IED region identification. The coarse and fine block diagrams of the proposed method are brought in Figure 5.5. The IED and non-IED onsets and offsets are determined visually by the epileptologist from iEEG recordings. Using these labels the wavelet coefficients of temporal iEEG signals related to each frequency level j, $j = 1, \ldots, J$ are segmented to IED and non-IED segments. Based on these segments, J DCGs are constructed for J frequency levels. These J DCGs are characterized providing J-dimensional measure values for all of the nodes. The leading IED (ℓIED) regions are estimated by multiple graph analysis using multi-objective optimization method based on these measure values.

5.4 Conclusion

In this chapter we explained the basic idea of DCG which included the existing problems of the available methods and the proposed solutions. Furthermore, the main blocks of the proposed iEEG analysis method are introduced. In the next chapter we explain how to calculate the DCG.



(a)



(b)

Figure 5.4: DCG of a focal epileptic patient (the same patient of Figures 5.1-5.3) in 4-8 HZ for two different layout representations in parts (a) and (b). The layout representation of part (a) and Figures 5.1-5.3 are the same for a better comparison.

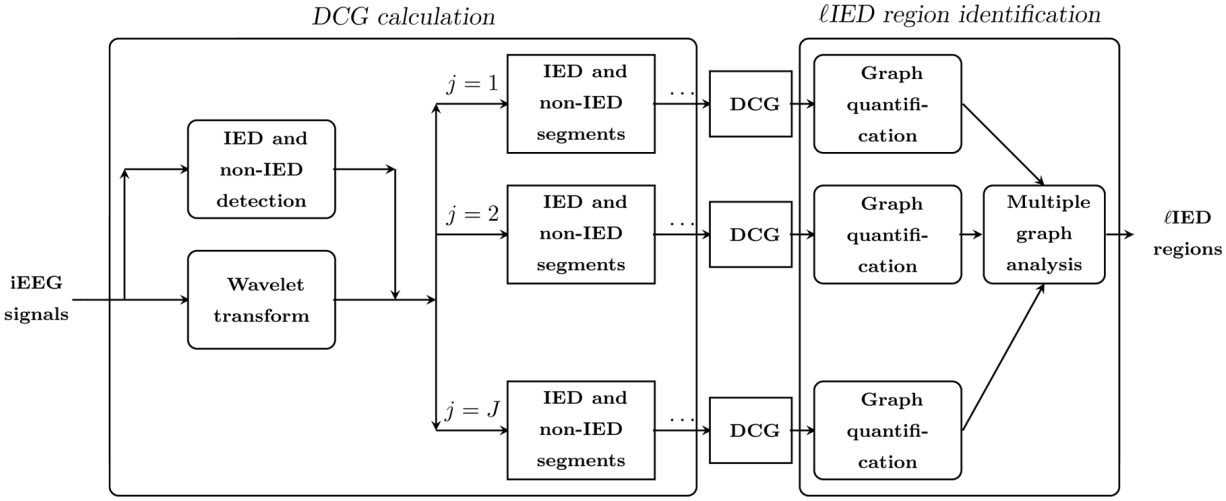


Figure 5.5: block representation of the main steps of the proposed method for iEEG analysis. First, the IED and non-IED onsets and offsets are identified visually from iEEG recordings by epileptologist. The wavelet coefficients are segmented according to these IED and non-IED time intervals. Next DCG related to each frequency level j, $j = 1, \ldots, J$ is constructed based on IED and non-IED wavelet coefficients. Please notice that all of the sub-blocks of DCG calculation are not demonstrated in this block diagram. For more details about the DCG calculation refer to the next chapter. Finally, the J DCGs are characterized and the ℓIED regions are estimated by multiple graph analysis based on these characterized measure values.

Chapter 6

Differential connectivity graph (DCG)

6.1 Introduction

The main problem of available graphs for interictal analysis of iEEG recordings is lack of reliability (for details see section 5.2, page 43). The other important problem is that classic graphs are based on the analysis of only IED time intervals. The interpretation of such graphs is difficult since the connections related to unwanted events like background activity is included as well as connections related to IED events. In this chapter we explain how to analyze both IED and non-IED time intervals using permutation-based multiple testing [145] for solving these problems.

This chapter is organized as follows. First the different steps of DCG calculation is explained. Then we describe the method that is used to estimate the directions of the edges of DCG to provide directed DCG (dDCG).

6.2 DCG calculation

DCG calculation consists of three stages:

1. IED and non-IED segmentation
2. Couplings computation
3. DCG inference

The diagram of the proposed stages are demonstrated in Figure 6.1.

6.2.1 IED and non-IED segmentation

The diagram depicted in Figure 6.2 summarizes the procedure of IED and non-IED segmentation.

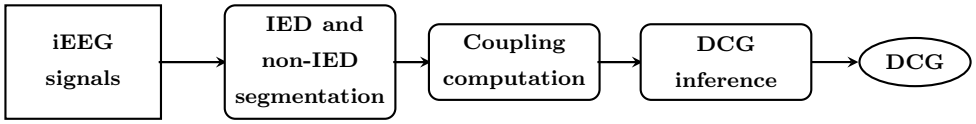


Figure 6.1: Diagram of DCG calculation.

IED and non-IED detection

The IED and non-IED time intervals are IED and non-IED onsets and offsets identified from the time series of iEEG recordings. This is done manually by the epileptologist for each patient. An IED period is a non-stationary time interval including at least one IED. A non-IED period is a homogeneous time interval without any IED or abnormal event. Two typical IED and non-IED time intervals are depicted in Figure 6.3. The number of IED and non-IED periods are on average 304 and 174 per patient, respectively, which provides statistically significant results.

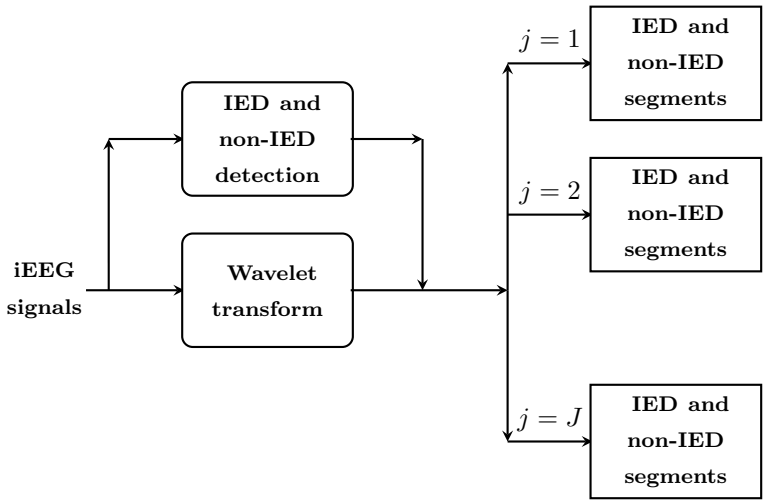


Figure 6.2: Diagram of IED and non-IED segmentation.

IED and non-IED segments

IED or non-IED segments or periods refer to the wavelet coefficients during IED or non-IED time intervals explained above. In the following we explain about the reason of using wavelet transform and the wavelet transform applied in this procedure.

Maximal overlap discrete wavelet transform The IEDs often take place within a specific frequency band and can best be determined by decomposition of the signal to the frequency band of interest through Fourier transform [61, 119], filter banks [83] or appropriate wavelet transform [89, 94, 156–158].

The average of power spectral density (PSD) of IED time intervals for all of bipolar iEEG channels are demonstrated in Figure 6.4 for a typical patient. The PSD of electrode leads of each electrode are depicted with the same color. Considering the shape of PSD of IED time intervals,

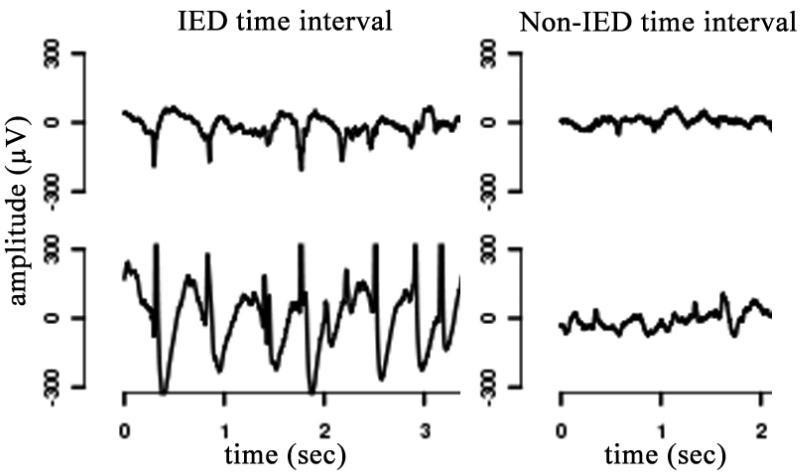


Figure 6.3: Two typical IED and non-IED time intervals from two channels of iEEG recordings.

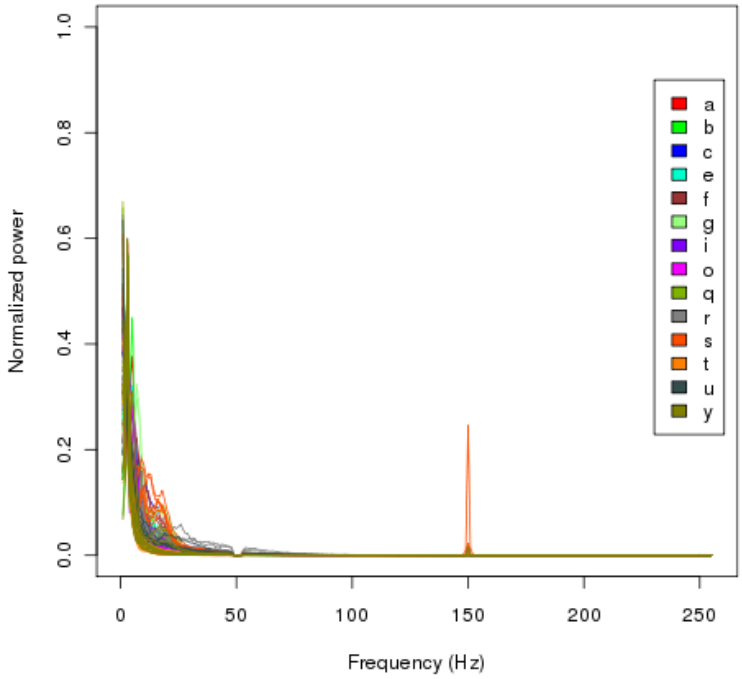


Figure 6.4: Power spectral density of IED time intervals for a typical patient.

the power is concentrated in the range [2 60] Hz. In this range the frequency contribution of IED time intervals is greater for lower frequencies. Wavelet transforms which are well-known for analyzing non-stationary EEG signals [26, 89, 92, 94, 156–164] provide automatic frequency band selection which adapts narrower bands for lower frequencies and larger bands for higher frequencies. Furthermore, mother wavelets like Daubechies [165, 166] are a proper choice for filtering IED signals due to the shape of the filter [162]. Consequently, in this work we prefer to use wavelet transforms.

Here, we use the maximal overlap discrete wavelet transform (MODWT) [89,167]. MODWT is a non-decimated version of orthonormal discrete wavelet transform (DWT) that relaxes orthogonality to gain properties like the ability to handle the time series of arbitrary sample size and translation invariance [160,167,168]. In contrast to DWT, the MODWT coefficients are associated with zero phase filters that provide consistent time-shift of MODWT coefficients with temporal signals. In MODWT, the set of coefficients in each level of decomposition includes the same number of samples as the temporal signal minus the samples affected by boundary effect[1]. The start and end time points of IED and non-IED time intervals in temporal signal are consistent with the MODWT coefficients and these coefficients can be segmented according to temporal labeling. This property does not hold for the DWT. In addition the MODWT variance estimator is asymptotically more efficient than the same estimator upon DWT [160, 167]. Nevertheless, the number of independent samples of each level is the same for DWT and MODWT but the MODWT coefficients are more redundant (due to the lack of orthogonality) than the DWT coefficients.

We assume N-dimensional observations $\mathbf{X} = [\mathbf{x}_1 \ \dots \ \mathbf{x}_N]$, where $\mathbf{x}_i = [x_i[1] \ \dots \ x_i[T]]^T$, the ith column of matrix $\mathbf{X} \in \mathbb{R}^{T\times N}$, contains T samples of the signal recorded from the ith channel, and N is the number of channels associated with bipolar iEEG electrode leads (see section 8.2, page 83). The J-level MODWT of $\mathbf{x}_i$ is $W_i^J = MODWT_J[\mathbf{x}_i] = [\mathbf{c}_i^J \quad \mathbf{d}_i^J \dots \mathbf{d}_i^1]$ ($\mathbf{c}_i^j$ and $\mathbf{d}_i^j$ are the approximate (scaling) and detailed (wavelet) coefficient sequences) such that

$$\left\|W_i^J\right\|^2 = \left\|\mathbf{c}_i^J\right\|^2 + \sum_{j=1}^{J} \left\|\mathbf{d}_i^j\right\|^2 \tag{6.1}$$

W_i^J is calculated recursively as:

$$c_i^{j+1}[k] = h^j[-k] * c_i^j[k], \quad j = 0, \dots, J-1 \tag{6.2}$$

$$d_i^{j+1}[k] = g^j[-k] * c_i^j[k], \quad j = 0, \dots, J-1 \tag{6.3}$$

where $\mathbf{h}$ and $\mathbf{g}$ are the scaling and wavelet filters of an orthonormal wavelet transform, respectively. $\mathbf{h}^0 = \mathbf{h}$, $\mathbf{g}^0 = \mathbf{g}$, $\mathbf{c}_i^0 = \mathbf{x}_i$, and * denotes convolution. The MODWT scaling filter of level

[1]The boundary effect may cause a shift difference between the events in the original time series and wavelet coefficients. To compensate this effect, the wavelet coefficients are shifted by half of the MODWT filter length at each level j ($0.5(2^j - 1)(N_f - 1) + 1$, N_f is the wavelet filter length at level $j = 1$ [89]) to have the same time instant in the temporal signal and in the wavelet coefficients. Here a periodic boundary effect is used.

$(j+1)$ is the up-sampled versions of the previous level:

$$h^{j+1}[k] = \begin{cases} h^j[\frac{k}{2}], & k \text{ even} \\ 0, & k \text{ odd} \end{cases} \tag{6.4}$$

Wavelet coefficients of level j are associated with frequency interval $[\frac{f_s}{2^{j+1}}, \frac{f_s}{2^j}]$, where f_s is the sampling rate (Hz).

MODWT is applied on the whole processed signal for J frequency bands providing wavelet coefficients, denoted as J matrices $\mathbf{M}^j_{\mathbf{X}}$, $j = 1, \ldots, J$ (Figure 6.5). Then columns of each matrix $\mathbf{M}^j_{\mathbf{X}}$ are segmented into L^1 IED and L^2 non-IED time intervals (section 6.2.1, page 51) called segments. The N-dimensional IED or non-IED segment m is denoted as $\mathbf{S}^l_m = \left[\mathbf{s}^l_{1m} \ \ldots \ \mathbf{s}^l_{Nm}\right]$, $m = 1, \ldots, L^l$, where the upper index l denotes the IED (l=1) or non-IED (l=2) variable. L^l and T^l_m are the number of time intervals or segments and the sample number of each segment m, respectively. $\mathbf{s}^l_{im} = \left[s^l_{im}[1] \ \ldots \ s^l_{im}[T^l_m]\right]^T$ is the ith column of matrix $\mathbf{S}^l_m \in \mathbb{R}^{T^l_m \times N}$ that contains T^l_m samples of wavelet coefficients at a given frequency level during IED or non-IED segment m. Figure 6.5 shows how to extract $\mathbf{S}^l_m$ from $\mathbf{X}$.

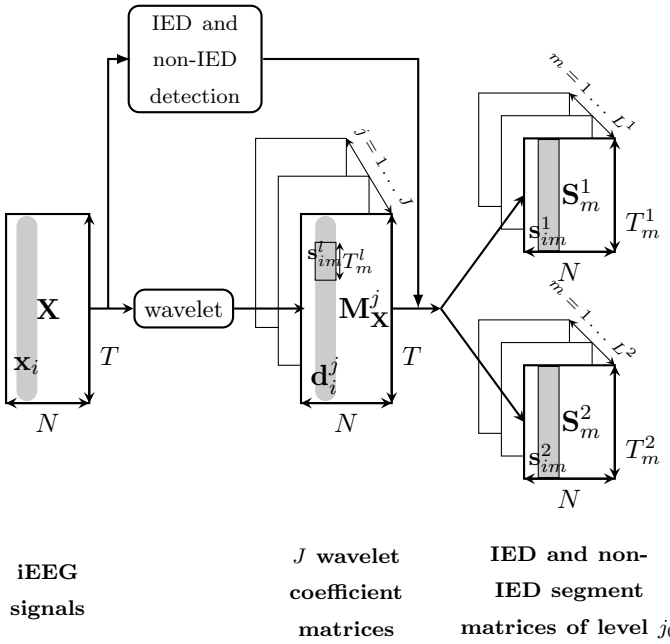


Figure 6.5: IED and non-IED segment matrices. The MODWT transform is applied on the T samples of N-dimensional iEEG recordings, $\mathbf{X}$, for the J frequency levels providing J matrices denoted as $\mathbf{M}^j_{\mathbf{X}}$, $j = 1, \ldots, J$. $\mathbf{x}_i$ and $\mathbf{d}^j_i$ are the ith columns of matrices $\mathbf{X}$ and $\mathbf{M}^j_{\mathbf{X}}$, respectively. $\mathbf{d}^j_i$ contains wavelet coefficients of $\mathbf{x}_i$ at frequency level j. Columns of each matrix $\mathbf{M}^j_{\mathbf{X}}$ (at each given level j) are segmented according to IED and non-IED time intervals providing L^1 and L^2 matrices denoted as $\mathbf{S}^1_m$ and $\mathbf{S}^2_m$, respectively. T^l_m is the number of rows of matrix $\mathbf{S}^l_m$ which is equal to the length of each IED or non-IED segment m denoted as $\mathbf{s}^l_{im}$, where $i = 1, \ldots, N$ is the channel index. j_0 is a given frequency level.

6.2.2 Coupling computation

Different coupling measures [169] may be selected between pairs of time series (3.8, page 24) like measures of linear coupling (e.g., cross-correlation [86, 170], wavelet correlation coefficient [20, 63, 89, 148, 158, 171], coherence [93, 172–177]), nonlinear coupling (e.g., non-linear correlation [13, 178, 179], phase synchrony [61, 96, 97, 102, 108, 113, 172, 180], mutual information [25, 110, 181–185]), and directed coupling (e.g., Granger causality [112, 186], directed coherence [57,117], partial directed coherence [187,188], directed transfer function (DTF) [84,117,146], generalized synchrony [120–124], transfer entropy [125]). In this work, we focus on linear coupling and directional properties are estimated by the estimation of the time delays. We chose linear coupling since there are several studies indicating no preference of nonlinear coupling measures over linear ones during interictal period of EEG analysis [8, 10, 100, 126, 127]. We also tested wavelet-Hilbert synchrony (nonlinear coupling) [97] and found almost similar final results (leading IED regions) confirming the latter studies. Further more linear measures like cross-correlation can be estimated directly from the time series without any partitioning (quantization, coarsegraining), that is valuable compared to some other nonlinear methods like mutual information [189]. An important practical advantage of cross-correlation is the number of parameters. The only parameter of cross-correlation is the maximum time lag, while other measures like Granger causality measures and generalized synchrony need more number of parameters.

Here, our formal coupling measure is maximum wavelet correlation-coefficient. The MODWT cross-correlation coefficients [20,89] are estimated for each channel pair $(i,j) \in \{1,\dots,N\}^2$, $i \neq j$ and during IED or non-IED segments in terms of different time lags (τ):

$$\hat{\rho}_m^l\left(\mathbf{s}_{im}^l, \mathbf{s}_{jm}^l, \tau\right) = \frac{\widehat{cov}\left\{s_{im}^l[k], s_{jm}^l[k-\tau]\right\}}{\sqrt{\widehat{var}(s_{im}^l[k])\widehat{var}(s_{jm}^l[k-\tau])}} \tag{6.5}$$

where $\widehat{cov}$ and $\widehat{var}$ [89] are estimated covariance and variance (by empirical average), respectively. For each channel pair and IED or non-IED segment, the maximum of MODWT cross-correlation (6.5) is obtained as:

$$\begin{aligned} \tau_{ij}^* &= \arg\max_\tau\left(\left|\hat{\rho}_m^l(\mathbf{s}_{im}^l, \mathbf{s}_{jm}^l, \tau)\right|\right) \\ \hat{\rho}_m^{lmax}\left(\mathbf{s}_{im}^l, \mathbf{s}_{jm}^l\right) &= \hat{\rho}_m^l(\mathbf{s}_{im}^l, \mathbf{s}_{jm}^l, \tau_{ij}^*) \end{aligned} \tag{6.6}$$

where $\hat{\rho}_m^{lmax}$, the maximal MODWT cross-correlation (MMCC), is considered as our formal coupling measure. The τ_{ij}^* is the lag between $\mathbf{s}_{im}^l$ and $\mathbf{s}_{jm}^l$ at which the maximum of the absolute value of $\hat{\rho}_m^l\left(\mathbf{s}_{im}^l, \mathbf{s}_{jm}^l, \tau\right)$ occurs. For non-stationary fractionally differenced signals (the definition is given in [89]), the confidence interval of MODWT cross-correlations can be approximated based on [89]. Here, the non-stationary time series do not hold this *particular property*, thus a permutation method is used as explained in section 6.2.3, page 57.

For each $\mathbf{S}_m^l$, the coupling measure (MMCC), $\hat{\rho}_m^{lmax}$, calculated for each channel pair $(i,j) \in \{1,\dots,N\}^2$, provides a square symmetric $N \times N$ matrix, $\Gamma_m^l = \left[\gamma_m^l[i,j]\right]$, where $\gamma_m^l[i,j] = \hat{\rho}_m^{lmax}\left(\mathbf{s}_{im}^l, \mathbf{s}_{jm}^l\right)$. Due to the symmetry, only entries of the upper triangle of matrix Γ_m^l are

considered. We build vector $\mathbf{c}^l_{m:}$ as the concatenation of the upper triangle columns of Γ^l_m. $\mathbf{c}^l_{m:}$ includes the MMCC between all of channel pairs during IED or non-IED time interval m (Fig. 6.6). Length of $\mathbf{c}^l_{m:}$ is the number of possible connections that is $N_c = \frac{N^2-N}{2}$.

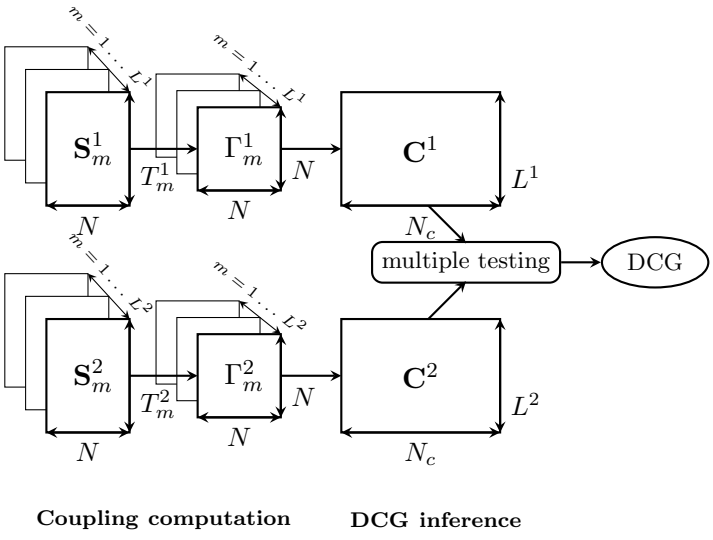


Figure 6.6: Coupling computation and DCG inference at a given frequency band. Each wavelet coefficient matrix of IED or non-IED segment m ($m = 1, \ldots, L^l$), $\mathbf{S}^l_m \in \mathbb{R}^{T^l_m \times N}$ provides one symmetric $N \times N$ coupling measure matrix, Γ^l_m. Each Γ^l_m builds the mth row of the matrix $\mathbf{C}^l \in \mathbb{R}^{L^l \times N_c}$, which contains all of the information concerning couplings between all of the possible connections at each frequency level. The DCG connections are identified by comparing $\mathbf{C}^1$ and $\mathbf{C}^2$, using permutation-based multiple testing.

6.2.3 DCG inference

Multiple testing

The scheme depicted in Figure 6.6 demonstrates how to infer the DCG [20, 63, 148, 190] for a given frequency band. We denote $\mathbf{C}^l = [c^l_{mn}]$, $m = 1, \ldots, L^l$, $n = 1, \ldots, N_c$. $\mathbf{C}^l$ includes MMCC between activities recorded from all of channels in the columns and during all of IED or non-IED time intervals (temporal epileptic activities) in the rows. $\mathbf{C}^1$ and $\mathbf{C}^2$ are served to a multiple hypothesis test (appendix F, page 147) for testing whether or not to assign a connection between channels i and j, $(i, j) \in \{1, \ldots, N\}^2$ (denoted as connection n for simplicity). We test the null hypothesis that the connection n is unchanged under IED and non-IED states against the alternative that their connection changes such that:

$$\begin{cases} H^n_0 : & \mu^1_n = \mu^2_n \\ H^n_1 : & \mu^1_n \neq \mu^2_n \end{cases} \tag{6.7}$$

where μ_n^l is the average of $\mathbf{c}_{:n}^l$, which is the nth column of matrix $\mathbf{C}^l$. The test statistic t_n for each connection n between pair of nodes i and j is defined as:

$$t_n = \frac{\widehat{\mu}_n^1 - \widehat{\mu}_n^2}{\sqrt{\frac{(\widehat{\sigma}_n^1)^2}{L^1} + \frac{(\widehat{\sigma}_n^2)^2}{L^2}}} \tag{6.8}$$

where $\widehat{\mu}_n^l$ and $\left(\widehat{\sigma}_n^l\right)^2$ are the empirical estimated mean and variance of $\mathbf{c}_{:n}^l$, respectively. The simple and ideal method to do each connection test is to assume a theoretical distribution for the test statistic under null hypothesis. Since we do not know the distribution of t_n, we use permutation tests, which require fewer assumptions about the data, thereby yielding more reliable procedures by using data-based distributional characteristics [144, 145, 191].

Permutation resampling method

Figure 6.7 demonstrates different stages of the permutation procedure. We permute N_p (number of permutations) times the coupling measures computed for IED and non-IED groups. More precisely, for each connection n, we permute N_p times the $L^1 + L^2$ entries of $\mathbf{c}_{:n}^1$ and $\mathbf{c}_{:n}^2$, then regroup the first L^1 and L^2 entries as IED and non-IED groups, respectively. For each permutation we calculate the test statistic t_n according to (6.8). The N_p values of t_n form a "permutation" distribution to which we compare the observed t_n of the original data and estimate a raw p-value as follows:

$$p[n] = \frac{card(\{n_p | \left|t_n^{n_p}\right| > |t_n|\})}{N_p} \tag{6.9}$$

where $p[n]$, $n = 1, \ldots, N_c$ denotes the raw p-value of each connection n. $t_n^{n_p}$, $n_p = 1, \ldots, N_p$ is the calculated test statistic of n_pth permutation and $card(.)$ refer to the number of elements. The raw p-value is estimated as the number of permutations which the absolute value of their corresponding test statistic, $t_n^{n_p}$, exceeds the absolute value of the observed t_n (calculated from the original data) divided by N_p. Large number of permutations (N_p is experimentally chosen equal to one million for about 5000 connections) may provide better estimation of p-values. The total number of possible permutations for $L^1 + L^2$ samples, with L^1 and L^2 samples of each group is $\frac{(L^1+L^2)!}{L^1!L^2!}$, i.e. in our experiment over 10^{50}. The N_p permuted sets are chosen randomly from the sets of possible permutations.

Computing the distribution of t_n values for all of the node pairs is time consuming. For instance, if the graph includes 105 nodes (number of channels for iEEG recordings of patient 2, P2), then we construct a distribution for each t_n of the $N_c = \frac{105^2-105}{2} = 5460$ connections. If construction of each distribution (for a single connection) requires $N_p = 10^6$ repetitions then we compute t_n for 5460×10^6 times. Each computation of t_n taking 2.2 μs (on a shared 3 GHz, 4 core Xeon 64 bits processor), then the calculation of raw p-values (for all of connections) takes 200 minutes.

Exchangeability Permutation test relies on the assumption that the distribution of observations does not change by permutation under the null hypothesis [144, 192]. The permutation-

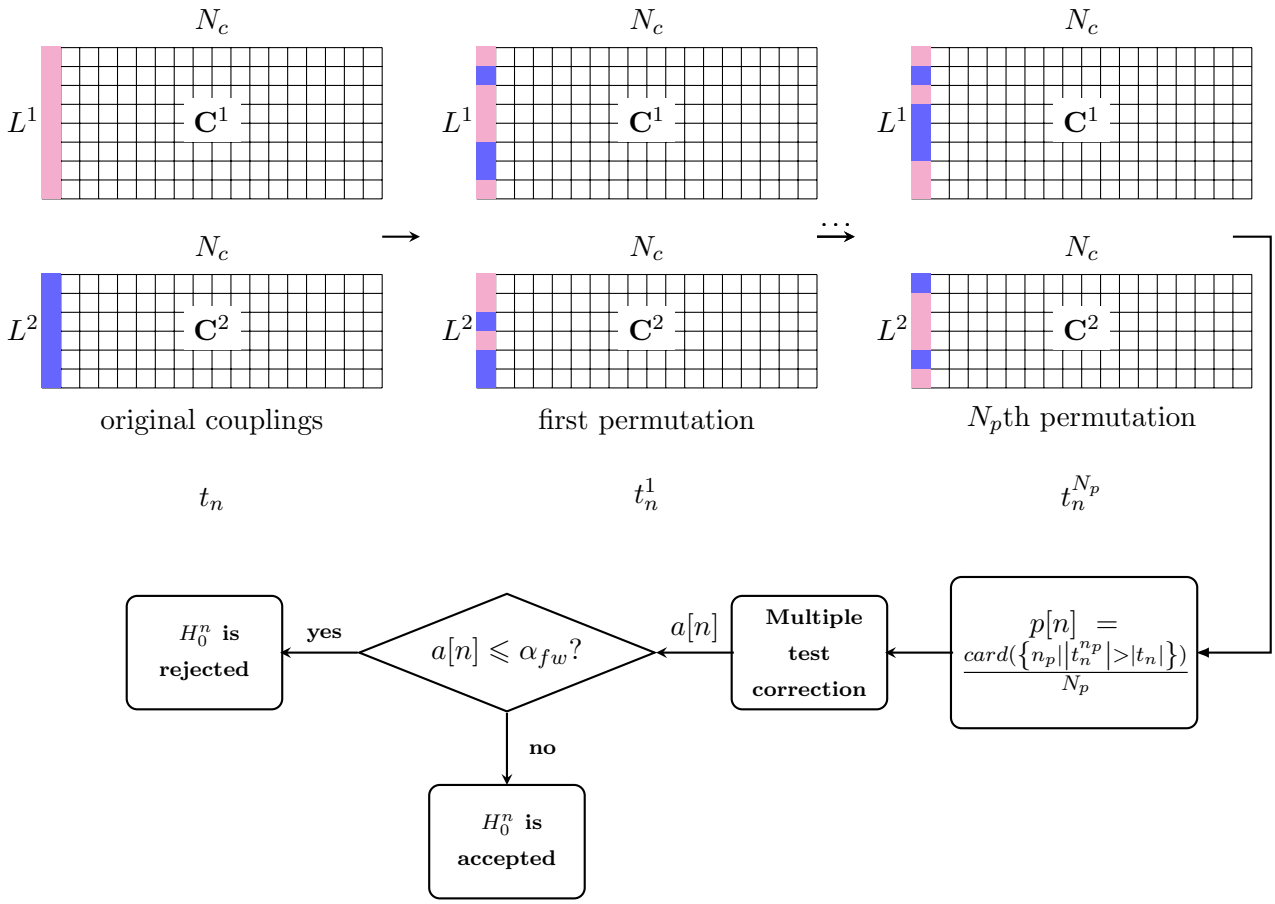


Figure 6.7: Permutation based multiple testing. For each connection n, L^1+L^2 entries of $\mathbf{c}^1_{:n}$ and $\mathbf{c}^2_{:n}$ ($\mathbf{c}^1_{:n}$ and $\mathbf{c}^2_{:n}$ are the colored columns of matrices $\mathbf{C^1}$ and $\mathbf{C^2}$) are permuted and regrouped to L^1 IED and L^2 non-IED entries. The permutation is repeated N_p times. A schematic demonstration of first ($n_p = 1$) and N_pth ($n_p = N_p$) permutation for one connection is depicted in the first row. For the original couplings and each permutation the test statistic (6.8) is calculated denoted as t_n and $t_n^{n_p}$, respectively. Using the estimated distribution of t_n under null hypothesis from N_p values of $t_n^{n_p}$ and its comparison with t_n, the raw p-value of each connection, $p[n]$ is obtained. $p[n]$ is defined in the first right block of second row. $card(.)$ denotes the number of elements. These raw p-values are corrected for each connection using multiple test correction providing adjusted p-values, $a[n]$. The raw p-values are adjusted to keep FWER equal to α_{fw}. The null hypothesis (H_0^n) is rejected if its adjusted p-value is smaller than α_{fw}. As such an edge is inserted between the related nodes in the graph.

based null distribution can be correctly estimated when the number of IED and non-IED time intervals are equal (balanced sample sizes) [145]. However, unbalanced number of samples can be problematic when the sample size is small. Here, the number of time intervals is unbalanced, but large (Table 8.2). The permutation tests for large unbalanced and balanced number of samples were computed to test the performance of the method (section 9.3.1, page 93). The results are comparable based on high similarity percentages between the DCGs related to two conditions of large unbalanced and balanced number of samples. Similarity percentage is defined as the normalized sum of common number of significant or non-significant t-values over number of possible connections. The latter result based on similar performance of permutation test for large unbalanced and balanced number of samples was previously confirmed in [145, 190] for simulated data.

Multiple testing correction

For multiple testing, we need to control the abundance of false positives [145], i.e. the scalar α (significance level or Type I error). Since the probability of making Type I error increases in multiple testing, we adjust raw p-values in order to keep FWER equal to α_{fw} where α_{fw} is familywise α. FWER is the probability of making at least one Type I error for the whole family of N_c tests. For more details please see appendix F.3, page 148.

We suspect that the couplings between signal pairs for different time intervals are non-independent (non-independent tests). As we explained in appendix F.3.1 (page 148), the Šidàk correction method becomes conservative for non-independent tests compared to independent tests. Here we use Šidàk step down method [145, 193] which is more powerful (smaller adjusted p-values) while keeping strong control of the FWER.

Let $a[n]$, $n = 1, \ldots, N_c$ be the adjusted p-values for N_c connections. Šidàk step down adjustments (more details can be found in [145]) are defined as:

$$\begin{cases} a[i] = \max(a[i-1], 1-(1-p[i])^{N_c-i+1}) & 2 \leqslant i \leqslant N_c \\ a[1] = 1-(1-p[1])^{N_c} & i = 1 \end{cases} \tag{6.10}$$

For connection n, we reject the null hypothesis in the test (6.7) when its related adjusted p-value is less than or equal to α_{fw}. The DCG is constructed by keeping the connections whose MMCC values significantly change between IED and non-IED time intervals. In the following, the significant connections which $\left|\widehat{\mu}_n^1\right| > \left|\widehat{\mu}_n^2\right|$ or $\left|\widehat{\mu}_n^2\right| > \left|\widehat{\mu}_n^1\right|$ are called positive or negative connections, respectively. A positive or negative connection demonstrates respectively a significant increase or decrease of coupling between channel pair (i, j) (connection n) during IED state.

For each interested frequency band, its related DCG is calculated, thus for J interested frequency bands, J DCGs are obtained. In the next section, we explain how the direction of the edges of the DCG is estimated.

6.3 Directed DCG (dDCG) computation

6.3.1 Introduction

To estimate the direction of the edges of DCG, we try to estimate the time delay between the MODWT coefficients of the signal pairs observed at the two ends of each edge of DCG. To estimate the time lags, we use the classic method of maximum cross-correlation [58–60, 91].

6.3.2 Causal relationship

Let's assume DCG=$(\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ and $\mathbf{E}$ represent the set of DCG vertices (or nodes associated with bipolar iEEG channels) and edges (or connections), respectively. $d_a^j[k]$ and $d_b^j[k]$ are denoted as the jth level MODWT coefficients (6.1)-(6.3) of signals $\mathbf{x}_a$ and $\mathbf{x}_b$ observed at nodes a and b ($a, b \in \mathbf{V}$), where $k = 1, \ldots, T$, and T is the number of samples of the whole processed iEEG recording. The drive-response relationship between $\mathbf{x}_a$ and $\mathbf{x}_b$ for the resolution level j is defined through the estimation of time lag between $\mathbf{d}_a^j$ and $\mathbf{d}_b^j$, denoted as τ_{ab}^{*j}. τ_{ab}^{*j} is the time lag in which the absolute maximum cross-correlation between $\mathbf{d}_a^j$ and $\mathbf{d}_b^j$ occurs:

$$\widehat{\rho}\left(\mathbf{d}_a^j, \mathbf{d}_b^j, \tau\right) = \frac{\widehat{cov}\left\{d_a^j[k], d_b^j[k-\tau]\right\}}{\sqrt{\widehat{var}(d_a^j[k])\widehat{var}(d_b^j[k-\tau])}} \tag{6.11}$$

$$\tau_{ab}^{*j} = \arg\max_{\tau}\left(\left|\widehat{\rho}\left(\mathbf{d}_a^j, \mathbf{d}_b^j, \tau\right)\right|\right) \tag{6.12}$$

where $\widehat{\rho}$, $\widehat{cov}$, and $\widehat{var}$ are the empirical estimations of correlation coefficient, covariance and variance, respectively.

We estimate the time lag between $\mathbf{d}_a^j$ and $\mathbf{d}_b^j$ signals assuming that:

1. $\mathbf{d}_a^j$ and $\mathbf{d}_b^j$ include sufficient large number of samples
2. There is a connection between node pair (a, b) in the frequency level j.

Upon these two assumptions, the direction of the edge between node pair (a, b) is estimated from a to b if τ_{ab}^{*j} is negative and the reverse relationship if τ_{ab}^{*j} is positive. The reasoning of the above assumptions are explained through the study of the reliability of the time lag estimation in the next section.

We assume that the "transfer function" between signals recoded from different brain regions cannot be represented by pure delay, eventually we do not confine the time causality to be the same for all of the frequency bands, thus we estimate τ_{ab}^{*j} for each frequency level j.

The sign of time delay (τ_{ab}^{*j}) is used in this section to estimate the direction of DCG connections and the value of time delay (τ_{ab}^{*j}) is used to calculate the emittance measure in the next chapter. For simplicity in the following, we mark $\mathbf{d}_a^j$, $\mathbf{d}_b^j$, and τ_{ab}^{*j} as $\mathbf{d}_a$, $\mathbf{d}_b$, and τ_{ab}^{*}, respectively.

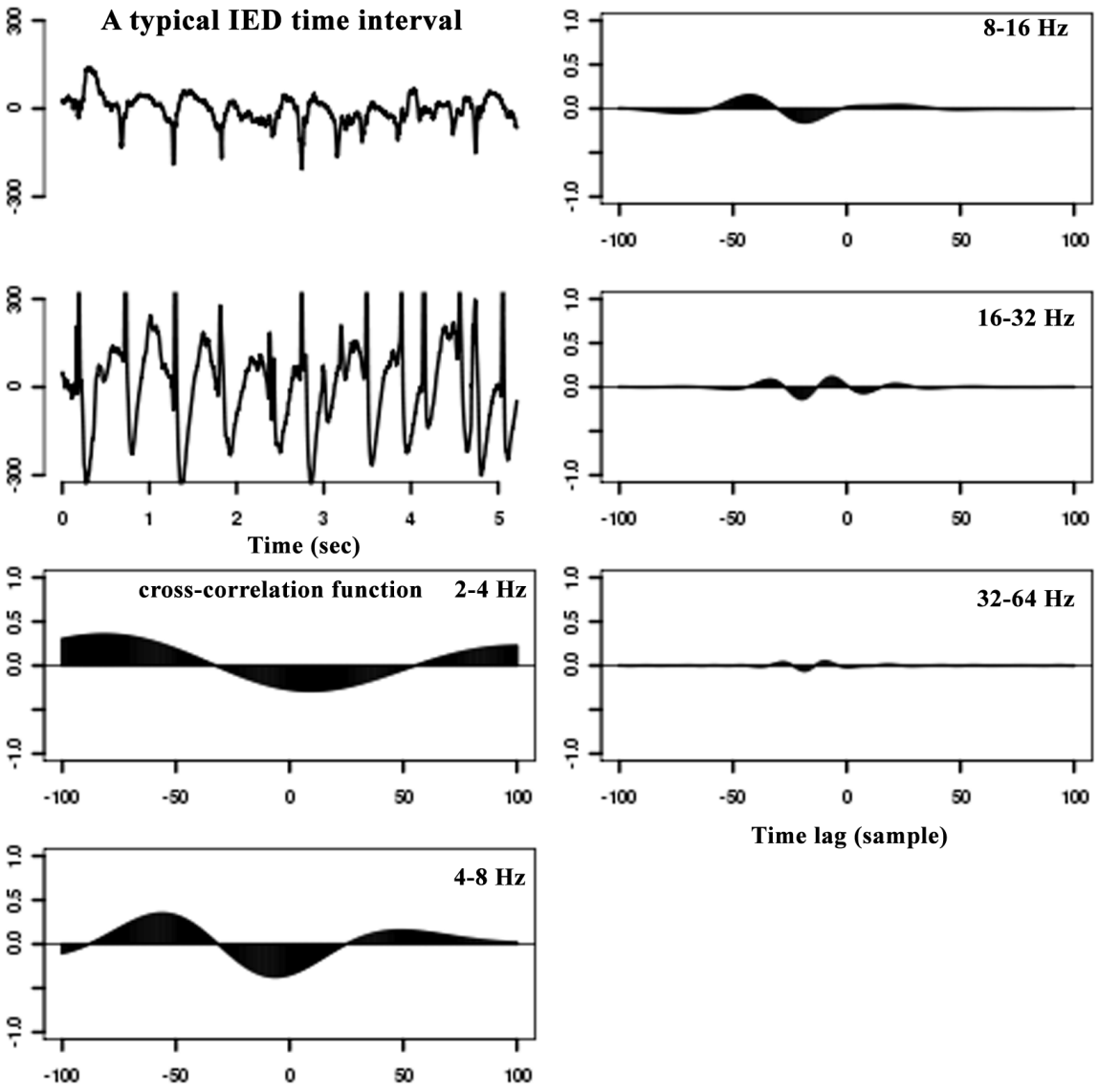


Figure 6.8: A typical temporal signal pair during IED events and the cross-correlation functions between the MODWT coefficients of this signal pair for the time interval of length T in 2-4 Hz, 4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz.

6.3.3 Time causality reliability

The signals studied here are non-stationary signals including biphasic waveforms. A typical temporal signal pair during IED time interval is demonstrated for a length of 5 sec in Figure 6.8. The cross-correlation function (CCF) between the MODWT coefficients of this signal pairs (6.11) for the whole selected time interval of length T is also demonstrated for 2-4 Hz, 4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz in the first and second columns, respectively. The CCF of such signal pairs which include byphasic waveforms oscillate around its maximum correlation and damps out (appendix G.3, page 155). Thus there would be an ambiguity between the two positive and negative extrema of CCF which are close (in absolute value), but one of them is numerically greater. To be confident how significant one of the extrema is greater than the other, the confidence interval of each CCF value is needed.

The confidence interval of MODWT cross-correlation (6.11) is estimated [89] (appendix H, page 157) for non-stationary fractionally differenced signals (for the definition please refer to [89]). For the non-stationary signals studied here, this property does not hold. Furthermore the estimation of the distribution of maximum MODWT cross-correlation for non-stationary signals is complicated. As such to test the reliability of time lags calculated in (6.12) a statistical jackknife method is used as follows.

N_w windows starting at random in the range $[1, T - W]$ are considered where T and W are the length of selected time series (about one hour) and the length of each window (20 minutes), respectively. The window length is considered large enough to include large number of IED time intervals. The time lag is estimated (6.12) for each of randomly chosen windows providing N_w time lag values. These N_w time lag values form a histogram-based probability distribution. For each edge of DCG between node pairs (a, b), $a, b \in \mathbf{V}$, this histogram is estimated and denoted as $\hat{p}_{\tau^*_{ab}}(u)$. The time lag which has the greatest probability is defined as:

$$\bar{\tau}^*_{ab} = \arg\max_u(\hat{p}_{\tau^*_{ab}}(u)) \tag{6.13}$$

For each node pair (a, b), we test if the N_w time lag values are significantly non-equal to zero. If the time lag values are significantly equal to zero then $\bar{\tau}^*_{ab}$ is set to zero, and we conclude that the direction of the edge between node pairs (a, b) cannot be estimated. For the statistical test a bootstrap method with N_b repetitions is used. We compare the $\bar{\tau}^*_{ab}$ values with the τ^*_{ab} estimated (6.12) between signals of length T. The percentage of similar time lags ($\bar{\tau}^*_{ab} \times \tau^*_{ab} > 0$) over number of edges of DCG is in the range [78 95]% for different frequency bands.

Consequently, (6.12) can provide reliable estimation of the most probable time lag between signal pairs of DCG edges if it is calculated for long enough signals and if τ^{max} is selected properly. A discussion on the choice of τ^{max} is brought in section 9.2, page 87.

6.4 Conclusion

Our formal coupling measure is MMCC which is calculated between signal pairs (MODWT coefficients) during IED and non-IED time intervals for DCG calculation and between signal pairs (MODWT coefficients) during the whole selected time period studied for time causality estimation.

DCG provides a statistically reliable method for identification of brain regions involved in interictal events and their relationships. The causal evaluation of these relationships is provided by dDCG. DCG is statistically reliable due to the large number of IED and non-IED time intervals and using permutation based multiple testing. Since DCG includes discriminated connections between IED and non-IED time intervals, its interpretation is easier than the graphs related to either IED or non-IED time interval.

Although in this work DCG is applied on the iEEG recordings of epileptic patients, it may generally be used for other applications. DCG can identify the connections whose statistical properties change between two states. Here these two states are IED and non-IED, but they might be two other states like task and rest. As such using DCG, the brain regions involved in the interested task and their relationships might be identified (see section 11.2, page 117).

dDCG include both source and sink nodes. To distinguish between source and sink nodes, in the next chapter we explain how to characterize the nodes of dDCG.

Chapter 7

Identification of ℓIED regions

7.1 Introduction

For each frequency band, DCG and eventually dDCG are calculated. Now we need to quantify the nodes of each dDCG to measure the strength of source and sink nodes. The usual graph measures do not include enough information for this purpose. Furthermore, once the dDCGs are characterized for source and sink identification, a data mining tool is needed to extract the information from the multi-dimensional or multi-frequency graph measure values (associated with different objective functions). A classic solution is to use single objective optimization methods. These methods provide a single solution for each objective function (related to graph measure values of each frequency band), while we need to optimize the multiple objective functions simultaneously to provide a set of solutions.

In this chapter, we explain how using lagged mutual information and the digraph structure can be informative for characterization of digraphs for the purpose of source and sink identification. Furthermore, we explain how multi-objective optimization method [149, 150] can solve the problem of classic methods especially single objective optimization methods for the estimation of leading IED regions using multi-dimensional graph measure values.

This chapter is organized as follows. First we introduce two new digraph measures to characterize the dDCG, then a multiple graph analysis method based on multi-objective optimization is explained for the estimation of leading IED regions.

7.2 Characterization of dDCG

Nodes of dDCG are related to IED events including source and sink nodes. A source is defined as a node which has high emittance contribution to the rest of digraph. Our aim is to quantify the relevance of each node as a source or to measure the emittance contribution of each node to the rest of digraph. Classic digraph measures like local efficiency, global efficiency [81], out-degree, and in-degree [77] are not designed for this purpose and therefore do not provided enough information to identify the source and sink nodes. Here we propose two new measures:

1. Local information denoted as LI;

2. Total global efficiency denoted as E_{tglob}.

LI measures the amount of information which passes through each node locally that is the difference between outgoing and incoming information, based on lagged mutual information and the digraph structure. This measure is introduced for the purpose of source and sink node identification from digraphs. Positive LI value for a node shows that the node emits information more than it receives which leads to its emittance contribution. The emittance contribution reveals the strength of a source node. Conversely negative LI value shows that the node receives more information than it emits which leads to the strength of a sink node.

One disadvantage of most of classic measures for the identification of source and sink nodes is that these measures consider only incoming connections. Here we wondered if including the effect of incoming paths in global efficiency may improve its precision. To this end we proposed total global efficiency. In E_{tglob}, we consider both outgoing and incoming paths in order to study the effect of incoming paths, in addition to outgoing ones on the characterization of source and sink nodes.

In the following we explain these new digraph measures.

7.2.1 Local information LI

We propose a new index to measure the amount of information that passes through each node locally which we called local information (LI). LI of each node depends on two types of knowledge:

1. Outgoing and incoming connections incident the node (dDCG structure)
2. The amount of information which is carried by each of these connections. The amount of information of each connection is calculated based on lagged mutual information (MI) between the signal pairs observed at the two ends of the connection.

Using these knowledge, LI measures the outgoing information minus incoming information of each node.

Let D_a and D_b be two random variables with probability density functions (PDF) $p_a(u)$, $p_b(v)$ and joint PDF $p_{ab}(u,v)$. The MI between D_a and D_b is defined as:

$$MI(D_a, D_b) = \int\int dudv p_{a,b}(u,v) \log \frac{p_{a,b}(u,v)}{p_a(u)p_b(v)}. \tag{7.1}$$

We assume samples of time series $\mathbf{d}_a = [d_a[1] \ \dots \ d_a[T]]^T$ and $\mathbf{d}_b = [d_b[1] \ \dots \ d_b[T]]^T$ as the observations of D_a and D_b. By partitioning the supports of $\mathbf{d}_a$ and $\mathbf{d}_b$ into bins of finite size and assuming ergodicity, MI (7.1) can be approximated as the finite sum:

$$MI(D_a, D_b) \approx MI_{binned}(\mathbf{d}_a, \mathbf{d}_b) \equiv \sum_{ij} \hat{p}_{ab}(i,j) log \frac{\hat{p}_{ab}(i,j)}{\hat{p}_a(i)\hat{p}_b(j)} \tag{7.2}$$

where $\hat{p}_a(i)$, $\hat{p}_b(j)$, and $\hat{p}_{ab}(i,j)$ are estimated probabilities obtained by relative frequencies. In the following $MI_{binned}(\mathbf{d}_a, \mathbf{d}_b)$ is simplified by $MI(d_a[k], d_b[k-\tau])$, which emphasizes on the

shift τ between the two time series $\mathbf{d}_a = [d_a[k+1] \ \ldots \ d_a[k+T]]^T$ and $\mathbf{d}_b = [d_b[k+1-\tau] \ \ldots \ d_b[k+T-\tau]]^T$.

The local information of the node a, $a \in \mathbf{V}$, denoted as $LI[a]$, is then defined as:

$$LI[a] = \sum_{\mathbf{V}_{a\to b}} MI(d_a[k], d_b[k-\tau^*_{ab}]) - \sum_{\mathbf{V}_{b\to a}} MI(d_a[k], d_b[k-\tau^*_{ab}]) \tag{7.3}$$

where $\mathbf{V}_{a\to b} = \{b \in \mathbf{V} - \{a\} \mid \tau^*_{ab} < 0\}$. In $\mathbf{V}_{b\to a}$, nodes a and b have the reverse relationships, i.e. $\tau^*_{ab} > 0$. In theory, LI of the source or sink nodes are positive or negative, respectively and greater values demonstrate higher emittance contribution or higher strength of a source node. $LI[a]$ is zero when in and out information flows are equal.

LI can be interpreted as the total degree (outdegree minus indegree, section 3.7.1, page 22) of a weighted digraph in which the weight related to the edge from node a to node b is the lagged mutual information (7.3) between the signal pairs associated with nodes a and b ($MI(d_a[k], d_b[k-\tau^*_{ab}])$). This weight can represent the amount of information carrying by the edge.

To know the accuracy of LI estimation, its variance is estimated by jackknife resampling method. For each connected node of dDCG, we calculate the LI (7.3) for N_w windows. Let W be the window length and T the number of samples of the whole processed signal. The start time of each window is a random number in the range $[1, T-W]$. The standard deviation of LI for each connected node, is approximated as the standard deviation of N_w recomputed LI values.

7.2.2 Total global efficiency

To consider the incoming paths (in addition to the outgoing ones) in global efficiency, the total global efficiency is defined as an extension of classic global efficiency.

Here we remind the definition of path and shortest path length matrix. A path from node a to node b (or outgoing path of a, or incoming path of b) is an ordered sequence of non-repeated edges and nodes connecting node a to node b [17]. The shortest path length between nodes a and b is the minimum number of edges traversed to get from a to b.

Shortest path length matrix of dDCG, $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ is denoted as $\mathbf{L_G} = [l_{ab}]$, where l_{ab} is the shortest path length from node a to node b. If any path does not exist from a to b, then $l_{ab} = \infty$.

The information about outgoing paths of node a are in the ath row of $\mathbf{L_G}$, and the information about its incoming paths are in the ath column of $\mathbf{L_G}$. For each node a of $dDCG$, we define the total global efficiency as:

$$E_{tglob}[a] = \frac{1}{N-1} \sum_{b\in\mathbf{V}, b\neq a} \left(\frac{1}{l_{ab}} - \frac{1}{l_{ba}}\right) \tag{7.4}$$

$E_{tglob}[a]$ is in the range $[-1, 1]$. High positive values of $E_{tglob}[a]$ shows that we may reach from node a to the rest of digraph easier than reaching from the rest of digraph to this node. The low negative values show the inverse relationship.

As an example, we calculated the total global efficiency for the nodes of the sample digraph defined in section 3.6, page 21. The total global efficiency ($\mathbf{E}_{tglob} = [E_{tglob}[1], \ldots, E_{tglob}[7]]^T$) for the nodes of digraph $\mathbf{G}$ depicted in Figure 3.1 is equal to $\mathbf{E}_{tglob} = [0.2, -0.4, -0.75, 0.2, 0.3, 0.05, 0.3]^T$. Nodes 5, and 7 have higher positive values than other nodes. It can be seen in Figure 3.1 that these nodes have more outgoing paths than incoming ones. On the other hand node 3 has very low total global efficiency which shows that we can easily reach from other nodes of the digraph to node 3, while it is impossible to access the rest of digraph from this node.

7.2.3 Comparison of new measures with usual digraph measures

Global efficiency (E_{glob}), local efficiency (E_{loc}), and total degree (k_{tot}) are defined in section 3.7, page 22. The efficiency of these usual digraph measures [77, 81] are quantitatively compared with the two new measures. To evaluate the comparison, each of digraph measures is used for the characterization of dDCGs related to different frequency bands to estimate the ℓIED regions for five patients. The ℓIED regions based on the different measures are compared with visually inspected seizure onset zones by the epileptologist. Though the SOZ is theoretically the region to be removed in candidates to resective surgery, it is not easy to be defined in many patients partly because of the difficulties regarding recording seizure periods. Thus it is valuable to wonder if one can predict the SOZ by estimating the IED regions.

LI measure outperforms other usual measures and E_{tglob} in the identification of source and sink nodes (section 10.1, page 95). Since E_{glob} and k_{tot} measures provide more comparable results to LI, here we focus on advantages and disadvantages of LI measure over these two measures.

LI can be interpreted as the k_{tot} of a weighted digraph in which the weight related to each edge is the amount of information carried by the edge (section 7.2.1, page 66). Accordingly, E_{glob} can be interpreted as the K_{out} of a weighted digraph, except the weight of the edge between nodes a and b is the inverse of the shortest path length between these nodes. The other difference between E_{glob}, k_{tot} and LI is that E_{glob} takes into account the higher-order neighborhood (global properties) while k_{tot} and LI consider the first-order neighborhood (local properties). Finally, E_{glob} and k_{tot} are deterministic measures, while LI is a stochastic measure and needs to be estimated correctly. We estimate the accuracy of LI by the estimation of its empirical standard deviation using jackknife resampling method (section 7.2.1, page 66).

Advantages LI measure has several advantages over E_{glob} as follows.

1. LI evaluates each node based on its outgoing and incoming connections and the information which is carried by each of these connections. Therefore, in addition to the structure of the digraph, LI uses the information extracted from signal pairs associated with each node pair, while E_{glob} only uses the structure of the digraph. Consequently, the sensitivity of LI is supposed to be less than E_{glob} to the minor changes in the digraph structure.

2. LI measure evaluates the amount of information of both outgoing and incoming connections which increases its efficiency in the definition of source and sink nodes, while E_{glob}

considers neither the incoming paths, nor the amount of information related to each edge. A node with high E_{glob} may have negative, zero, or positive LI value because in addition to outgoing connections it may receive several incoming connections as well. Moreover, each outgoing or incoming edge may carry different amounts of information. Hence contrary to E_{glob}, LI measure has the potential to be used for distinguishing between source, transit and sink nodes. To analyze an IED related digraph like dDCG, this information is valuable to characterize the role of different brain regions. Although being the patients seizure-free after resective surgery is very valuable, currently there is no way to determine if the removed area is too large. If this is the case, more conservative surgery can be preferred and distinguishing between source, transit and sink regions can be valuable for this purpose.

3. LI is more suited for ranking the ℓIED nodes (section 10.1.4, page 103) than E_{glob}. The node which has greater amount of positive local information is more relevant to a source. Consequently LI which is designed for measuring the emittance contribution is more skilled in ranking the ℓIED nodes than E_{glob}.

Compared with k_{tot}, LI measure is preferred since LI is a weighted version of k_{tot}, whose weights are the amount of information. Thus, for source and sink definition, LI is more suitable than k_{tot}.

Disadvantage The main disadvantage of LI over E_{glob} and k_{tot} is the computation load. Furthermore, LI is more time consuming since for a proper estimation of mutual information including the variance, a long period (about one hour with sampling rate equal to 512 Hz) of signal pairs is supposed to be selected. We compare the computation time for E_{glob}, k_{tot} and LI for example for the dDCG related to patient 3 in 4-8 Hz (Figure 10.1 in section 10.1.4, page 100). This digraph includes 29 nodes and 62 edges. On a shared 3 GHz, 4 core Xeon 64 bits processor, the computation times for E_{glob} and k_{tot} are equal to 0.6 seconds, while it is 9.7 minutes for LI.

7.3 Multiple graph analysis

7.3.1 Introduction

dDCG provides sparser graphs comparing to separated IED and non-IED complex graphs and may provide simpler interpretation, but still concluding to leading IED regions from the dDCGs of different frequency bands needs careful analysis. Proposed LI measure characterizes the multi-frequency dDCG nodes to quantitatively identify the source nodes of each frequency band. The aim is to estimate the source nodes or leading IED regions from multi-dimensional graph measure values associated with different objective functions. For this purpose, different available methods can be considered as follows.

1. A simple method that one can consider is clustering method like k-means (appendix A, page 123). As such, the source nodes or leading IED regions can be defined as the cluster including the nodes which have similar measure values in all of frequency bands. This method has several problems: 1) the number of clusters should be determined, 2) labeling the clusters cannot be automatic, because the data does not include this level of information, 3) clustering is not suited for the extraction of source nodes from multi-dimensional measure values. The preference between different dimensions is unknown. Thus two nodes that each has a maximum value in one of the dimensions (not the same dimensions for the two nodes) will not be in the same cluster while both can be important.

2. The second method can be single-objective optimization. This method results in a unique solution for a single objective function. As such a set of individual optimum solutions of each objective (related to each frequency band) can be provided which is not efficient. Furthermore, in our application the preference between different objectives (related to different frequency bands) is unknown, consequently there is no preference between the individual optimum solutions of each objective.

3. The third method can be scalarization of multiple objective functions into a single objective function. We can define a single objective function like Euclidean norm over the measure values of different frequency bands. This method provides a single solution. Moreover this solution depends on the importance of the objective functions. Either, the importance of all of the objective functions can be assumed equally which is not necessarily true or one can consider some weights for different objectives. However since the preference between objective functions is unknown, this method is not suitable.

Here, we propose a multiple graph analysis method using multi-objective optimization (in the following we recall it Pareto method) [149,150,194–196]. In the next section we explain how this method solves the above problems of classic methods. The optimum source nodes obtained using this method are the estimated leading IED regions.

7.3.2 Multi-objective optimization

Multi-objective optimization method [149,150] considers different objectives simultaneously and provides a set of optimum solutions or a set of alternatives with different trade-offs, instead of one single solution for a single-objective function as in single-objective optimization problems. The set of optimal solutions are called Pareto front or Pareto optimum solutions.

Pareto front is the set of non-dominated solutions, i.e. the solutions which do not dominate each other. Dominance can roughly be defined for two nodes X and Y as follows. Regardless the dimensions which nodes X and Y have equal LI values, node X dominates node Y if its LI values of all other dimensions are greater than related values for Y. The detailed definition of dominance is given in section 7.4.1, page 73. Here, we illustrate the explained aspect with an example shown in Figure 7.1. Let consider the case of two objectives: LI values of two frequency levels j_1 and j_2, which both are to be maximized.

Node A represents a solution with great LI value in frequency level j_2, but low LI value in frequency level j_1. On the other hand, point B represents the converse situation. When the preference between frequency levels j_1 and j_2 is unknown, one cannot choose a unique solution between A and B. In other words there is no dominance or priority between solutions that one is better than the other in one objective, but worse in another objective. In the illustrated example, the set of nodes A, B, C, and F construct the Pareto front, since there is no other node which dominates these nodes and moreover these nodes do not dominate each other. In other words the Pareto front is the set of non-dominated nodes.

In Figure 7.1, node D is an example of non-Pareto-front node. If we compare node D with node B or F, again there is no preference between them (none of them dominate the other one). But there exist other nodes (A, C, and E) that have greater values of LI in both frequency levels, and thus dominate node D. Another non-Pareto-front node is node E. By comparing node C with node E, node C has greater value of LI in frequency level j_1 and equal LI value in frequency level j_2, thus node C dominates node E.

To show one example of the differences between multi-objective optimization methods and classic methods like scalarizing the multiple objective functions (the third method explained in section 7.3.1, page 69), we assume that the search space includes only three nodes C, E, and F. For the scalarization method, we assume using Euclidean norm over graph measure values of different frequency bands. Based on Euclidean norm, node C has the greatest norm and nodes E and F have equal norms. Here we compare the optimum solutions for these two methods assuming two following cases:

1. *Node C does not exist in the search space*: if node C did not exist, then the set of optimum solutions based on scalarization and multi-objective optimization would include both nodes E and F. In scalarization, nodes E and F are selected since both have equal norms and in multi-objective optimization, these nodes are selected since there is no preference between them and moreover there is no other node dominating these nodes.

2. *Node C exists in the search space*: if node C exists, then the set of optimum solutions based on scalarization would include only node C since it has the greatest norm. Conversely, the set of optimum solutions based on multi-objective optimization would include nodes C and F since these nodes do not dominate each other and there is no other node dominating these nodes. Node E is not a Pareto front node because node C dominates this node.

Multi-objective optimization includes three stages:

1. Model building

2. Optimization

3. Decision making.

First the model including the formulation of the optimization problem with specifying decision variables, and objectives are explained. Second the optimization or the algorithm which finds

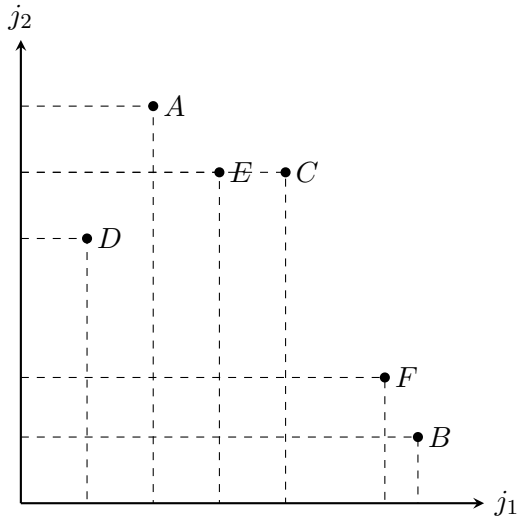


Figure 7.1: Multi-objective optimization.

the optima of the optimization problem is described. Third since the Pareto optimum solutions have no preference to each other, the decision making methods are specified to rank the optimal solutions. In the following, we explain these three stages in the way we applied for solving our problem.

7.4 Application of multi-objective optimization method for the estimation of ℓIED nodes

This section includes two main subsections as follows.

1. *Classic (deterministic) multi-objective optimization method*: in this subsection, we explain how the classic (deterministic) multi-objective optimization method is used for the estimation of ℓIED nodes. To this end, two optimization algorithms are used: 1) the classic Pareto optimization algorithm, and 2) the new neighbor-Pareto optimization algorithm.

2. *New statistical multi-objective optimization method*: in this subsection, we introduce a new statistical multi-objective optimization method in comparison with the deterministic multi-objective optimization method.

7.4.1 Multi-objective optimization method (deterministic)

Model building

The local information measure defined in (7.3) is calculated for the connected dDCG nodes of all of frequency levels, forming a matrix $\mathbf{LI} = [LI^j[n]] \in \mathbb{R}^{N\times J}$, $n = 1, \ldots, N$, $j = 1, \ldots, J$. We denote $\mathbf{LI}^j[:]$ the jth column of matrix $\mathbf{LI}$ including the LI values of all of the nodes in each frequency level j. The row n of the matrix $\mathbf{LI}$, denoted $\mathbf{LI}^{\cdot}[n]$ is a $J-$dimensional feature vector including the local information values of node n in all of J frequency levels.

The multi-objective optimization problem has the following form:

$$\text{maximize} \quad \left\{LI^{j=1}[n], LI^{j=2}[n], \ldots, LI^{j=J}[n]\right\} \tag{7.5}$$

consisting of J objective functions that are aimed to be maximized simultaneously. Generally the objective values (objective function values) are the functions of the decision vector (variable). In our application the objective vector and decision vector are the same. The decision vector $\mathbf{LI}^{\cdot}[n]$ belong to the search space $P \subset \mathbb{R}^J$.

In the following, two optimization algorithms to find the set of optimum solutions are explained.

Optimization

Here, two Pareto optimization algorithms are explained which the first one is the classic Pareto optimization algorithm and the second one is the new neighbor-Pareto optimization algorithm. We introduce the neighbor-Pareto optimization algorithm by improving the classic one to include the neighbor nodes of Pareto front in the set of optimum solutions.

Pareto optimization algorithm: Roughly speaking, a Pareto optimal set is a subset of N nodes whose $\mathbf{LI}^{\cdot}[n]$ is significantly large. More precisely the Pareto optimal set denoted as $D(P)$ can be obtained with the following Pareto optimization algorithm. Let consider N J-dimensional decision vectors, $\mathbf{LI}^{\cdot}[n]$ as N nodes in the search space P.

1. Initialize $D(P)$ with $\mathbf{LI}^{\cdot}[n_0]$, n_0 can be any node.
2. Compare $\mathbf{LI}^{\cdot}[n] \in P$, $n \neq n_0$ with the members of $D(P)$ using the following conditions. In the first iteration the only member of $D(P)$ is node n_0.
   (a) If $\mathbf{LI}^{\cdot}[n] \geqslant \mathbf{LI}^{\cdot}[n_0]$ for all of j levels, and $\mathbf{LI}^{\cdot}[n] > \mathbf{LI}^{\cdot}[n_0]$ for at least one level j, then node n dominates node n_0 and thus $\mathbf{LI}^{\cdot}[n_0]$ is replaced with $\mathbf{LI}^{\cdot}[n]$[1].
   (b) Else if $\mathbf{LI}^{\cdot}[n] > \mathbf{LI}^{\cdot}[n_0]$ for at least one level j, then $\mathbf{LI}^{\cdot}[n]$ is added to $D(P)$, since there is no preference between nodes n and n_0.
3. Increment n and go to previous step till n reaches to N.
4. The members of $D(P)$ are the Pareto optimal solutions denoted as $\mathbf{LI}^{\cdot}[n^*]$ which build the Pareto front or estimated ℓIED nodes.

The block representation of the above algorithm is demonstrated in Figure 7.2.

Neighbor-Pareto optimization algorithm: A margin can be assumed for the Pareto front. To identify this margin, we test the neighborhood of the Pareto optimal solutions during the optimization process. The neighborhood test in terms of the normalized square distance between a pair of decision vector is defined as:

[1]Please note that in the two inequalities, one of them is a strict inequality

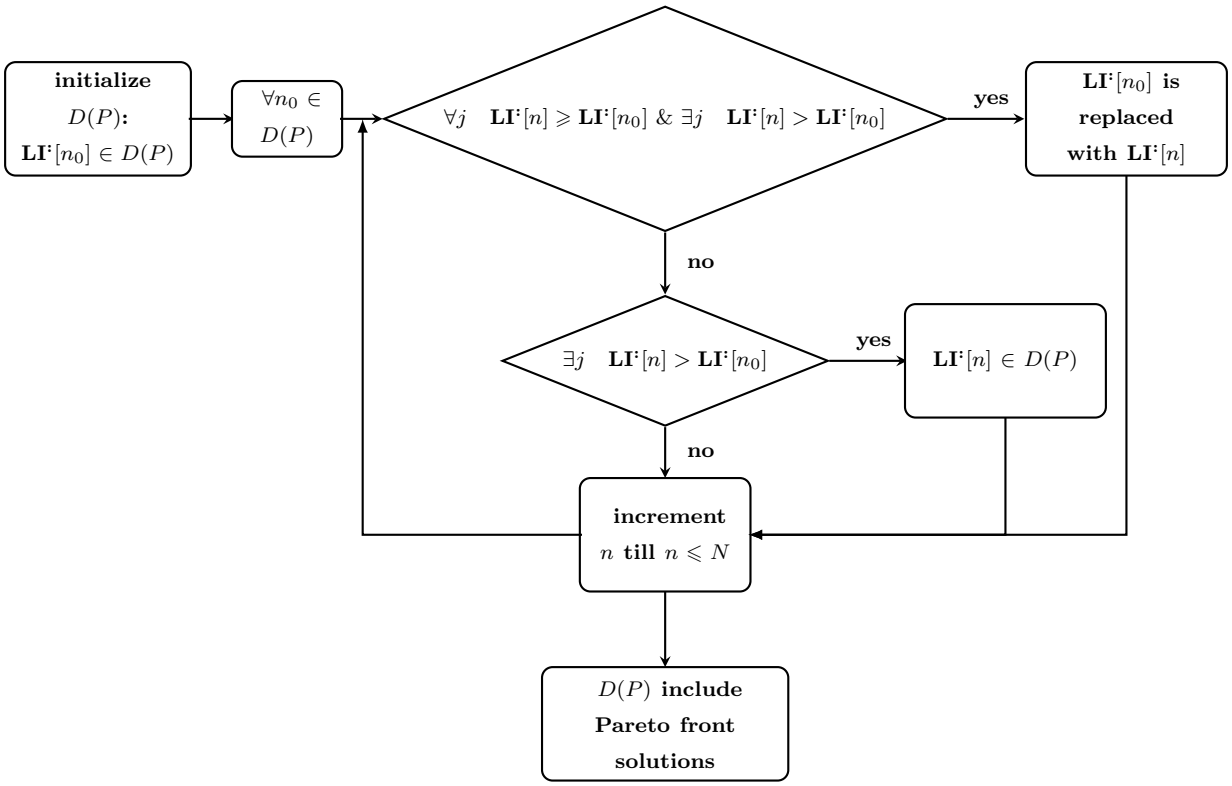


Figure 7.2: The block representation of the Pareto optimization algorithm in multiple graph analysis.

$$\frac{\|\mathbf{LI}^{\cdot}[n] - \mathbf{LI}^{\cdot}[n_0]\|_2^2}{\|\mathbf{LI}^{\cdot}[n]\|_2 \ \|\mathbf{LI}^{\cdot}[n_0]\|_2} \leqslant \epsilon \tag{7.6}$$

where ϵ is a small positive value and $\|\cdot\|_2$ refers to Euclidean distance. The normalized distance is simply the square Euclidean distance between a pair of decision vectors divided by the norms of each of decision vectors. As such the normalized distance is adapted to each node pair, i.e. thicker margin is assumed for the nodes which have greater norms.

Using the neighborhood test, the Pareto optimization algorithm can be improved to include the neighbor nodes of Pareto front in the set of optimum solutions. In part (a) of second step of Pareto optimization algorithm, the neighborhood test is added to see if node n_0 is a neighbor of node n. If node n dominates node n_0 and the neighborhood test is true, then non-dominated node (n_0) can be kept as the member of $D(P)$ since it is in the neighborhood of the dominant node n. Accordingly an extra condition should be added in the second step to do the neighborhood test for the case node n_0 dominates node n. The improved algorithm called as *neighbor-Pareto optimization* is as follows:

1. Initialize $D(P)$ with $\mathbf{LI}^{\cdot}[n_0]$, n_0 can be any node.

2. Compare $\mathbf{LI}^{\cdot}[n] \in P$, $n \neq n_0$ with the members of $D(P)$ using the following conditions. In the first iteration the only member of $D(P)$ is node n_0.

   (a) If $\mathbf{LI}^{\cdot}[n] \geqslant \mathbf{LI}^{\cdot}[n_0]$ for all of j levels, and $\mathbf{LI}^{\cdot}[n] > \mathbf{LI}^{\cdot}[n_0]$ for at least one level j, then

i. If $\frac{\|\mathbf{LI}^{\cdot}[n]-\mathbf{LI}^{\cdot}[n_0]\|_2^2}{\|\mathbf{LI}^{\cdot}[n]\|_2 \|\mathbf{LI}^{\cdot}[n_0]\|_2} > \epsilon$, $\mathbf{LI}^{\cdot}[n_0]$ is replaced with $\mathbf{LI}^{\cdot}[n]$, since node n dominates node n_0.

ii. Else if $\frac{\|\mathbf{LI}^{\cdot}[n]-\mathbf{LI}^{\cdot}[n_0]\|_2^2}{\|\mathbf{LI}^{\cdot}[n]\|_2 \|\mathbf{LI}^{\cdot}[n_0]\|_2} \leqslant \epsilon$, $\mathbf{LI}^{\cdot}[n]$ is added to $D(P)$ (node n_0 is not rejected) since although node n dominates node n_0, node n_0 is in the neighborhood of node n.

(b) Else if $\mathbf{LI}^{\cdot}[n] > \mathbf{LI}^{\cdot}[n_0]$ for at least one level j, then $\mathbf{LI}^{\cdot}[n]$ is added to $D(P)$, since there is no preference between nodes n and n_0.

(c) Else if $\frac{\|\mathbf{LI}^{\cdot}[n]-\mathbf{LI}^{\cdot}[n_0]\|_2^2}{\|\mathbf{LI}^{\cdot}[n]\|_2 \|\mathbf{LI}^{\cdot}[n_0]\|_2} \leqslant \epsilon$ then $\mathbf{LI}^{\cdot}[n]$ is added to $D(P)$, since node n is in the neighborhood of node n_0.

3. Increment n and go to previous step till n reaches to N.

4. The members of $D(P)$ are the neighbor-Pareto optimal solutions ($\mathbf{LI}^{\cdot}[n^*]$).

The block presentation of this algorithm is depicted in Figure 7.3.

In part (b) of Figure 7.4 a typical Pareto front ($\mathbf{LI}^{\cdot}[n^*]$) according to the Pareto optimization algorithm is demonstrated. The $\mathbf{LI}^{\cdot}[n]$ for all of the nodes of dDCG are shown in part (a). Each $\mathbf{LI}^{\cdot}[n]$ is demonstrated by a pentagonal which is related to an electrode lead. The pentagonals of electrode leads related to one electrode are depicted in unique color. In part (c), the neighbor-Pareto optimal solutions are demonstrated. It can be seen that the pentagonal related to electrode c is not considered in part (b) since one of the nodes related to electrode a dominated this node. However the node related to electrode c is included in the neighbor-Pareto optimal solutions since it is in the neighborhood of one of the nodes related to electrode a (according to the neighborhood test).

The Pareto optimal solutions are a subset of neighbor-Pareto optimal solutions. The neighbor-Pareto optimization algorithm is less conservative compared to Pareto optimization algorithm, i.e. it includes the nodes which are in the neighborhood of Pareto front nodes. However in neighbor-Pareto optimization algorithm the parameter ϵ is added. This parameter is associated with the thickness of the margin. If ϵ is set equal to zero, both optimization algorithms provide the same optimum solutions.

Decision making: ranking the estimated ℓIED nodes

There is no preference between the Pareto front solutions. To rank the solutions according to their relevance, a post-optimal technique or some decision-making considerations are required.

Here we focus on two usual techniques [149]:

1. Compromise programming

2. Utility functions.

In the following we explain these methods and the way we applied here.

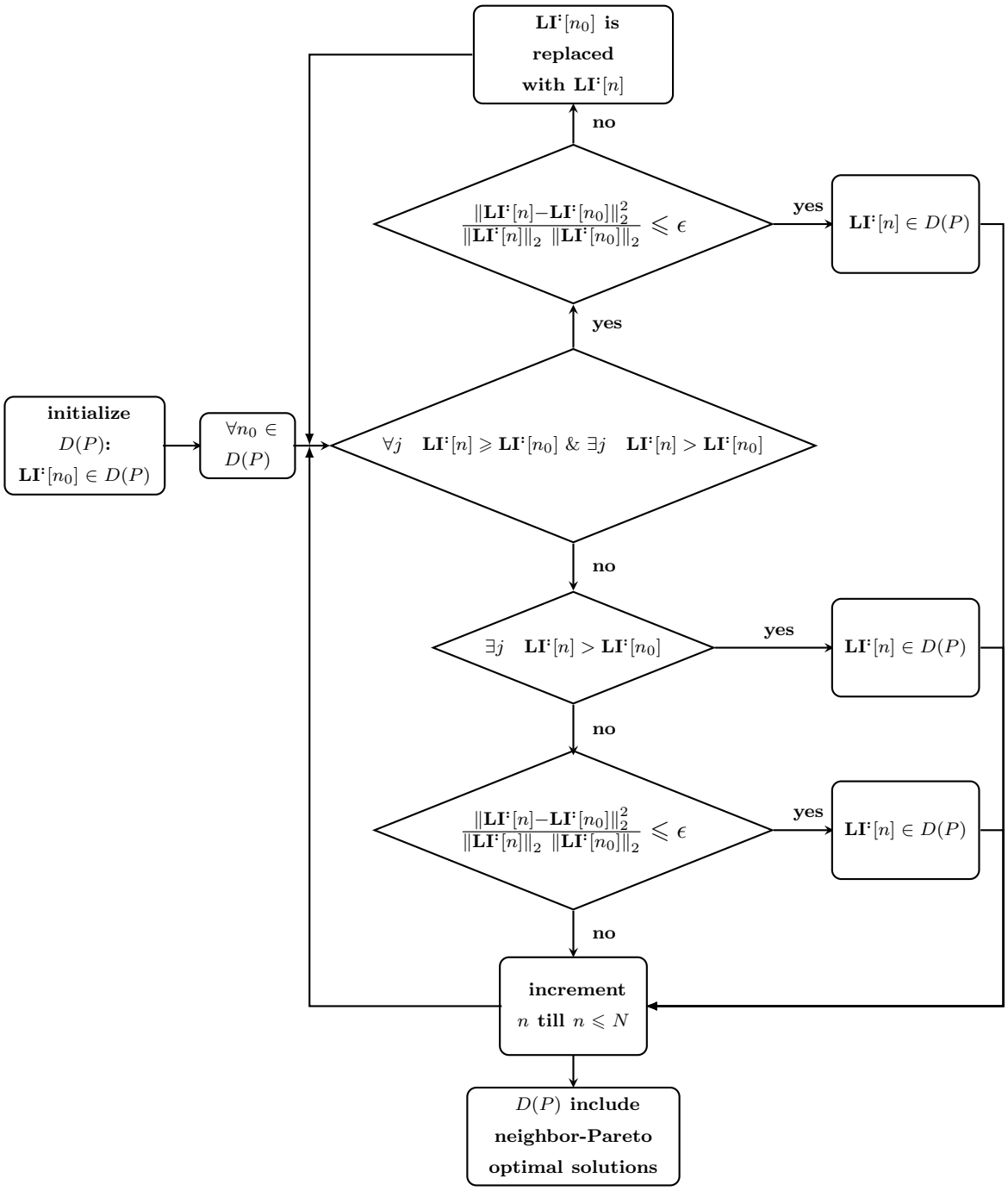


Figure 7.3: The block representation of the neighbor-Pareto optimization algorithm in multiple graph analysis. The difference between this algorithm and Pareto optimization algorithm (Figure 7.3) is that in the former the nodes in the neighborhood of Pareto front solutions are also included using the neighborhood test ($\frac{\|\mathbf{LI}^{\cdot}[n]-\mathbf{LI}^{\cdot}[n_0]\|_2^2}{\|\mathbf{LI}^{\cdot}[n]\|_2 \ \|\mathbf{LI}^{\cdot}[n_0]\|_2} \leqslant \epsilon$).

Compromise programming: In compromise programming method [197], we select a solution from Pareto front which locates closer to an ideal point. The ideal point is compromised of the individual best values of each dimension that is usually infeasible solution. The definition of ideal point and the distance measure is application-dependent and different definitions could be considered. Here, we compute the ideal point, $\mathbf{z} = [z_1, \ldots, z_J]$ as:

$$\begin{aligned} z_j &= \max_{n \in N^*}(LI^j[n]) \\ N^* &= \{n | \mathbf{LI}^{\cdot}[n] \in D(P)\} \end{aligned} \tag{7.7}$$

then we rank $\mathbf{LI}^{\cdot}[n^*]$ according to their increasing distance (Euclidean distance) to $\mathbf{z}$:

$$d(\mathbf{LI}^{\cdot}[n^*], \mathbf{z}) = (\sum_{j=1}^{J} \left| LI^j[n^*] - z_j \right|^2)^{\frac{1}{2}}. \tag{7.8}$$

The Pareto front solution with the smaller value of $d(\mathbf{LI}^{\cdot}[n^*], \mathbf{z})$ is preferred.

Utility functions: The solution which maximizes the utility function has the most relevance. Equivalently the multiple objective functions are reduced to a single-objective function that is maximizing the utility function. The definition of utility function is problem-dependant [149]. Here, we use a utility function $U(\mathbf{LI}^{\cdot}[n^*])$ which allows to order the $\mathbf{LI}^{\cdot}[n^*]$ vectors by comparing their norms, e.g.

$$\begin{aligned} U_1(\mathbf{LI}^{\cdot}[n^*]) &= \|\mathbf{LI}^{\cdot}[n^*]\|_1 = \textstyle\sum_{j=1}^{J} LI^j\,[n^*] \\ U_2(\mathbf{LI}^{\cdot}[n^*]) &= \|\mathbf{LI}^{\cdot}[n^*]\|_\infty = \max_{1 \leqslant j \leqslant J}(LI^j\,[n^*]) \end{aligned} \tag{7.9}$$

The solution with greater utility function (norm) value deserves greater attention.

7.4.2 Statistical multi-objective optimization method

The statistical multi-objective optimization method has the same concept of classic (deterministic) multi-objective optimization method in a statistical context, i.e. instead of comparing two deterministic scalar values, we compare two random variables or the realizations related to these random variables.

Here we explain more details about the statistical multi-objective optimization method. The LI value of each node n at each frequency band is assumed to be a random variable which its realizations are denoted as vector $\mathbf{LI}^j[n, \mathbf{:}] \in \mathbb{R}^{1 \times N_w}$. To calculate this vector, and for each window the local information of the node in the related dDCG is calculated using jackknife method (for the details of jackknife method refer to section 7.2.1, page 66). The LI values of all of the nodes in all of frequency bands form a matrix $\mathbf{LI} = [LI^j[n, m]] \in \mathbb{R}^{N \times J \times N_w}$, $n = 1, \ldots, N$, $j = 1, \ldots, J$, $m = 1, \ldots, N_w$. We assume $\mathbf{LI}^{\cdot}[n, \mathbf{:}]$ belonging to the search space $P \subset \mathbb{R}^{J \times N_w}$ as the decision matrix of node n which includes the LI values of N_w windows in J frequency levels for node n.

The multi-objective optimization problem can be in the following form:

$$\text{maximize} \quad \left\{LI^{j=1}[n,:], LI^{j=2}[n,:], \ldots, LI^{j=J}[n,:]\right\} \tag{7.10}$$

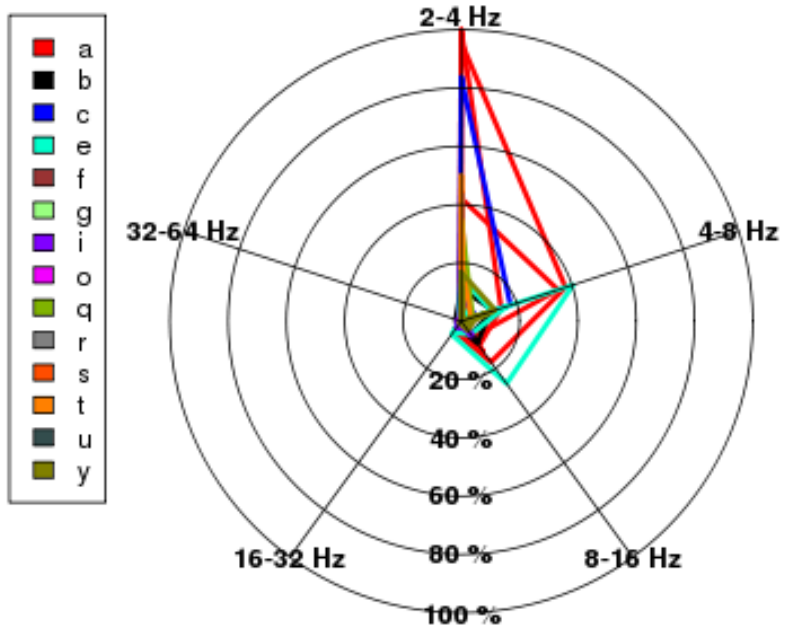
consisting of J objective vectors that are aimed to be maximized simultaneously.

The optimization algorithm is applied on N $J \times N_w$ decision matrices in the search space P. For comparing two nodes say n and n_0 in the second step of the Pareto optimization algorithm, a statistical test based on permutation is applied between vectors $\mathbf{LI}^{\,j}[n,:]$ and $\mathbf{LI}^{j}[n_0,:]$. We test the alternative hypothesis that the LI related to node n and n_0 are not equal at frequency level j against the null hypothesis that they are equal. If the null hypothesis is rejected and the related t-value is positive, the LI of node n is significantly greater than LI of node n_0 at frequency level j, otherwise it is smaller. If the null hypothesis is not rejected, the LI of node n and n_0 are equal at frequency level j. The $D(P)$ of the output of the optimization algorithm includes the statistically significant Pareto optimal solutions.

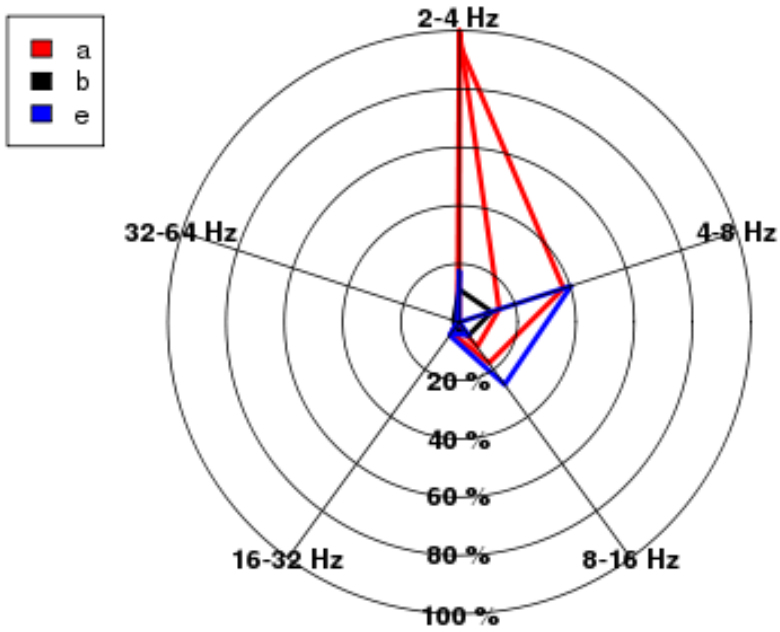
7.5 Conclusion

In this chapter, we introduced two new digraph measures for the identification of source and sink nodes from digraphs: local information (LI) and total global efficiency. Comparison of LI with usual measures as well as total global efficiency revealed that LI is more informative and skilled for the definition of source and sink nodes (see section 10.1, page 95) since it considers incoming connections as well as outgoing ones and the information which is carried by each connection.

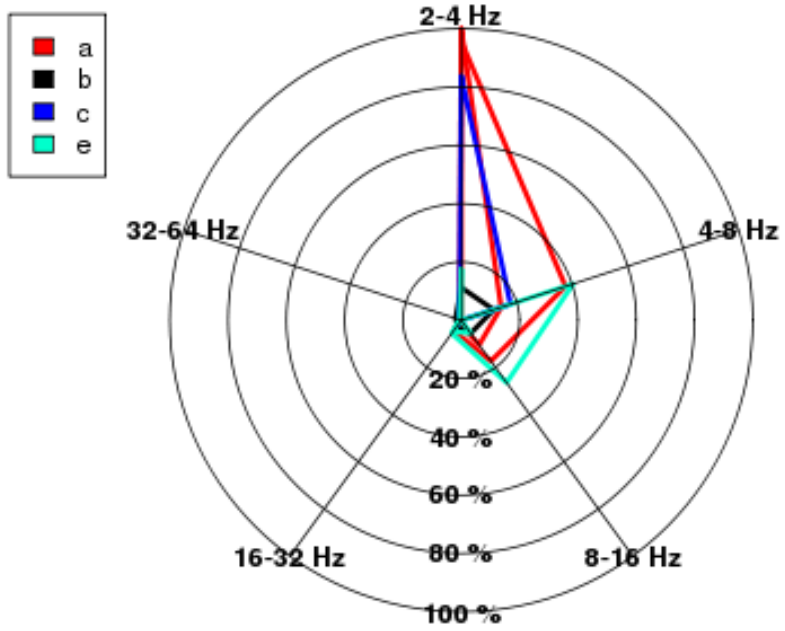
The multiple graph analysis method based on multi-objective optimization [149, 150] is used to extract the information from multi-dimensional or multi-frequency digraph measure values. Using multi-objective optimization method, the multiple objective functions related to digraph measure values of different frequency bands are optimized simultaneously to provide a set of optimum solutions. We found multi-objective optimization method [149, 150] more useful than classic methods for the estimation of leading IED regions using multi-dimensional measure values. In this method, two optimization algorithms are used. The first algorithm is the classic Pareto optimization and the second one is the new neighbor-Pareto optimization. The Pareto optimal solutions are a subset of neighbor-Pareto optimal solutions since the latter includes the nodes in the neighborhood of Pareto front as well as Pareto front nodes. Both of these algorithms can be used for identification of 1) strong ℓIED regions using Pareto optimization algorithm and 2) less strong ℓIED regions using neighbor-Pareto optimization algorithm.



(a) bipolar iEEG channels



(b) Pareto front (Pareto optimization algorithm)



(c) neighbor-Pareto optimum solutions (neighbor-Pareto optimization algorithm)

Figure 7.4: The demonstration of five dimensional local information measure vectors (a) $\mathbf{LI}^{\cdot}[n]$, $n \in \{1, 2, \ldots, N\}$, (b) $\mathbf{LI}^{\cdot}[n^*]$ (Pareto optimization algorithm), and (c) $\mathbf{LI}^{\cdot}[n^*]$ (neighbor-Pareto optimization algorithm) related to P1 in web plot. The five dimensions correspond to five frequency bands from 2 to 64 Hz. The pentagonals related to electrode leads of one electrode are depicted with the same color.

Part III

Experimental results

Chapter 8

Data and parameters

8.1 Introduction

In this chapter, the protocol of iEEG recordings, the parameters of the patients' iEEG, and the parameters of the method are reported.

Part III includes two other chapters as follows:

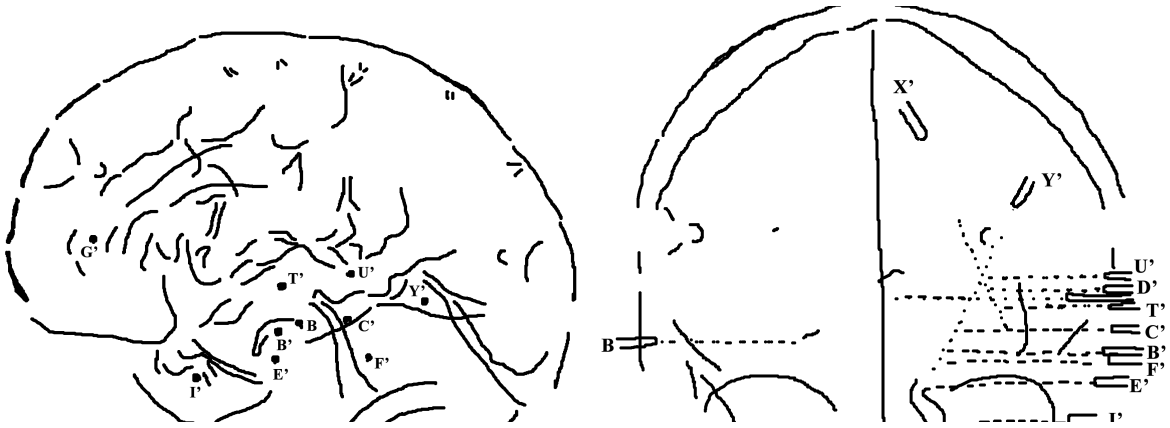
1. dDCG
2. Estimated ℓIED regions.

Chapter 9 (dDCG) is devoted to the dDCG results, choice of the parameters and the reliability of DCG. In chapter 10 (estimated ℓIED regions), we report the results of the estimated ℓIED regions and their reliability. Further more we compare our method with other classic methods.

8.2 Data

The iEEG recordings were obtained from five patients suffering from focal epilepsy. The type of epilepsy of the patients are brought in Table 8.1. The patients underwent pre-surgery evaluations with the iEEG recordings [15,198]. They are seizure free after resective surgery. Eleven to fifteen semi-rigid multi-lead intracerebral electrodes with 0.8 mm diameter were bilaterally implanted in suspected seizure origins based on clinical considerations. The multi-lead electrodes (Dixi, Besançon, France) include 5, 10, 15 or 18 leads. Each lead has 2 mm length and the leads are evenly spaced with inter-space of 1.5 mm. The iEEG were recorded with an audio-video-EEG monitoring system (Micromed, Treviso, Italy) with a maximum of 128 channels and sampled at 512 Hz. The electrode leads were recognized on the patient's implantation scheme, and localized in the Montreal Neurological Institute (MNI) atlas. As an example, the implantation scheme of patient 3 (P3) is brought in Figure 8.1. Bipolar derivations were considered between adjacent leads within each electrode, i.e. each recording site is referenced to its nearest neighbor (bipolar montage) [199]. For simplicity, these adjacent bivariate derivations are represented as e_i instead of $e_{i+1} - e_i$. Bipolar montage in iEEG provides higher local precision compared to common reference which one remote site is used as reference for all of the channels [200–202].

The 50 Hz is removed by a 5th-order notch Butterworth filter with 3dB lower, and upper band stop frequencies equal to 48 Hz and 52 Hz, respectively.



(a) Sagittal view (b) Coronal view

Figure 8.1: Implantation scheme of the iEEG electrodes for P1 (a) Sagittal view; (b) Coronal view.

Table 8.1: Type of epilepsy of the patients. L/R:left/right; T:temporal; midInsG: middle short gyrus of insula.

patient	focal epilepsy
P1	LT
P2	LT
P3	LT
P4	RT
P5	RmidInsG

Table 8.2: Parameters of the patients' iEEG. N: number of the bipolar channels; T: length of the original (non segmented) iEEG signal (minutes); N_c: number of possible connections; L^l: number of IED or non-IED time intervals.

	P1	P2	P3	P4	P5	mean
N	104	105	111	109	100	106
T (minutes)	61	56	42	90	66	55.44
N_c	5356	5460	6105	5886	4950	5551
L^1	298	614	223	160	223	304
L^2	143	200	195	183	148	174

8.3 Parameters

The parameters of the patients' iEEG and method are reported in Tables 8.2 and 8.3, respectively.

Table 8.3: Parameters of the method. τ^{max}: maximum number of time lags (samples) examined in cross-correlation function analysis; f_s: Sampling rate (Hz); N_p: Number of permutations; N_w: number of windows in jackknife method; N_b: number of bootstrap repetitions in jackknife method; W: length of each window (minutes) in jackknife method; SP: statistical multi-objective optimization method.

Method parameter	
Wavelet filter	'la8'
Number of wavelet levels	5
False positive error (α)	0.05
α_{fw} (familywise α)	0.05
τ^{max} for DCG (samples)	27
τ^{max} for dDCG (samples)	100
f_s (Hz)	512
N_p for DCG	10^6
N_p for SP	10^4
N_w	100
N_b	10^4
W for reliability test of τ^* (minutes)	20
W for reliability test of LI and SP (minutes)	33

Chapter 9

dDCG

In this chapter, first the resulted dDCGs are presented. Next the choice of parameter τ^{max} in DCG calculation and time causality estimation are discussed. Finally the reliability of DCG is explained.

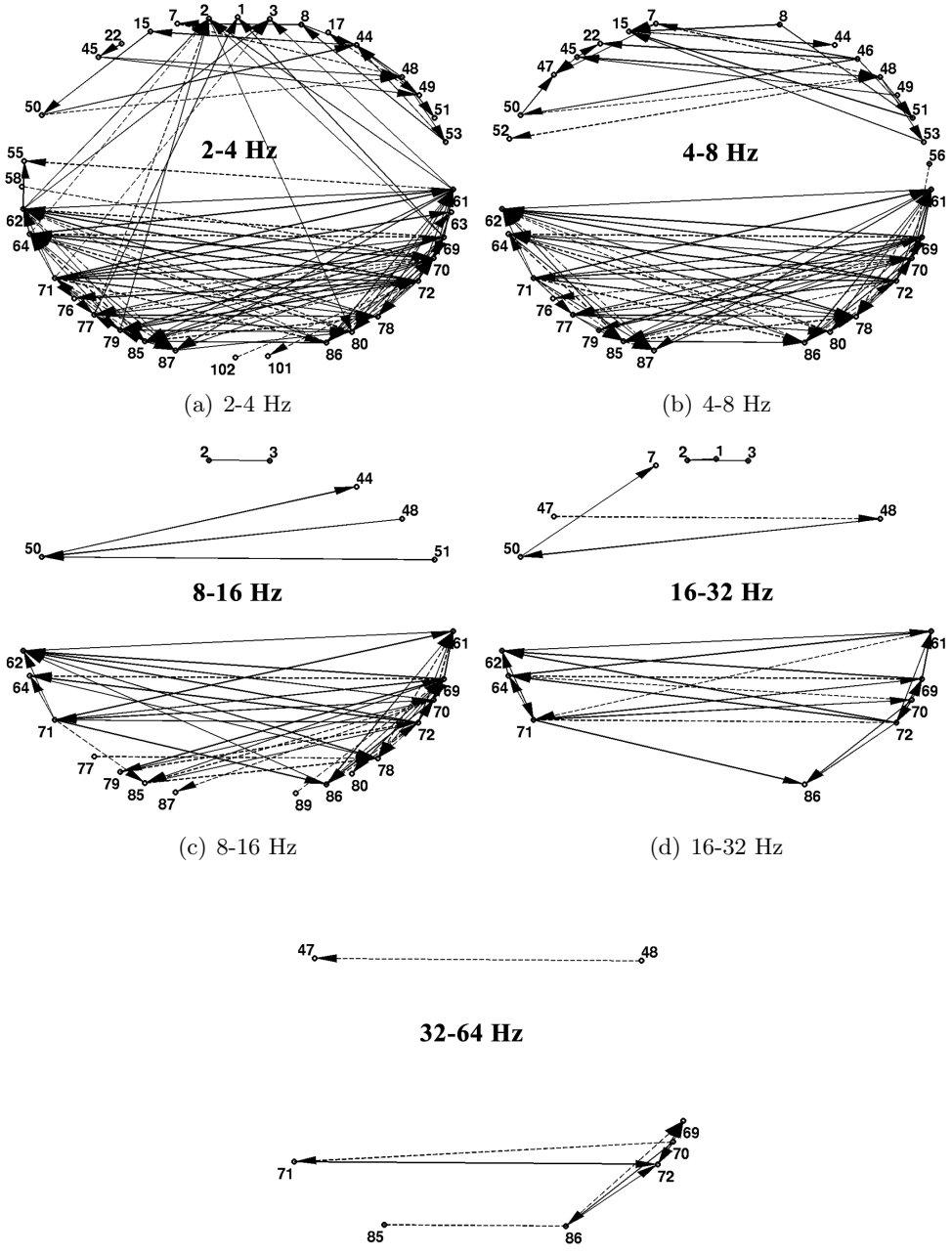
9.1 Results

The frequency contribution of IED time intervals from the iEEG recordings studied here is mostly in the range [2 60] Hz (section 6.2.1, page 52). Therefore, the dDCGs are calculated for five frequency bands in this range: 2-4, 4-8, 8-16, 16-32, and 32-64 Hz. These frequency bands can be assigned to EEG rhythms as following: 2-4 Hz: delta, 4-8 Hz: theta, 8-16 Hz: alpha, 16-32 Hz: beta, and 32-64 Hz: gamma. The related directed DCGs are shown in Figures 9.1 and 9.2 for iEEG analysis of P2 and P3, respectively. The nodes are iEEG bipolar channels (bivariate derivations of the leads of each electrode) and edges represent connections between nodes. The solid and dashed lines show the positive and negative connections, respectively (defined in section 6.2.3, page 57).

9.2 Choice of τ^{max}

The variation of calculated MMCC (6.5)-(6.6) for different time windows (IED or non-IED time intervals) can be due to the estimation error and to the variability of the signal pair couplings between two brain regions in time and frequency. Since our signals are non-stationary, the estimation of the variance of MMCC is very complicated. To overcome this difficulty, we used multiple testing based on permutation to estimate the distribution of the test statistic (6.8) related to MMCC providing a reliable procedure for detecting the statistically significant connections (section 6.2.3, page 58). For a better estimation of MMCC, we studied experimentally the effect of the different parameters on the variance of MMCC estimation (appendix G, page 151).

In the following, first we explain the effect of maximum number of time lags ($\tau^{max} \geqslant |\tau|$) in MMCC estimation [60, 203], and how τ^{max} is experimentally chosen for the calculation of DCG.



(a) 2-4 Hz

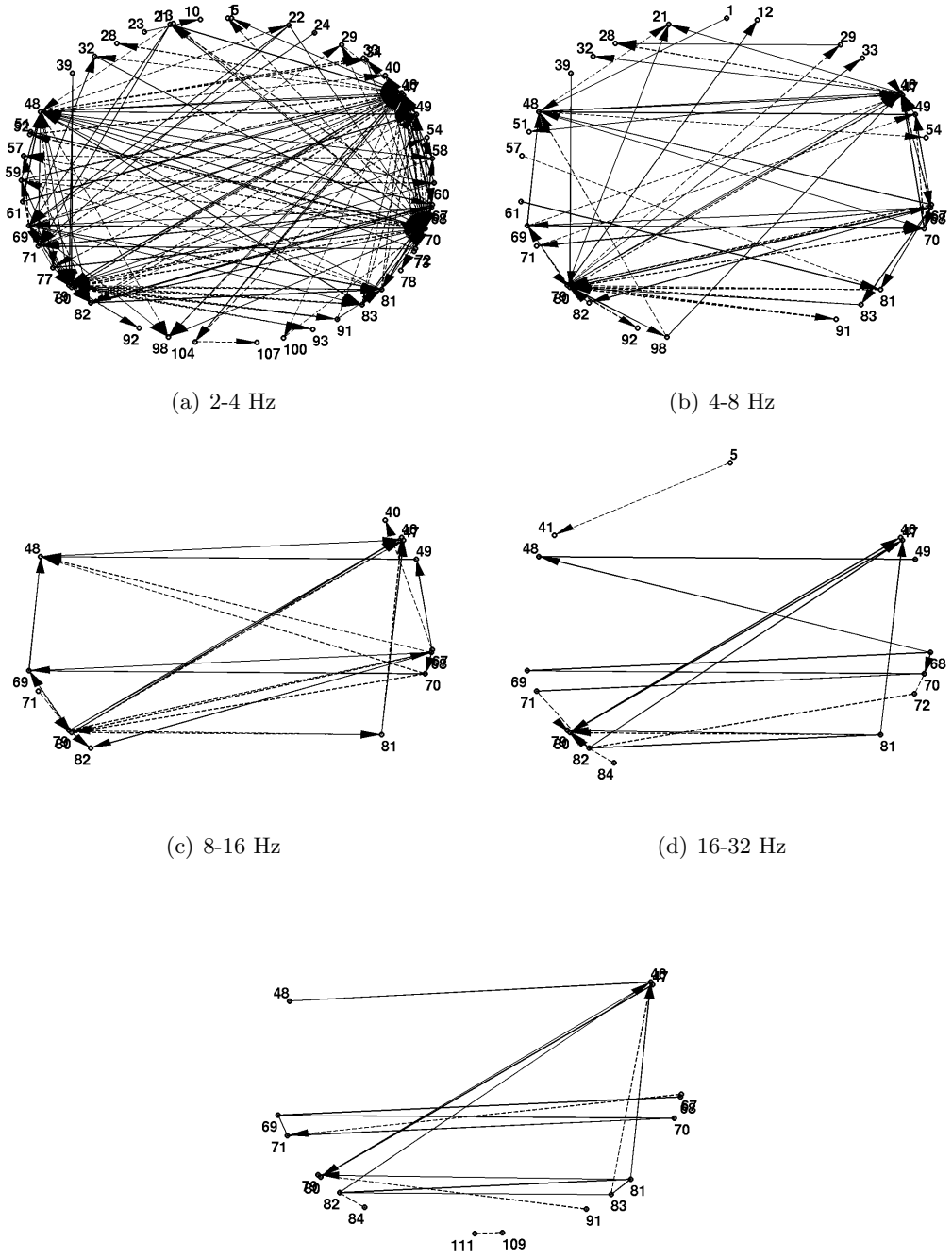
(b) 4-8 Hz

(c) 8-16 Hz

(d) 16-32 Hz

(e) 32-64 Hz

Figure 9.1: dDCG related to (a) 2-4 Hz, (b) 4-8 Hz, (c) 8-16 Hz, (d) 16-32 Hz, and (e) 32-64 Hz for P2. The nodes are iEEG bipolar channels and edges represent connections between nodes. The solid and dashed lines show the positive and negative connections, respectively. The location of nodes do not correspond to their real spatial locations for a better visualization of the graphs.



(a) 2-4 Hz

(b) 4-8 Hz

(c) 8-16 Hz

(d) 16-32 Hz

(e) 32-64 Hz

Figure 9.2: dDCG related to (a) 2-4 Hz, (b) 4-8 Hz, (c) 8-16 Hz, (d) 16-32 Hz, and (e) 32-64 Hz for P3. The nodes are iEEG bipolar channels and edges represent connections between nodes. The solid and dashed lines show the positive and negative connections, respectively. The location of nodes do not correspond to their real spatial locations for a better visualization of the graphs.

Next, the choice of this parameter in time causality estimation is discussed.

9.2.1 Choice of τ^{max} in the estimation of MMCC

Generally, in cross-correlation computation (6.5), τ^{max} is suggested to be large enough to include the true lag between a signal pair. Contrarily, increasing the number of time shifts (τ^{max}), inflates the confidence interval of correlation estimation for non-stationary fractionally differenced signals [89] (appendix H, page 157). Although non-stationary signals do not hold this particular property in this work, the experimental results for MMCC estimation confirmed the above statement, i.e. increase of maximum number of time lags, or decrease of number of samples may increase the variance of MMCC estimation. Here we explain our experimental results for non-stationary signals.

The similarity percentage is defined as the normalized sum of common number of significant or non-significant t-values over number of possible connections. The similarity percentage between DCGs based on $\tau^{max} = 27$ and 100 samples (the reason of these choices are explained later in this section) are compared for two patients (P1 and P2). Contrary to $\tau^{max} = 100$, less empirical variance of MMCC and more number of connections is provided by $\tau^{max} = 27$, especially for lower frequencies where the similarity percentage decreases for both patients. The empirical variance is the variance of estimated MMCC of different IED or non-IED time intervals along the processed time signal. One reason for increase of variance can be the decrease of the number of samples especially in lower frequencies where the number of independent samples decreases.

Eventually a proper selection of τ^{max} can be the smallest maximum number of time lags examined in cross-correlation analysis including the true lag [203]. Satisfying this condition requires the physiological knowledge about the time lags between interictal events recorded from different recording electrode leads. The duration of propagation of action potentials along axons can be assumed less than 40 msec and multi-synaptic transmission within a cortical region may lead to a delay less than 30 msec. As in EEG the post-synaptic potentials are recorded, the delay between IEDs can be assumed as the propagation delay of action potentials (maximum 40 msec) plus the synaptic delay (maximum 30 msec) based on [204]. 70 msec can be considered as the maximum delay of a direct connection. However the connections are not necessarily direct which this increases the upper limit of the time lag. This physiological lag between recorded IED events from different brain regions can be at most 100-200 msec i.e. about 50-100 samples with $f_s = 512$ Hz, typically less than 50 msec i.e. about 25 samples based on [22].

In the calculation of DCG (section 6.2, page 51), the aim is to estimate the maximum coupling between each signal pair during IED and non-IED time intervals. In this step, the smallest maximum number of time lags examined in cross-correlation analysis including the assumed true lag (about 25 samples) [22] is selected equal to 27 samples. Although we observe high similarity percentages between DCGs-based 27 and 100 samples and almost similar final results (estimated ℓIED regions), we prefer $\tau^{max} = 27$. Using $\tau^{max} = 27$ provides more probable connections (less conservative) and less time consuming procedure in the calculation of DCG. We prefer to be less conservative and more powerful in the calculation of DCG to identify as

many discriminated connections as possible. Once such connections are found, a higher level processing (section 6.3, page 61 and chapter 7, page 65) is adopted to extract nodes with optimum emittance contribution.

9.2.2 Choice of τ^{max} in the estimation of time causality

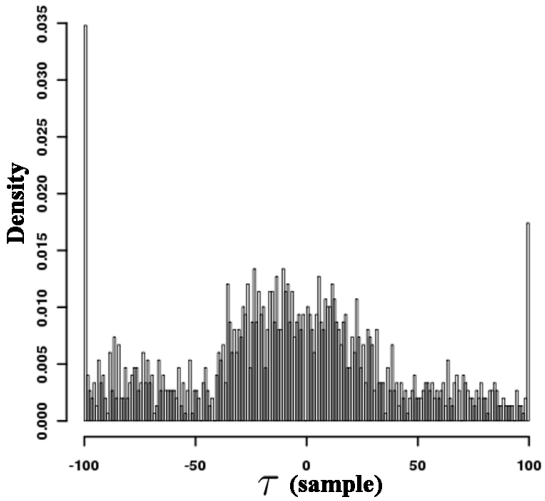
The time lag in which the absolute maximum of cross-correlation function occurs (6.12) is considered as the time causality between underlying signals. For the estimation of time causality, the cross correlation is calculated between MODWT coefficients pair of the whole length of selected data for processing (infinitely long signals). The mean of the length of selected data for processing (Table 8.2) over all of patients is 55.44 minutes or 1 703 117 samples. Large number of samples is considered for a better estimation of MMCC.

The histogram of the τ^*_{ab} for the edges of DCG ($a, b \in \mathbf{V}$) related to P1 for two different values of τ^{max} (16 samples and 100 samples) are compared in Figure 9.3 for 2-4 Hz and 4-8 Hz. Each bar of these histograms demonstrates the number of edges which their τ^*_{ab} is equal to the time sample of the bin (the width of each bin is equal to one time sample) over the number of edges of the graph. The two peaks of each side of the histograms may demonstrate the number of edges that their direction could not be estimated. In the histograms related to $\tau^{max} = 16$ samples (parts (c) and (d) of Figure 9.3), the two peak bars at 16 and -16 are corresponded to 59% and 30% of the edges for 2-4 HZ and 4-8 Hz, respectively. It shows that for example in 4-8 Hz, we could not estimate the direction of the edges for 30% of them. These percentages for $\tau^{max} = 100$ samples is equal to 5% and 1% for 2-4 Hz and 4-8 Hz, respectively. In the estimation of ℓIED regions, the direction of the edges which their τ^*_{ab} is equal to τ^{max} are not considered.

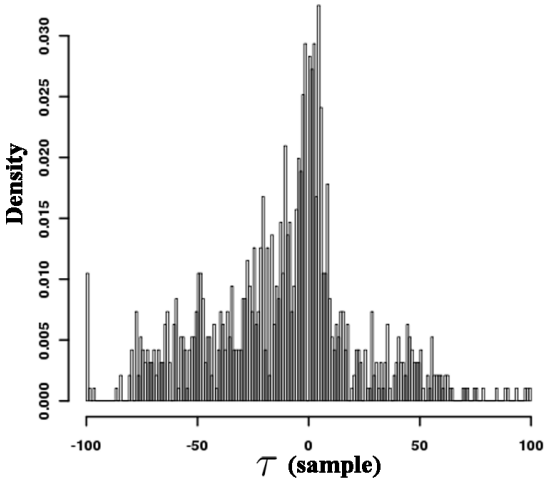
Increasing the τ^{max} up to the physiological constraint provides more information to estimate the τ^*_{ab}. The estimated ℓIED regions for all of the patients are compared for $\tau^{max} = 27$ and 100 samples. The more the histogram of τ^*_{ab} (based on $\tau^{max} = 100$ samples) is concentrated on lower time delay values, the less the estimated ℓIED regions are changed between $\tau^{max} = 27$ and 100 samples. In detail, the set of ℓIED regions are not changed for P3, P4, and P5, while this result is changed for P1 and P2. For P1 and P2, the ℓIED regions based on $\tau^{max} = 27$ samples include the result related to $\tau^{max} = 100$ samples plus two extra regions. One of the extra regions related to $\tau^{max} = 27$ samples for P1 is not removed during surgery. Since the patients are all seizure-free after resective surgery, this result reveals the decrease of precision of the method for $\tau^{max} = 27$ (for P1). The ℓIED regions related to $\tau^{max} = 27$ include the results related to $\tau^{max} = 100$ since only mostly the directions of the edges of lower frequencies are affected by changing the τ^{max}. Note that ℓIED regions are estimated considering the LI values of dDCG of all of frequency bands.

By using $\tau^{max} = 27$ samples, we cannot determine the time causality for a part of the edges of DCG (no maximum cross-correlation can be found in the range $[-\tau^{max}\ \tau^{max}]$), therefore the estimation of ℓIED regions would be based on the rest of the edges of DCG. In other words dDCG based on $\tau^{max} = 100$ samples has more common number of edges with DCG compared to $\tau^{max} = 27$ samples. Therefore, the dDCG used for the estimation of ℓIED regions based on

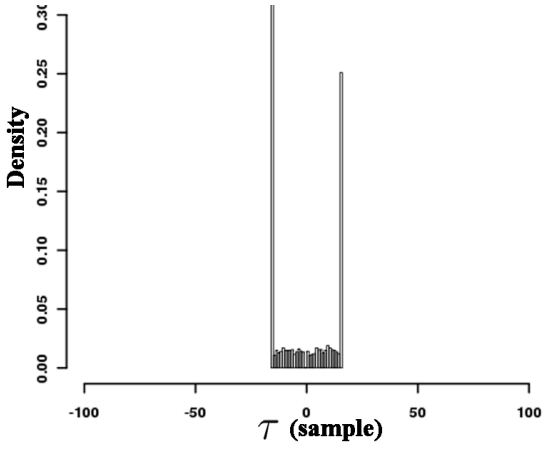
$\tau^{max} = 100$ samples is more similar to its related DCG. Consequently, in the time causality estimation, we chose $\tau^{max} = 100$ samples that is the upper limit of physiological time lag [22].



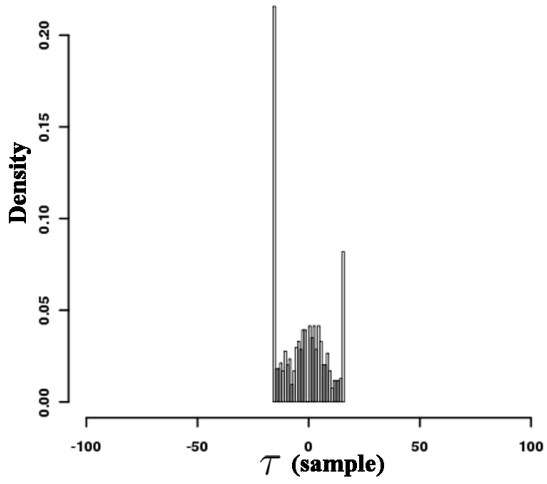
(a) 2-4 Hz, $\tau^{max} = 100$ samples



(b) 4-8 Hz, $\tau^{max} = 100$ samples



(c) 2-4 Hz, $\tau^{max} = 16$ samples



(d) 4-8 Hz, $\tau^{max} = 16$ samples

Figure 9.3: The histogram of τ^*_{ab} values of the edges of DCG for $\tau^{max} = 16$ samples for (a) 2-4 Hz and (b) 4-8 Hz. The histogram related to $\tau^{max} = 100$ samples for (c) 2-4 Hz and (d) 4-8 Hz.

9.3 Reliability of DCG

In this section the reliability of DCG based on the effect of balanced and unbalanced number of IED and non-IED time intervals and different sets of IED time intervals are explained. Furthermore the significance of each edge of DCG is described.

9.3.1 Effect of large balanced and unbalanced number of IED and non-IED time intervals

The effect of large balanced and unbalanced (section 6.2.3, page 58) number of time intervals (L^1 and L^2) in multiple testing is verified by comparing the DCG constructed with unbalanced number of time intervals and ten (an arbitrary choice) DCGs with random balanced number of time intervals. The balanced number of time intervals is fixed equal to the smaller number of IED and non-IED time intervals ($min(L^1, L^2)$). Say $L^2 = min(L^1, L^2)$, therefore L^2 IED time intervals are chosen randomly among L^1 IED time intervals and the related DCG is constructed. We repeated this procedure ten times. The similarity percentage is calculated between the original DCG-based unbalanced number of time intervals and ten recalculated DCG-based balanced number of time intervals. The mean of these ten percentages for P1 at frequency band equal to 4-8 Hz is 97.2%± 0.3. This result shows that the large unbalanced number of time intervals is not problematic for the permutation test [145].

9.3.2 Effect of different sets of IED time intervals

To verify the reliability of DCGs for different random sets of IED time intervals, the similarity percentage between each pairs of ten recalculated DCGs (ten recalculated DCG-based balanced number of time intervals in previous paragraph) is obtained. The mean of these 45 similarity percentages for the same patient and frequency band is 96.9%± 0.25.

9.3.3 Significance of each edge of DCG

In DCG, an edge is inserted if the couplings of the related signal pairs change significantly between IED and non-IED time intervals. To test the significance of the couplings between signal pairs during IED or non-IED time intervals, we do a pretest before the multiple testing (Figure 6.6, section 6.2.3, page 57). The couplings of each connection ($\mathbf{c}^l_{:n}$) are tested to see if they are greater than a threshold equal to 0.3 with the following test:

$$\begin{cases} H^n_0: & \mu^l_n \leqslant 0.3 \\ H^n_1: & \mu^l_n > 0.3 \end{cases} \tag{9.1}$$

where μ^l_n is the average of $\mathbf{c}^l_{:n}$ (the nth column of matrix $\mathbf{C}^l$). The bootstrap method is used for the estimation of p-values for 1000 repetitions and false positive rate (α) is set equal to 0.05. The connections whose couplings are significantly greater than the threshold for both IED and non-IED states are entered the multiple testing to define the DCG connections. The resulted DCG is compared with the DCG based on multiple testing without the pretest. The similarity percentages between these two DCGs for P1 and P2 in different frequency bands are reported in Table 9.1.

These similarity percentages show that most of the connections of the DCG have significantly large couplings in both IED and non-IED time intervals. However DCG is designed to detect the connections which their couplings change significantly between IED and non-IED time intervals,

Table 9.1: Significance of each edge of DCG. The effect of adding a pretest before multiple testing in the procedure of DCG inference (section 6.2.3, page 57).

Similarity percentage	2-4	4-8	8-16	16-32	32-64
P1	96.83	92.53	94.45	97.65	99.27
P2	100	99.95	99.54	99.69	99.87

consequently based on the results of Table 9.1, the pretest is not necessary and will not be used.

Chapter 10

Estimated ℓIED regions

Once the dDCGs related to different frequency bands are obtained, we need to quantify the nodes of the dDCGs for the estimation of the ℓIED regions. To this end two new measures (local information and total global efficiency) are proposed (see section 7.2, page 65). In this chapter first these new measures are compared with classic measures. Since comparison between estimated ℓIED regions based on local information revealed more informative and efficient results compared to other graph measures studied, we used this measure for quantification of dDCGs and eventually for the estimation of ℓIED regions. Second, the reliability of estimated ℓIED regions based on the choice of maximum time lag is explained. Third, the estimated ℓIED regions based on Pareto and neighbor-Pareto optimization algorithms in the multiple graph analysis method are compared. Finally, we compared the estimated ℓIED regions based on local information (using Pareto optimization algorithm) with other methods of SOZ detection. Two methods are considered in this comparison. The first method is SOZ detection based on visually inspection and second method is SOZ detection using induced ictal iEEG recoding by stimuli [1].

To summarize, in this chapter we explain about the following issues:

1. Comparison of estimated ℓIED regions based on different digraph measures
2. The reliability of estimated ℓIED regions based on the choice of maximum time lag
3. Comparison between the estimated ℓIED regions based on Pareto and neighbor-Pareto optimization algorithms
4. Comparison of estimated ℓIED regions with other methods of SOZ detection.

10.1 Comparison of estimated ℓIED regions based on different digraph measures

In this section the efficiency of usual digraph measures [5, 77, 81] like global efficiency (E_{glob}), local efficiency (E_{loc}), and total degree (k_{tot}) are compared with two proposed measures: local information (LI) and total global efficiency (E_{tglob}).

For each patient, five dDCGs related to five frequency bands ($J = 5$) from 2-64 Hz (2-4, 4-8, 8-16, 16-32, and 32-64 Hz) are obtained. These 25 dDCGs (5 dDCGs for each of 5 patients) are characterized with five digraph measures (three classic ones and two proposed measures).

For any measure studied, we associate matrix $\mathbf{Q} \in \mathbb{R}^{J\times N} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N]$, so that each column $\mathbf{q}_n = [q^1[n], q^2[n], \ldots, q^J[n]]^T$, $n = \{1, \ldots, N\}$ corresponds to the measure values of node n for J different frequency bands. E_{glob}, E_{tglob}, E_{loc}, k_{tot}, and LI are calculated for each of dDCGs. The negative E_{tglob}, negative k_{tot}, and negative LI values are set to zero since we are not interested in sink nodes. LI values are normalized to their maximum.

For the estimation of ℓIED regions based on each measure, the multi-objective optimization method [149,150] is applied on matrix $\mathbf{Q}$ of each measure. To evaluate the efficiency of different graph measures for the identification of source and sink nodes, the estimated ℓIED regions based on different digraph measures are compared with visually inspected SOZ (vSOZ) by the epileptologist (details are explained in section 10.4.1, page 110). Though the SOZ is theoretically the region to be removed in candidates to resective surgery, it is not easy to be defined in many patients partly because of the difficulties regarding recording seizure periods. Thus it is valuable to wonder if one can predict the SOZ by estimating the ℓIED regions.

The estimated ℓIED regions for five patients are presented in Table 10.1. The vSOZs are also reported for comparison. For each patient, the column labels correspond to either ℓIED regions estimated by digraph measures or the vSOZs of the patient.

The resected regions always include the vSOZ, and all of the patients are seizure free after the resective surgery. Therefore measures which suggest ℓIED nodes congruent with vSOZ are preferred, i.e. more number of common regions (true positives, or TP), smaller number of uncommon regions (false negatives, or FN), and smaller number of ℓIED regions which are not included in vSOZ (false positives, or FP). The precision and the sensitivity defined as $(\#\text{TP})/(\#\text{TP}+\#\text{FP})$, and $(\#\text{TP})/(\#\text{TP}+\#\text{FN})$, respectively are reported in Tables 10.2 and 10.3. The number of regions is denoted by #. Although both FP and FN are aimed to be minimized, due to the existing trade off between these errors, we prefer the measures providing less FP compared to FN. In fact FP indicate normal brain regions that are wrongly detected as SOZ which has the risk of removing normal regions in resective surgery. Conversely, FN show the missed SOZ which has the risk of a second surgery supposing the presurgery evaluations are based on only estimated ℓIED regions. However, the presurgery evaluations are based on different complementary clinical knowledge including iEEG, fMRI, semiology, etc. Estimated ℓIED regions are aimed to provide the information extracted from iEEG analysis based on the functional connectivity related to interictal events. Therefore since ℓIED regions provide complementary information for conducting the resection, we prefer to estimate the ℓIED regions with minimum FP to increase the precision of the estimation in comparison with its sensitivity.

Furthermore, although in this study all patients being seizure free after surgery reveals that resected regions (based on vSOZ) included the essential regions generating the seizures, currently there is no way to determine how large was the removed regions during surgery. Therefore it is valuable to study the properties of the uncommon nodes between vSOZ and the set of estimated ℓIED nodes (FN) based on different measures to analyze their properties more profoundly.

LI measure is designed for source and sink node identification from digraphs, while E_{glob}, E_{loc}, and k_{tot} are not exactly suited for this purpose. In the following we explain about the comparison results of LI measure with each of these usual measures.

10.1.1 Local efficiency

The estimated ℓIED regions based on E_{loc} are consistent with vSOZ for P2, P3 and P4, while this is not true for P1 and P5. For P1, the estimated ℓIED based on E_{loc} includes the vSOZ with the addition of five other regions (FP): {anterior cingulate gyrus, anterior superior temporal gyrus, insular cortex, white matter mesial frontal, frontal operculum}. For avoiding long Table, we do not report all of ℓIED regions detected by E_{loc} for P1 in the Table 10.1. The local efficiency measure was not applicable for P5 due to great sparseness of related graphs. Here we explain the reason of this sparseness. P5 is a very specific patient, treated for right operculo-insular cavernous malformation after the first resective surgery. After nine years, the patient got reflex seizures [205]. The iEEG recordings for the second presurgery evaluations include extensive sampling of the insular, opercular and temporal cortices. The SOZ is recognized visually in middle short gyrus of right insula and this region is removed in the second resective surgery. The patient is seizure free after this surgery. dDCG of this patient from iEEG recordings before second surgery include one edge between two adjacent electrode leads in middle short gyrus of right insula in the three frequency bands: 8-16, 16-32 and 32-64 Hz. For such sparse digraphs, E_{loc} is not applicable.

The definition of E_{loc} as the evaluation of the efficiency of the connections between the first-order outgoing neighbors of each node is not well suited for the source and sink node detection. Indeed a node with high E_{loc} may not be a good candidate for a source node. Furthermore, E_{loc} considers neither the incoming paths nor the amount of information of each edge. Finally, this measure may suggest regions that should not be removed during surgery (P1), which reduces its precision. Although E_{loc} is highly sensitive, we prefer the other measures with higher precision.

10.1.2 Total degree

For most of the patients (all of the patients except P4), all of the ℓIED regions based on k_{tot} are included in the set of related vSOZ which leads to high mean precision over patients. However this measure provides a smaller number of related vSOZ than LI (less TP) which makes it less powerful. Total degree considers both outgoing and incoming connections, but without weights (the amount of information). Two nodes with the same k_{tot} may have different amounts of local information (Table 10.4, first and second columns), since connections may carry different amounts of information. To benefit from such knowledge (amount of information), LI measure was introduced.

Table 10.1: Comparison between visually inspected seizure onset zones (vSOZ) and estimated leading IED (ℓIED) regions for five patients (P1 to P5) based on different measures: E_{glob}: global efficiency; E_{tglob}: total global efficiency; E_{loc}: local efficiency; k_{tot}: total degree; LI: local information. For each patient, the column labels correspond to either ℓIED regions estimated by digraph measures or the vSOZ of the patient. Each cross shows that the region related to its column is detected by the digraph measure related to its row or by the vSOZ. amyg: amygdala; ant/post/m: anterior/posterior/mesial; CG: cingulate gyrus; entCx: entorhinal cortex; fusi: fusiform gyrus; HC: hippocampus; Ins: insula; midInsG: middle short gyrus of insula; pHcG: parahippocampal gyrus; T: temporal; TP: temporal pole.

P1	antHC	postHC	amyg	pHcG	mTP			
E_{glob}			×	×	×			
E_{tglob}			×	×	×			
E_{loc} [1]	×	×	×	×	×			
k_{tot}			×	×				
LI	×		×	×				
vSOZ	×	×	×	×	×			
P2	**antHC**	**postHC**	**amyg**	**pHcG**	**Ins**	**fusi**		
E_{glob}	×							
E_{tglob}	×			×	×	×		
E_{loc}	×	×		×				
k_{tot}	×			×				
LI	×							
vSOZ	×	×	×	×				
P3	**antHC**	**postHC**	**pHcG**	**TP**	**postT4**			
E_{glob}	×	×	×					
E_{tglob}	×	×		×	×			
E_{loc}	×	×	×					
k_{tot}	×	×						
LI	×	×						
vSOZ	×	×	×					
P4	**antHC**	**postHC**	**amyg**	**entCx**	**mTP**	**antCG**	**postT1**	**T4**
E_{glob}	×	×	×	×		×		
E_{tglob}	×	×	×	×	×	×	×	×
E_{loc}	×	×		×				
k_{tot}		×	×	×				
LI	×	×	×	×				
vSOZ	×	×	×	×	×			
P5	**midInsG**							
E_{glob}	×							
E_{tglob}	×							
E_{loc}	NA							
k_{tot}	×							
LI	×							
vSOZ	×							

Table 10.2: Precision of different digraph measures for five patients. NA: not applicable.

Precision	E_{glob}	E_{loc}	k_{tot}	E_{tglob}	LI
P1	1	0.5	1	1	1
P2	1	1	1	0.5	1
P3	1	1	1	0.5	1
P4	0.8	1	1	0.62	1
P5	1	NA	1	1	1
mean	0.96	0.88	1	0.72	1

Table 10.3: Sensitivity of different digraph measures for five patients. NA: not applicable.

Sensitivity	E_{glob}	E_{loc}	k_{tot}	E_{tglob}	LI
P1	0.6	1	0.4	0.6	0.6
P2	0.25	0.75	0.5	0.5	0.25
P3	1	1	0.67	0.67	0.67
P4	0.8	0.6	0.6	1	0.8
P5	1	NA	1	1	1
mean	0.73	0.84	0.63	0.75	0.66

Table 10.4: Comparison of E_{loc}, E_{glob}, E_{tglob}, $k_{tot} = k_{out} - k_{in}$, and LI values of node 46 located in left parahippocampal gyrus of P3. LI values are normalized to their absolute maximum.

	2-4	4-8	8-16	16-32	32-64
E_{loc}	0.02	0	0	0	0
E_{glob}	0.5	0.35	0	0	0.06
E_{tglob}	0.04	-0.05	-0.33	-0.20	-0.12
$k_{out} - k_{in}$	$12 - 12$	$5 - 5$	$0 - 3$	$0 - 1$	$1 - 3$
LI	-0.67	-0.09	-0.08	-0.02	-0.01

10.1.3 Total global efficiency

The total global efficiency is designed to add the information of incoming paths to the knowledge extracted from outgoing ones. Although it was introduced to refine the global efficiency, the results of this measure are not satisfactory: despite the inclusion of the proposed regions based on this measure in vSOZ, it provides additional regions that do not correspond to the related vSOZ. In other words, although adding the information of incoming paths has increased the sensitivity of total global efficiency in comparison with global efficiency, this information does not increase its precision (Tables 10.2 and 10.3).

This result shows that considering the incoming paths in addition to the outgoing ones in global efficiency is not sufficient for increasing its precision in source and sink node detection. Consequently, a node with high total global efficiency may not be a good candidate for a source node.

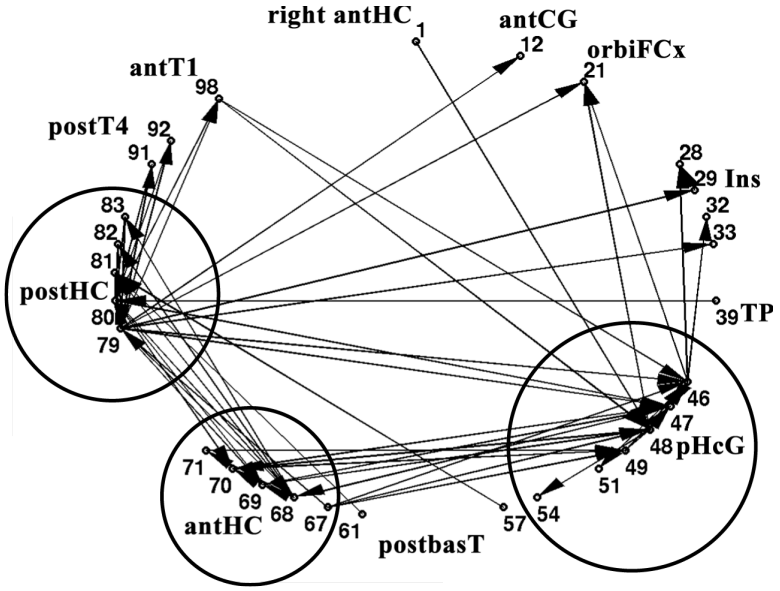
10.1.4 Global efficiency

Global efficiency and LI, both seem to have congruent results with vSOZ (Tables 10.1-10.3), but LI can provide more information about the emittance contribution of the nodes than global efficiency, since E_{glob} does not take into account the incoming paths and the information carried by each edge.

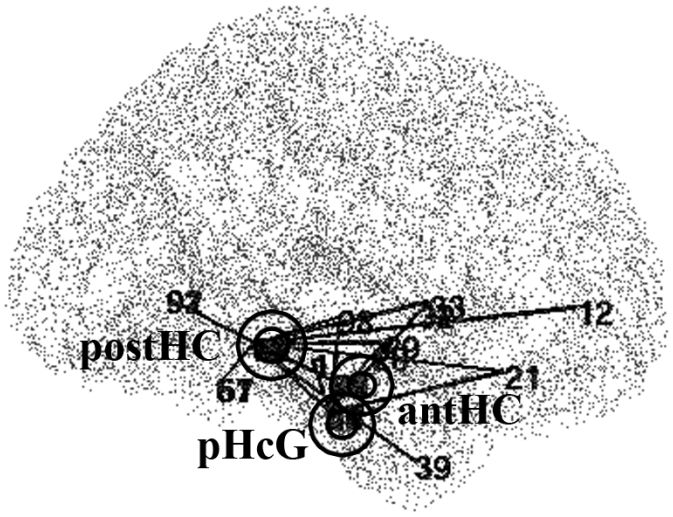
Source and sink node detection

The differences between LI and global efficiency is highlighted in the nodes with both several outgoing and incoming connections and with negative or zero amount of local information, i.e. incoming information is greater or approximately equal to outgoing information. For such nodes, LI values are negative or zero, while E_{glob} can be high if we reach from these nodes to the rest of digraph through the short-length outgoing paths. To show one example of such nodes, we focus on the iEEG analysis of P3. In this patient, node 46 located in left parahippocampal gyrus is included in the estimated ℓIED nodes by E_{glob}, while it is not estimated by LI. Here we explain the details about this node in the dDCG of different frequency bands. We analyze the role of node 46 based on comparing its different measure values (Table 10.4).

We remind that E_{glob}, E_{loc}, E_{tglob}, k_{tot} ($k_{out} - k_{in}$) and LI are in the range $[0, 1]$, $[0, 1]$, $[-1, 1]$, $[-(N-1), N-1]$, and $[-\infty, \infty]$ respectively. For comparing the values of different measures, LI values are normalized to their absolute maximum (range $[-1, 1]$) in Table 10.4. The subgraph of node 46 ($\mathbf{G}_{46^+}$) from dDCG of P3 in 4-8 Hz (Figure 10.1) is depicted in Figure 10.2. Node 46 has five outgoing and five incoming connections in this frequency band. The incoming and outgoing information are approximately equal which leads to low LI value. Furthermore we can get to 23 nodes among 29 nodes from node 46 through short-length outgoing paths, which gives relatively high E_{glob} comparing to other nodes of the digraph in this frequency band. Besides from 20 nodes of the digraph, we can also get to node 46 through short-length incoming paths. Global efficiency considers outgoing paths, therefore this measure is blind to the knowledge about the



(a)



(b)

Figure 10.1: (a) dDCG for P3 in 4-8 Hz, (b) digraph of (a) plotted in real channel coordinates superimposed on the 3D anatomical mesh (sagittal view). The Figure of part (b) is added for an anatomical representation to show where the graph and especially the three important regions of the dDCG (anterior/posterior hippocampus and parahippocampal gyrus) are located in the brain, while the details of the dDCG can be seen in part (a). These three regions are enclosed with circles in parts (a) and (b). The 3D rotatable version of Figure of part (b) is available at http://www.gipsa-lab.inpg.fr/∼ladan.amini/mes_images/dDCGPatient3scale4_8Hz.fig. The Matlab software is needed to see this rotatable Figure. ant/post/bas/m: anterior/posterior/basal/mesial; CG: cingulate gyrus; Cx: cortex; F: frontal; HC: hippocampus; Ins: insula; orbiF: orbitofrontal; pHcG: parahippocampal gyrus; T: temporal; TP: temporal pole. 12: antCG; 21: orbFCx; 28-33: InsmF1; 39: TP; 46-54: antbasT; 57-61: postbasT; 67-71: antHC; 79-83: postHC; 91-92: postT4; 98: antT1. All of bipolar channels are located in left side of the brain except bipolar electrode lead 1 (right antHC).

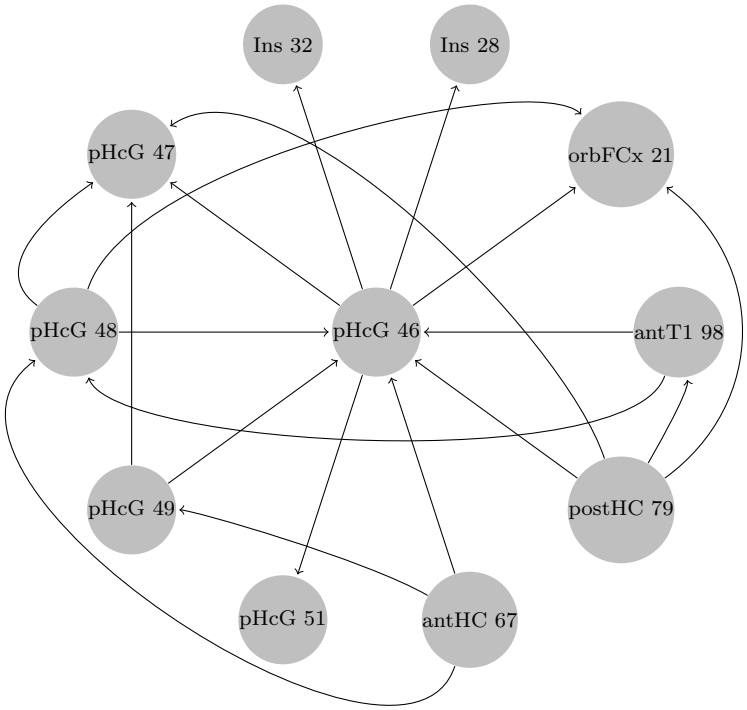


Figure 10.2: Directed subgraph of node 46 located in parahippocampal gyrus ($\mathbf{G}_{46^+}$) from the dDCG related to P3 in 4-8 Hz (Figure 10.1). Each node is represented by its number of bipolar electrode lead and the related brain region. ant/post: anterior/posterior; pHcG: parahippocampal gyrus; Ins: insula; orbiF: orbitofrontal; Cx: cortex; T: temporal; HC: hippocampus.

incoming paths and specially to the amount of information related to each connection, while LI value is based on the outgoing minus incoming amount of information.

Node 46 is selected as a source node by the multiple digraph analysis based on E_{glob} while LI of this node (Table 10.4) shows that it may not be considered as a source node. Comparing LI values with other usual measure values reveals that node 46 might be mostly a transit node than a source one. A transit node exerts approximately all of the information it receives. However distinguishing between source and transit nodes is challenging and might not be easy to be explored neither by E_{glob} nor by visual inspection.

To end with, E_{glob} is not as precise as LI in source nodes detection since the incoming connections and the amount of information are not considered.

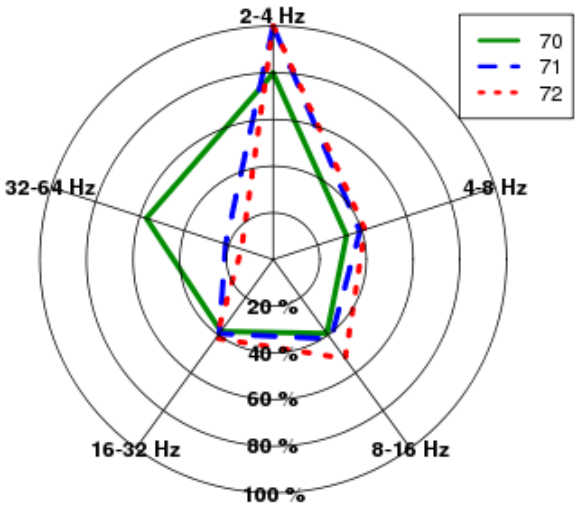
Ranking

Another advantage of LI over E_{glob} lies in ranking the estimated ℓIED nodes in terms of measure values. To explain this issue, we focus on the results of iEEG analysis of P2. This patient is chosen since the related ℓIED nodes based on E_{glob} and LI are the same: nodes 70, 71, and 72 located in left anterior hippocampus (Table 10.1). E_{glob} and LI values of these nodes are reported in Table 10.5 for different frequency bands. E_{glob} and LI values are normalized to their maximum (negative values of LI are set to zero before normalization). To rank the estimated ℓIED nodes, one way can be to order the ℓ_1 norm of related measure values. ℓ_1 norm is calculated for $\mathbf{q}_n$, $n \in \{\text{estimated } \ell\text{IED nodes}\}$ as: $\sum_{j=1}^{J} q^j[n]$, where $q^j[n]$ are nonnegative real numbers. Normalized ℓ_1 norm values of E_{glob} and LI for ℓIED nodes related to P2 are reported in Table 10.6. The greater the ℓ_1 norm the more reliable the ℓIED node. As can be seen in Table 10.6, the ranking based on E_{glob} and LI is not the same. Considering the incoming connections and the amount of information lead to smaller LI values for node 70, therefore this decreases the relevance of this node to the emittance contribution. Conversely, E_{glob} of this node is high since the incoming paths and the importance of the edges are not considered which leads to its high relevance.

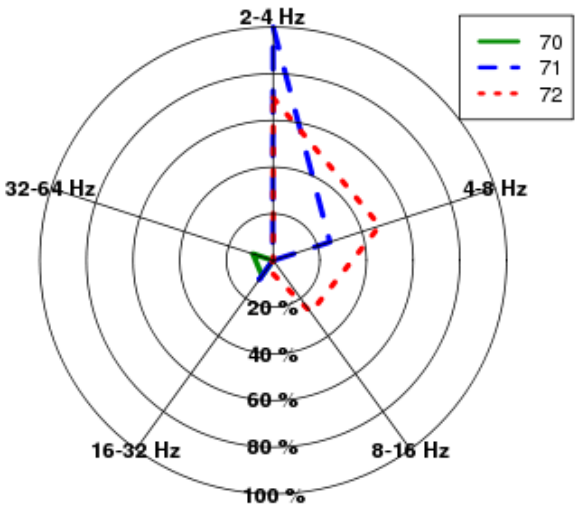
To easily compare the values of E_{glob} and LI for the estimated ℓIED nodes (reported in Table 10.5), these measure values are demonstrated in Figure 10.3. In Figure 10.3, the web plot of $\mathbf{q}_{70}$, $\mathbf{q}_{71}$, and $\mathbf{q}_{72}$ for E_{glob} and LI and for $J = 5$ are depicted. The pentagonals related to nodes 70, 71, and 72 are depicted in solid, dashed and dotted lines, respectively. More expanded pentagonals can refer to greater values of ℓ_1 norm. Based on E_{glob}, node 70 has the greatest ℓ_1 norm, while node 72 has the greatest ℓ_1 norm based on LI measure. To conclude, LI measure is preferred to be used for ranking the leading IED nodes than E_{glob}. This ranking can be valuable for the presurgery evaluations where there are several ℓIED nodes located in different brain regions.

10.1.5 Conclusion

We compared classic digraph measures with introduced LI measure for the source and sink node identification of digraphs. The comparison is evaluated using estimated ℓIED regions from



(a) E_{glob}



(b) LI

Figure 10.3: The demonstration of five dimensional measure vectors $\mathbf{q}_{70}$, $\mathbf{q}_{71}$, and $\mathbf{q}_{72}$ related to P2 for (a) E_{glob} and (b) LI in web plot. The five dimensions correspond to five frequency bands from 2 to 64 Hz.

Table 10.5: Comparison between E_{glob} and LI values of the three ℓIED nodes related to P2 in different frequency bands.

		2-4	4-8	8-16	16-32	32-64
E_{glob}	70	0.68	0.58	0.69	0.67	1
	71	0.85	0.68	0.74	0.69	0.38
	72	0.85	0.72	0.92	0.73	0.25
LI	70	0	0	0	0.08	0.09
	71	1	0.26	0	0.10	0
	72	0.7	0.48	0.27	0.04	0

Table 10.6: Comparison between ranking estimated ℓIED nodes based on E_{glob} and LI using ℓ_1 norm for P2.

	E_{glob}	ℓ_1 norm	LI	ℓ_1 norm
source1	70	1	72	1
source2	72	0.96	71	0.9
source3	71	0.92	70	0.12

dDCGs of different frequency bands. The estimated ℓIED regions based on different measures are compared with visually inspected SOZ by epileptologist. In this perspective, LI measure is more informative in comparison with usual measures and E_{tglob} for the purpose of source and sink node identification. However, comparing the values of different measures like E_{glob}, K_{out}, K_{in}, and LI provides complementary characterizing information based on different definitions for more profound analysis of the role of each node in the digraph and their specifications.

For source and sink distinction, E_{glob} and k_{tot} outperform the conventional measures and E_{tglob}. E_{tglob} does not increase the precision of E_{glob} in the estimation of ℓIED regions, which shows that including the incoming paths may not be sufficient without considering the amount of information of each edge. However, LI is preferred in comparison with E_{glob} and k_{tot}, since it is more informative and skilled for the source and sink node detection and ranking the estimated ℓIED nodes. The power of LI relies on taking into account the amount of information carried by each edge and considering incoming connections in addition to outgoing ones.

10.2 Reliability of estimated ℓIED regions for the choice of τ^{max}

In this section we explain the effect of maximum time lag (τ^{max}) selection in DCG calculation and time causality estimation on the estimated ℓIED regions.

10.2.1 Choice of τ^{max} in DCG calculation

LI measure values may change for two dDCGs based on $\tau^{\max} = 27$ and 100 samples (in DCG calculation) due to the minor changes of the topology of DCG, but nodes with relative high LI values are mostly consistent between these two graphs. Consequently the related ℓIED regions

are mostly similar. However as explained in section 9.2.1 (page 90), $\tau^{\max}$ is set to 27 samples to provide less time consuming and less conservative method for DCG calculation.

In Figure 10.4, we demonstrate the LI values for all of the nodes of P1 in different frequency bands for dDCGs based on $\tau^{\max} = 27$ and 100 samples in dashed and solid lines, respectively. Note that $\tau^{\max}$ is set equal to 100 samples in the estimation of time causality (section 9.2.2, page 91) in both dDCGs of all of frequency bands.

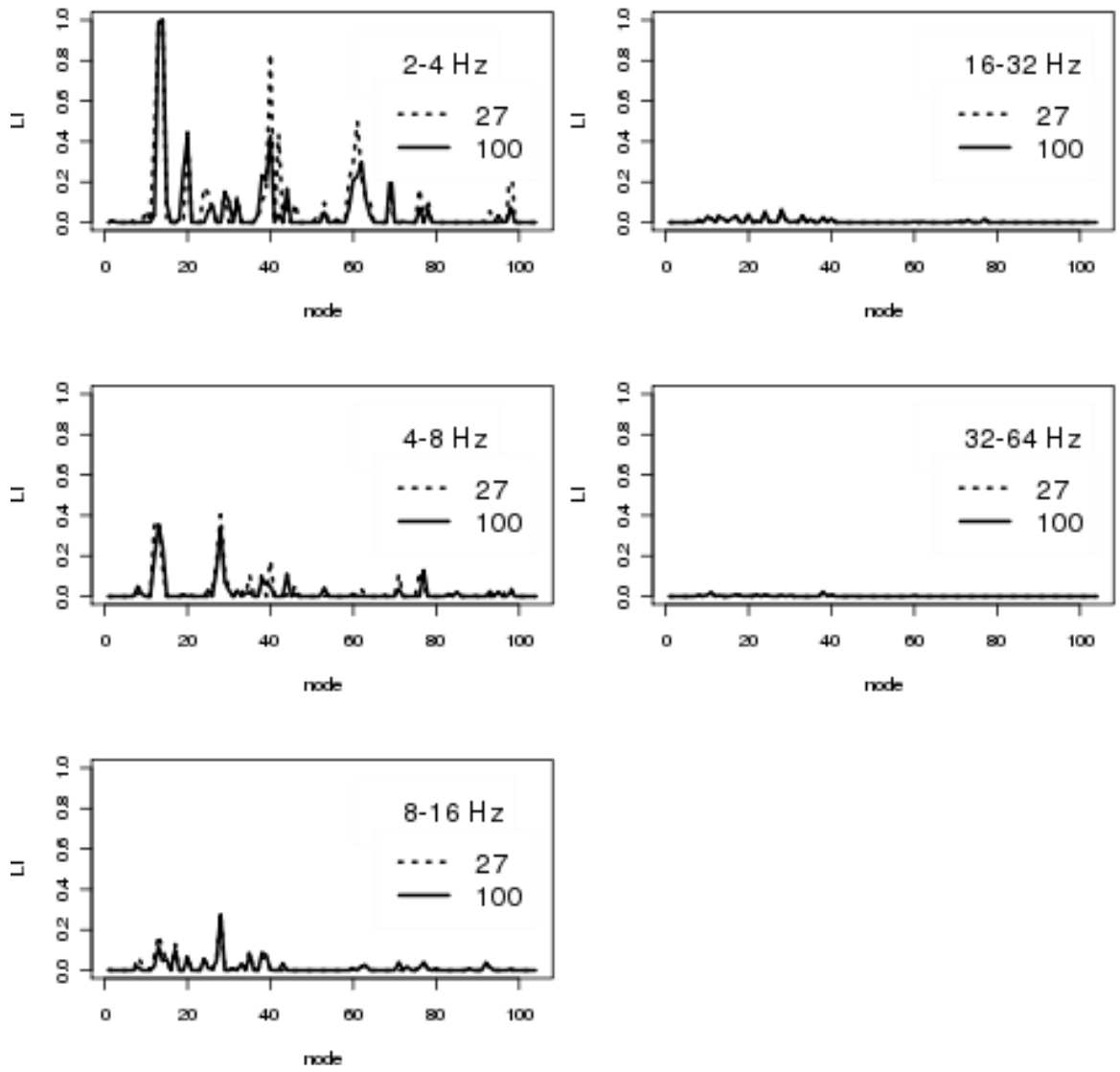


Figure 10.4: Comparison of LI measure values between dDCG based on $\tau^{\max} = 27$ and 100 samples (in DCG calculation) for P1. Note that $\tau^{\max}$ is set equal to 100 samples in the estimation of time causality (section 9.2.2, page 91). Nodes with relative high LI values are consistent in the two dDCGs, in spite of minor changes in the topology of the DCGs especially in lower frequencies.

10.2.2 Choice of τ^{max} in time causality estimation

Here we report the result of estimated ℓIED regions based on the choice of $\tau^{max} = 27$ and 100 samples in time causality estimation (Table 10.7). The estimated ℓIED regions based on $\tau^{max} = 27$ and 100 samples are the same for P3, P4, and P5. For P2, $\tau^{max} = 27$ provides more common regions between estimated ℓIED regions and vSOZ while for P1, it provides an extra region in the estimated ℓIED regions which is not included in the related vSOZ and not even removed during surgery. This result shows that $\tau^{max} = 27$ is not a proper choice which decreases the precision of ℓIED detection for P1 even if it increases the sensitivity of ℓIED detection for P2. As explained in section 9.2.2 (page 91), by choosing $\tau^{\max} = 27$ samples, we cannot estimate the time causality for a part of DCG connections especially in lower frequencies and thus the related DCG and dDCG may have less number of common connections. Therefore for $\tau^{\max} = 27$, the ℓIED detection would be based on a dDCG which is far from its related DCG. As such $\tau^{\max}$ is set to 100 samples in time causality selection.

Table 10.7: Comparison between vSOZ and estimated ℓIED regions for five patients (P1 to P5) based on two choices of τ^{max} (27 and 100 samples) in the time causality estimation. For each patient, the column labels correspond to either ℓIED regions estimated based on $\tau^{max} = 27$ or $\tau^{max} = 100$, or the vSOZ of the patient. Each cross shows that the region related to its column is detected based on the τ^{max} related to its row or by the vSOZ. amyg: amygdala; ant/post/m: anterior/posterior/mesial; CG: cingulate gyrus; entCx: entorhinal cortex; fusi: fusiform gyrus; HC: hippocampus; Ins: insula; midInsG: middle short gyrus of insula; pHcG: parahippocampal gyrus; T: temporal; TP: temporal pole.

P1	antHC	postHC	amyg	pHcG	mTP	Ins
$\tau^{max} = 27$	×	×	×	×		×
$\tau^{max} = 100$	×		×	×		
vSOZ	×	×	×	×	×	
P2	**antHC**	**postHC**	**amyg**	**pHcG**		
$\tau^{max} = 27$	×		×	×		
$\tau^{max} = 100$	×					
vSOZ	×	×	×	×		
P3	**antHC**	**postHC**	**pHcG**			
$\tau^{max} = 27$	×	×				
$\tau^{max} = 100$	×	×				
vSOZ	×	×	×			
P4	**antHC**	**postHC**	**amyg**	**entCx**	**mTP**	
$\tau^{max} = 27$	×	×	×	×		
$\tau^{max} = 100$	×	×	×	×		
vSOZ	×	×	×	×	×	
P5	**midInsG**					
$\tau^{max} = 27$	×					
$\tau^{max} = 100$	×					
vSOZ	×					

10.3 Comparison of estimated ℓIED regions based on Pareto and neighbor-Pareto optimization algorithms

Here we compare the results of estimated ℓIED regions based on Pareto and neighbor-Pareto optimization algorithms in multiple graph analysis (section 7.3.2, page 70). In neighbor-Pareto optimization algorithm, the nodes in the neighborhood of Pareto front nodes are also considered. Here we report the estimated ℓIED regions based on Pareto and neighbor-Pareto optimization algorithms in Table 10.8. The parameter ϵ of neighbor-Pareto optimization algorithm is set equal to 0.3 in this Table. The normalized square distance, $\frac{\|\mathbf{LI}^{\cdot}[n]-\mathbf{LI}^{\cdot}[n_0]\|_2^2}{\|\mathbf{LI}^{\cdot}[n]\|_2 \ \|\mathbf{LI}^{\cdot}[n_0]\|_2}$ in this algorithm is normalized in the range $[0, 100]$. For $\epsilon = 0$, the neighbor-Pareto optimization algorithm performs the same as Pareto optimization algorithm. For P2, P4, and P5 the estimated ℓIED regions are the same for both Pareto and neighbor-Pareto optimization algorithms and the results of neighbor-Pareto algorithm does not change for different values of ϵ in the range $[0, 0.6]$. For each of P1 and P3, one extra region which is included in the related vSOZ is obtained by neighbor-Pareto optimization compared to Pareto optimization. These extra regions are obtained for $0.3 \leqslant \epsilon \leqslant 0.6$ in neighbor-Pareto optimization or equivalently the Pareto and neighbor-Pareto optimization algorithms provide the same results for P1 and P3 for $0 \leqslant \epsilon < 0.3$.

Both of the algorithms provide consistent ℓIED regions with the related vSOZs, while since neighbor-Pareto algorithm provides the neighbor nodes of Pareto front in addition to Pareto front nodes, it provides more common number of regions with vSOZ. However Pareto optimization algorithm is parameter independent in comparison with neighbor-Pareto algorithm although the latter one is not very sensitive to its parameter ϵ. It is useful to apply both of these algorithms: Pareto optimization algorithm for the identification of strong leading IED regions (Pareto front) and neighbor-Pareto optimization algorithm for less strong ones (neighbor-Pareto optimal solutions excluding Pareto front). Such high-level information can be useful to identify the importance of each of visually detected SOZs.

We also compared estimated ℓIED regions based on multi-objective optimization and statistical multi-objective optimization methods. The parameters of the statistical method are given in Table 8.3. For the estimation of ℓIED regions based on the statistical method, the LI values are tested to exclude the nodes which their related LI values are significantly equal to zero before the application of optimization algorithm. The estimated ℓIED regions based on these methods provide mostly similar results for different patients. This result confirms the significance of LI values for the nodes of dDCG.

We prefer the multiple-objective optimization method in comparison with the statistical one since it has less computational load and time and moreover provides mostly similar ℓIED regions.

Table 10.8: Comparison between vSOZ and estimated ℓIED regions for five patients (P1 to P5) based on Pareto and neighbor-Pareto optimization algorithms in multiple graph analysis (section 7.3.2, page 70). For each patient, the column labels correspond to either ℓIED regions estimated by Pareto or neighbor-Pareto optimization algorithm, or the vSOZ of the patient. Each cross shows that the region related to its column is detected by the algorithm related to its row or by the vSOZ. amyg: amygdala; ant/post/m: anterior/posterior/mesial; CG: cingulate gyrus; entCx: entorhinal cortex; fusi: fusiform gyrus; HC: hippocampus; Ins: insula; midInsG: middle short gyrus of insula; pHcG: parahippocampal gyrus; T: temporal; TP: temporal pole.

P1	antHC	postHC	amyg	pHcG	mTP
neighbor-Pareto	×	×	×	×	
Pareto	×		×	×	
vSOZ	×	×	×	×	×
P2	antHC	postHC	amyg	pHcG	
neighbor-Pareto	×				
Pareto	×				
vSOZ	×	×	×	×	
P3	antHC	postHC	pHcG		
neighbor-Pareto	×	×	×		
Pareto	×	×			
vSOZ	×	×	×		
P4	antHC	postHC	amyg	entCx	mTP
neighbor-Pareto	×	×	×	×	
Pareto	×	×	×	×	
vSOZ	×	×	×	×	×
P5	midInsG				
neighbor-Pareto	×				
Pareto	×				
vSOZ	×				

10.4 Comparison of estimated ℓIED regions with other methods of SOZ detection

The resective surgical procedure is aimed at removing the SOZ without creating neurological deficits. The best way to detect SOZ is to record seizures from iEEG recordings. However, recording ictal EEG is time-consuming and requires an enormous effort when seizures are infrequent. Moreover, very long (several weeks) recording periods may increase the risk of complications like brain hemorrhage and infection [128–131]. Eventually, the number of seizures can be too limited, non-representative of usual seizures of a patient. Thus, obtaining statistically reliable results from studying seizure time intervals could be difficult. Consequently, other approaches based on interictal analysis for the prediction of SOZs by estimating the ℓIED regions remain warranted.

There are several studies [1, 8, 9, 12, 13, 22–24, 128, 133, 134, 206] for analysis of interictal periods to predict the SOZ from iEEG or ECoG. Comparison of different methods using different recordings is difficult due to the different resolution of recordings and different evaluation measures adapted to each method and related results. However, the results of previous studies and our findings are encouraging toward using the information of quantitative analysis of interictal periods for delinition of epileptogenic zone. Here, we compare the estimated ℓIED regions with two methods using the iEEG recordings of common patients (five and three common patients for first and second method, respectively).

1. SOZ detection by visual inspection (vSOZ)
2. SOZ detected by [1] from induced ictal periods by electrical stimuli (eSOZ)

In the following, we report the results of comparison of estimated ℓIED regions with vSOZ and eSOZ.

10.4.1 SOZ detection by visual inspection

Repeated electrical stimulation of the cortex during iEEG recording may induce transient changes of neural network properties which may be associated with increased likelihood of epileptic events in patients [207]. This clinical exam is performed for pre-surgical evaluations in the epilepsy department of Grenoble hospital (CHUG, Grenoble hospital, France) [208]. In this center, one of the methods of SOZ detection is based on visual inspection of pathological brain excitability (how a region is responsive to stimulations) during single pulse or repetitive stimuli [15]. In the classic method of vSOZ detection, SOZs are defined as the regions where the first electrophysiological changes detected at ictal onset (spontaneous ictal period). For the vSOZs reported in this work, the spontaneous ictal periods are used and induced ictal periods are used as a complementary tool.

The quantitative comparison of results between vSOZ, estimated ℓIED regions computed by our method, and removed regions during resective surgery are brought in Table 10.9 for the five patients (P1 to P5). The different quantitative measures are as follows. First, *dis* (mm) is the

average of minimum distances between ℓIED and vSOZ nodes, in which the proximity of the ℓIED nodes to vSOZ nodes is measured. The average of *dis* over five patients is 6.4 mm. This indicates the accuracy of agreement between vSOZ and estimated ℓIED regions. Second, *ovp* (%) is the average percentage of number of ℓIED nodes which are in the neighborhood ($\leqslant$ 1.5 cm) of at least one of the vSOZ nodes. The large percentage of *ovp* reveals that a large number of ℓIED nodes are in the neighbourhood of at least one of vSOZ nodes. Third, *ovp*2 (%) shows a similar percentage as *ovp* except *ovp*2 considers the vSOZ nodes which are in the neighborhood of at least one of the ℓIED nodes. We found an average of 84.4% and 84.8% over patients for *ovp* and *ovp*2, respectively.

10.4.2 Conclusion

All of the five patients are completely seizure-free after surgery and the estimated ℓIED regions are included in the resected regions for all of the patients. This result confirms that ℓIED regions can be valuable in pre-surgery evaluations while regions related to late secondary propagated activity (sink nodes) might not be necessary to be removed [12, 22, 23].

10.4.3 SOZ detected by [1] from induced ictal periods by electrical stimuli

David et al. [1] proposed two indices to measure the brain excitability and plasticity using 1 Hz stimulation-induced evoked responses of iEEG recordings from implanted patients to detect the SOZs. Brain excitability and plasticity (changes of intrinsic properties of neural networks) are approximated for all of electrode leads. The eSOZs were defined as the set of electrode leads which their both excitability and plasticity measure values are greater than their related thresholds. To optimize the selection of threshold related to each measure, three parameters are considered: 1) sensitivity: the probability over patients in which at least one electrode lead is chosen; 2) accuracy: the average over patients of selected electrode leads with the minimum distance to SOZ; 3) the average number of electrode leads used for SOZ detection over patients.

The comparison between the ℓIED regions estimated from interictal periods and eSOZ proposed by [1] from induced ictal periods of iEEG recordings is brought in Table 10.9.

- *P2*: We compare the activation map of eSOZ based on excitability and plasticity indices with the activation map related to ℓIED regions for P2. To obtain each activation map, a N-dimensional vector, $\mathbf{v}$, is assumed for N electrode leads. For electrode lead n which its index value (excitability or plasticity) is greater than the threshold, $v[n]$ is set to one, otherwise $v[n] = 0$. For the activation map of eSOZ based on the multiplication of excitability and plasticity indices, the related $v[n]$ is set similarly. We wonder if multiplication of these two measures can provide useful information since the regions which have both high excitability and plasticity are interested. For the activation map of estimated ℓIED regions, $v[n]$ is set to one for the estimated ℓIED electrode leads and zero for other electrode leads.

  The activation map is the interpolation of the values of $v[n]$ at the 3D electrode lead positions, overlaid on MR images.

The activation map of estimated ℓIED regions, eSOZs based on excitability, plasticity and the multiplication of excitability and plasticity are demonstrated for iEEG analysis of P2 in Figure 10.5. The color bar of each map indicates the different intensities used in the map. Regions in brighter intensities have higher contribution in the related method. All of the maps (parts (a)-(d)) are plot in the same coordinates (coordinates of electrode lead 72) for better comparison. Parts (a) to (c) of Figure 10.5 are associated with the activation map of eSOZs based on excitability, plasticity, and their multiplication, respectively. In part (d) the activation map of estimated ℓIED regions is demonstrated.

The eSOZs based on excitability and plasticity measures individually are {left anterior/posterior hippocampus, left parahippocampal gyrus} and {left anterior/posterior hippocampus, left amygdala}, respectively [1]. Thus the eSOZs upon both of these measures are {left anterior/posterior hippocampus}. The estimated ℓIED regions and vSOZ defined by epileptologist are {left anterior hippocampus} and {left anterior/posterior hippocampus, parahippocampal gyrus, amygdala}, respectively. {left anterior/posterior hippocampus, parahippocampal gyrus, amygdala, temporal pole} are resected during surgery and patient is seizure free after surgery. Since it is unknown how large the removed region was, using complementary information like estimated ℓIED regions or eSOZ might be helpful toward more conservative brain surgery resection.

- *P1-P5*: eSOZ and ℓIED regions are compared in Table 10.9 for three common patients (P1 to P3). There are more common regions between estimated ℓIED regions and vSOZ than between vSOZ and eSOZ for P1 and P3. For example for P3, eSOZ, vSOZ, and estimated ℓIED regions are {anterior hippocampus}, {anterior/posterior hippocampus, parahippocampal gyrus}, and {anterior/posterior hippocampus}, respectively. For two other patients, P4 and P5, this comparison was not possible. P4 recently underwent surgery and eSOZ results were not available. The SOZ of P5 cannot be defined by the eSOZ method since no seizure occurred by stimuli during iEEG recordings. Conversely, our method estimate the vSOZ correctly (middle short gyrus of insula) without using seizure periods, instead using IED and non-IED time intervals. This patient (P5) has the most focused vSOZ and resected region among all of the patients.

10.4.4 Conclusion

The proposed method estimates the ℓIED regions, congruent with vSOZ, requiring simpler, faster, and less expensive iEEG recording comparing to eSOZ method which requires induced seizure periods. For the patients in whom no seizure occurs during the stimuli, the estimated ℓIED regions can be valuable.

Table 10.9: Comparison between leading IED (ℓIED) regions estimated by our method, seizure onset zones detected by using induced seizure periods by electrical stimuli (eSOZ) following [1] and visually inspected SOZ (vSOZ) by epileptologist. *dis* (mm): the average of minimum distances between ℓIED and vSOZ nodes; *ovp*: the average percentage of number of ℓIED nodes which are in the neighborhood ($\leqslant$ 1.5 cm) of at least one of the vSOZ nodes. *ovp*2: the average percentage of number of vSOZ nodes that are in the neighborhood ($\leqslant$ 1.5 cm) of at least one of the ℓIED nodes; Bold regions demonstrate the uncommon regions between vSOZ and ℓIED regions. remR: removed regions during resective surgery; NA: not applicable; amyg: amygdala; ant/post/int/sup: anterior/posterior/internal/superior; entC: entorhinal cortex ; Hc: hippocampus; pHcG: parahippocampal gyrus; T: temporal; TP: temporal pole; m: mesial; G: gyrus; midInsG: middle short gyrus of insula.

	ℓIED	vSOZ	eSOZ	remR	*dis*	%*ovp*	%*ovp*2
P1	antHC pHcG amyg	ant/**postHC** pHcG amyg **mTP**	postHC	ant/postHC pHcG amyg TP antsupTG	9.57	60	60
P2	antHC	ant/**postHC pHcG amyg**	ant/postHC	ant/postHC pHcG amyg TP	0	100	85
P3	ant/postHC	ant/postHC **pHcG**	antHC	ant/postHC pHcG TP	4	100	100
P4	ant/postHC amyg entC	ant/postHC amyg **mTP** entC	-	ant/postHC amyg TP entC	12.91	62	79
P5	midInsG	midInsG	NA	midInsG	5.32	100	100
mean					6.4	84.4	84.8

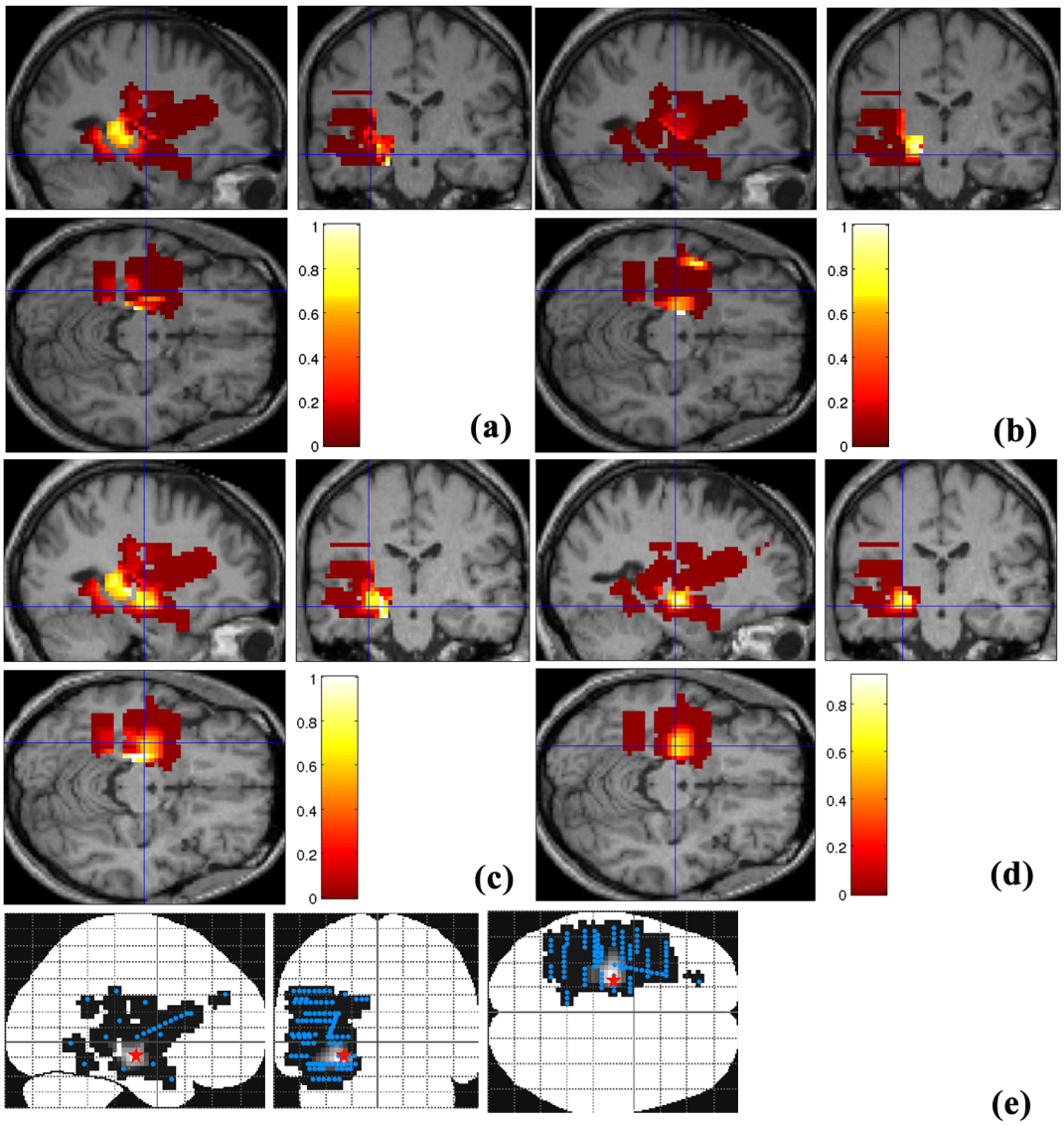


Figure 10.5: Activation map of eSOZs based on (a) excitability, (b) plasticity, (c) multiplication of excitability and plasticity and (d) activation map of estimated leading IED regions, overlaid on magnetic resonance (MR) image for P2. Part (e) shows the electrode lead locations (blue dots) overlaid on the map of estimated leading IED regions. The red star locates at the position of electrode lead number 72 that is located in left anterior hippocampus. The color bar of each map specifies the different intensities used in the map, zero is dark red, and one corresponding to white. For each activation map a N-dimensional vector, $v[n]$ is considered. $v[n]$ is set to one for the eSOZ electrode leads and zero for other electrode leads. The map presents the interpolation of $v[n]$ values at their 3D electrode lead positions, overlaid on MR images. The difference between maps is in the intensity change of corresponding interpolated $v[n]$ values. Regions showing in brighter colors, have greater contribution in the related method like left interior hippocampus in (d). All of the maps (part (a)-(d)) are plot in the same coordinate.

Chapter 11

Conclusions and Future Works

11.1 Conclusions

Drug resistant epileptic patients are recommended to do the surgery and the aim of this surgery is to remove SOZ without creating new neurological deficits. Recording seizures are difficult since they are not frequent in most of the patients. Therefore few number (1-2) of seizures are not sufficient for a statistical reliable analysis. A solution is to try to predict SOZ by estimating the IED regions from interictal analysis of iEEG recordings. There are several studies for this purpose [1, 8–10, 12, 13]. Since the problem is complicated, still prediction of SOZ by estimating IED regions is an open issue.

In this work we estimate the leading IED regions to help the epileptologists for the presurgery evaluations as a complementary information to EEG, fMRI, and semiology knowledge. To estimate the leading IED regions, we proposed a new reliable graph analysis method called Differential Connectivity Graph (DCG). DCG, the main original contribution of this work, is designed to identify the discriminated connections between two functional states. The connections whose statistical properties change significantly between two functional brain states construct DCG. As such, the effect of brain functions which are common between the two states are decreased.

DCG is implemented, tested and evaluated for differentiating between IED and non-IED states to extract the connections related to IEDs by decreasing the effect of common events like background activity, and its volume conduction effect. Therefore DCG provides sparser graph compared to separated IED and non-IED complex graphs providing simpler interpretation.

DCG is statistically reliable due to large number of IED and non-IED time intervals and using multiple testing based on permutation [145] . One of our contributions in this work is the application of permutation methods [145] for the graph inference. Permutation methods are powerful tools for identification of statistically significant tests (connections) assuming less assumptions about the underlying data contrary to other multiple testing methods which assume standard test statistic distributions. As such permutation method provides reliable identification of significantly discriminated connections of DCG. Moreover, contrary to threshold based graph inference methods, DCG is threshold independent.

Our two other contributions are as follows:

1. Introducing a new directed graph measure, LI for identification of source and sink nodes. LI measures the information which passes through each node locally (local information). This measure uses the outgoing and incoming connections of each node (digraph structure) and the information that is carried by each of these connections using lagged mutual information.

2. Proposing a multiple graph analysis to extract the information from the graph measure values of multiple graphs using multi-objective optimization method [149]. We found multi-objective optimization method very useful for extracting the information from measure values of the nodes of multiple graphs. In multi-objective optimization methods, different objectives whose preferences are not necessarily known can be optimized simultaneously providing a set of optimum solutions. As such the quantifications of multiple graphs can be summarized to the set of optimum nodes considering all of the graphs simultaneously.

   Furthermore, in the multi-objective optimization method applied in this work, we proposed a new optimization algorithm termed as neighbor-Pareto optimization to include the neighbor nodes of Pareto front in addition to the Pareto front.

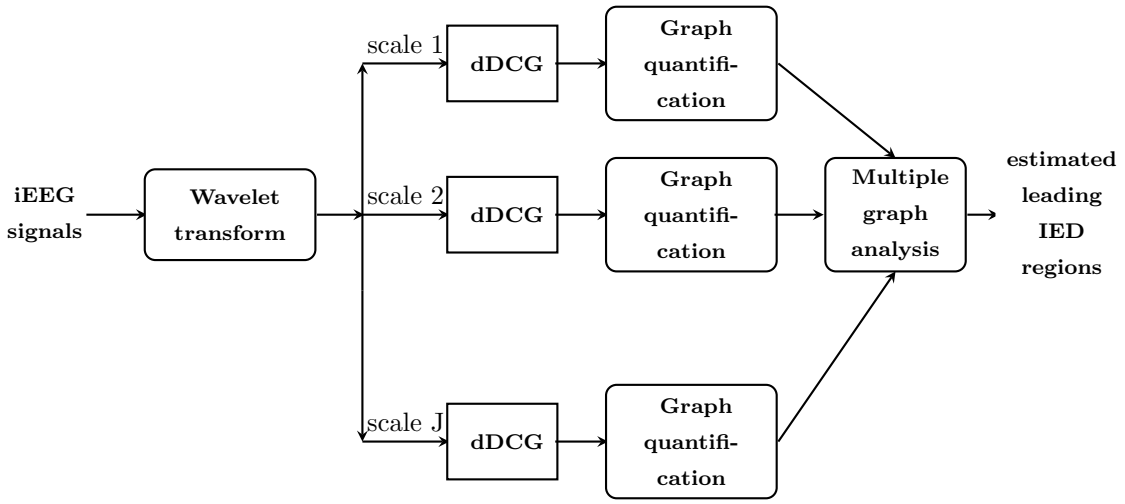


Figure 11.1: The block representation of the proposed method for iEEG analysis. The iEEG recordings are decomposed to different frequency bands (associated with different scales) using wavelet transform. Next for each scale the related directed DCG (dDCG) is obtained and the nodes of each graph is quantified. Finally the leading IED regions are estimated by multiple graph analysis using the graph measure values of all of the nodes of dDCGs.

Figure 11.1 summarizes the main blocks of the proposed iEEG analysis method. In this method, DCG is obtained for different frequency bands (associated with different scales). The direction of the edges of DCG are estimated by the estimation of the time causality between signal pairs related to each edge. Directed DCG (dDCG) nodes include source and sink nodes involved in the epileptic events. By characterization of dDCG nodes by local information and using multi-objective optimization method on these measure values in all of frequency bands, the leading IED nodes are estimated.

The efficiency of LI in comparison with other usual measures (global efficiency, local efficiency, outdegree, and indegree) is studied by comparing the leading IED regions based on different measures with visually inspected SOZs (vSOZ). In this perspective, LI is more informative and skilled compared to other classic measures for the identification of source and sink nodes since these usual measures are not well suited for this purpose. However comparing the measure values of different measures provides complementary characterizing information based on different definitions for more profound analysis of the role of each node. The advantage of LI over usual measures for source and sink identification is that it considers incoming connections as well as outgoing ones. Moreover LI considers the information which is carried by each of these connections.

The new neighbor-Pareto optimization algorithm in multi-objective optimization considers the regions in the neighborhood of Pareto front as well as Pareto front. The estimated ℓIED regions based on classic Pareto optimization and our new neighbor-Pareto optimization algorithm provide the same precision in the detection of the related vSOZs of each patient, while neighbor-Pareto optimization provides more common number of regions with vSOZs. However in the neighbor-Pareto optimization, the parameter ϵ should be chosen, while Pareto optimization algorithm is parameter independent. We prefer to apply both of these algorithms: Pareto optimization for identification of strong leading IED regions (Pareto front) and neighbor-Pareto optimization for identification of less strong ones (neighbor-Pareto optimal solutions excluding Pareto front). Such high-level information can be useful to identify the importance of each vSOZ.

The proposed method is applied on iEEG recording of five epileptic patients. We could provide statistically reliable leading IED regions using both IED and non-IED time intervals (without using seizure periods) congruent with visually inspected SOZ by the epileptologist and also with SOZ reported by [1] from induced ictal periods by electrical stimuli of iEEG recordings of the same patients. Moreover, estimated leading IED regions are all included in the removed regions (column remR of Table 10.9, section 10.4.1, page 110) during resective surgery. All of the patients being seizure free after resective surgery, one can conclude that estimated leading IED regions may present reliable information for the surgeon which should be useful for presurgical evaluations. Thus, the proposed method can provide efficient and repeatable analysis from iEEG recordings of epileptic patients.

11.2 Future works

Here is the list of the perspectives.

1. A simple perspective which leads to a fully automatic method is to replace the visual IED and non-IED labeling with automatic methods.

2. We are interested in considering large number of patients including the patients with primary failure surgery results and secondary seizure-free surgery output to test the performance of the method for difficult cases.

Although successful resective surgery or being the patients seizure-free is very valuable, currently there is no way to determine if the removed area is too large. Inclusion of estimated leading IED regions in the removed regions could be an argument toward more conservative surgery. To study about this perspective, we need to consider a clinical survey on large number of patients for a confident conclusion.

3. In the present method we are able to study the regions which the iEEG electrodes cover. Due to the difficulties of invasive electrodes, only limited suspected epileptic zones can be studied [14]. A very challenging perspective is to design an inverse model for localizing epileptic regions from non-invasive scalp EEG recordings.

4. In this work the idea of DCG is used for the estimation of leading IED regions, we think that DCG can be used for other applications, e.g. for identifying the source nodes related to a special task.

   More precisely, if we are interested in discriminating between one task (T1) and rest (instead of rest, a second task (T2) can be considered), the DCG can be extracted from couplings between signal pairs during T1 and rest. As such, we assume that DCG includes the discriminated connections between T1 and rest and its nodes include either sources or sinks related to T1. Using LI measure and multiple graph analysis, the source nodes related to T1 can be estimated. The procedure explained is the same procedure demonstrated in Figure 11.1, except the estimated leading IED regions should be replaced with the leading regions related to T1. It is important to mention that large number of time intervals related to T1 and rest are needed for statistically reliable results.

5. To estimate the directions of the edges of DCG, further investigations seem useful for comparing the current method of time causality estimation with other methods like transfer entropy.

6. In this work, we evaluated the efficiency of LI based on comparison of estimated leading IED regions related to underlying measure with visually inspected SOZ. We are interested in evaluating the efficiency of LI theoretically with other usual measures.

List of Related Publications

Journals

1. L. Amini, C. Jutten, S. Achard, O. David, P. Kahane, L. Vercueil, L. Minotti, G. A. Hossein-Zadeh, and H. Soltanian-Zade, Comparison Of Five Directed Graph Measures For Identification Of Leading Interictal Epileptic Regions, *Physiological Measurements*, *Physiol. Meas.*, vol. 31, pp. 1529-1546, 2010 **(chapter 7, page 65)**.

2. L. Amini, C. Jutten, S. Achard, O. David, H. Soltanian-Zadeh, G. A. Hossein-Zadeh, P. Kahane, L. Minotti, and L. Vercueil, Directed Differential Connectivity Graph Of Interictal Epileptiform Discharges, *under revision of IEEE Trans. Biomed. Eng.* **(part II, page 43)**.

Conferences

1. L. Amini, R. Sameni, C. Jutten, G. A. Hossein-Zadeh, and H. Soltanian-Zadeh, MR Artifact Reduction in the Simultaneous Acquisition of EEG and fMRI of Epileptic Patients, *Proc. of 16th European Signal Processing Conference (EUSIPCO), Lausanne, Switzerland*, August 25-29, 2008 **(appendix B, page 129)**.

2. L. Amini, S. Achard, C. Jutten, G.A. Hossein-Zadeh, and H. Soltanian-Zadeh, Connectivity Analysis of EEG Recordings for Epileptic Patients, *The 10th International Conference On Cognitive Neuroscience (ICON X), Bodrum, Turkey*, September 1-5, 2008 **(chapter 5, page 43)**.

3. L. Amini, S. Achard, C. Jutten, H. Soltanian-Zadeh, G. A. Hossein-Zadeh, O. David, and L. Vercueil, Sparse Differential Connectivity Graph of Scalp EEG for Epileptic Patients, *Proc. of the 17th European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium*, April 22-24, 2009 **(chapter 6, page 51 and appendix A, page 123)**.

4. L. Amini, C. Jutten, S. Achard, O. David, H. Soltanian-Zadeh, G. A. Hossein-Zadeh, P. Kahane, L. Minotti, and L. Vercueil, Directed Epileptic Network From Scalp And Intracranial EEG Of Epileptic Patients, *Proc. of the IEEE International Workshop On Machine Learning For Signal Processing (MLSP), Grenoble, France*, September 2-4, 2009 **(chapter 6, page 51)**.

Part IV

Appendices

Appendix A

Sparse differential connectivity graph of scalp EEG for epileptic patients

A.1 Abstract

The aim of the work is to integrate the information modulation of the inter-relations between EEG scalp measurements of two brain states in a connectivity graph. We present a sparse differential connectivity graph (SDCG) to distinguish the effectively modulated connections between epileptiform and non-epileptiform states of the brain from all the common connections created by noise, artifact, unwanted background activities and their related volume conduction effect. The proposed method is applied on real epileptic EEG data. Clustering the extracted features from SDCG may present valuable information about the epileptiform focus and their relations.

A.2 Introduction

Connectivity analysis using scalp EEG or fMRI data have been done based on different measures in the literature. The prominent among these measures are synchronization likelihood [209], correlation coefficients [20, 86], coherence [210], and Granger causality [211]. Furthermore, several types of evidences have been suggested in the literature proposing some measures to characterize topographical properties of the networks [5, 20, 81, 212]. There are also interests in simultaneous EEG and fMRI connectivity [212, 213].

In this paper, we present a sparse differential connectivity graph (SDCG) to study the relation between electrodes in two brain states based on the maximal overlap discrete wavelet transform (MODWT) [214], wavelet correlation estimation [20, 215] by means of connectivity measure, and multiple hypothesis t-test. In particular we address whether it is possible to benefit from connectivity graphs on scalp EEG for characterization of epileptiform sources.

The paper is organized as follows. In Section A.3, we describe the background, and the proposed approach. Section A.4 is devoted to the experimental results of the proposed method. Concluding remarks are presented in Section A.5.

A.3 Material and methods

A.3.1 Epilepsy

In epilepsy an area of the brain begins to discharge abnormally during a sudden and recurrent attack called seizure. Between two seizures, interictal epileptiform discharges (IED) may appear in the EEG measurements. The IEDs are waves or complexes (defined by International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN), 1974) discriminated from background activity. Since the appearance of IED has low probability, their quantitative analysis is rather challenging.

A.3.2 MODWT correlation estimation

In this purpose, wavelet correlation [215] is used as a measure of connectivity. The estimation of this measure is carried out utilizing the maximal overlap discrete wavelet transform (MODWT) [214], which is similar to discrete wavelet transform, but the signal is not subsampled and instead the filters are upsampled at each scale. Suppose $d_j^{s_1}[k]$ and $d_j^{s_2}[k]$ are the jth level MODWT coefficients of two stochastic processes with zero-mean stationary Gaussian backward differences ($s_1[k]$ and $s_2[k]$) [89]. The MODWT estimator of the correlation coefficient of $s_1[k]$ and $s_2[k]$ at scale j is:

$$\hat{\rho}_{s_1 s_2(j)} = \frac{\hat{Cov}\left\{d_j^{(s_1)}[k], d_j^{(s_2)}[k]\right\}}{\sqrt{\hat{Var}(d_j^{(s_1)}[k])\hat{Var}(d_j^{(s_2)}[k])}} \tag{A.1}$$

where $\hat{Cov}$ and $\hat{Var}$ are the estimations of covariance and variance respectively. This correlation coefficient estimation is asymptotically normally distributed with characterized confidence interval [89].

A.3.3 Proposed method

The flowchart of the proposed method is shown in Fig. A.1.

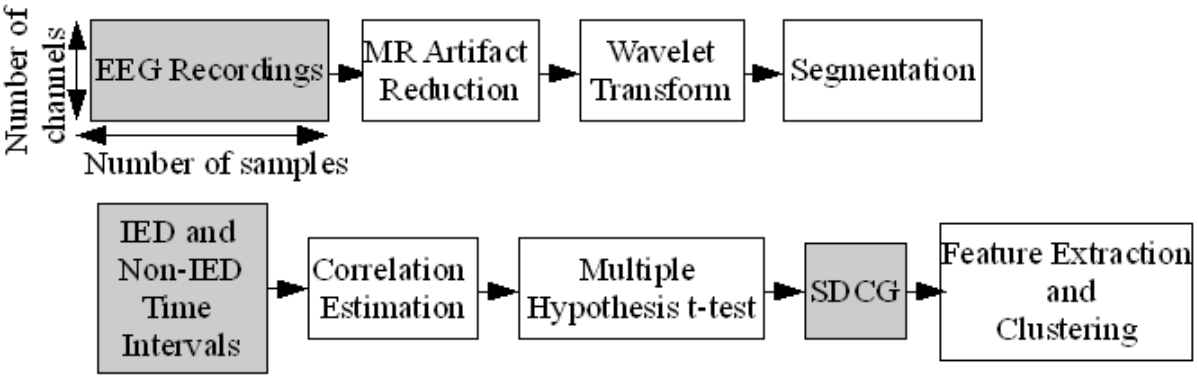


Figure A.1: Flowchart of the proposed method. The EEG preprocessing (top), graph computation and characterization (bottom).

1. **EEG data preprocessing**: Since our final project is simultaneous EEG and fMRI analysis, the EEG data was recorded inside the MR scanner. The first preprocessing step is to remove the MR artifacts using the method introduced in [216]. Next the expert neurologist labels the cleaned data by determining the start and end time points of IED occurrences as IED labels and time intervals without any IED as Non-IED labels. At the end, the 2-4 Hz MODWT coefficients of cleaned data are segmented using these labels.

   We have experimentally found 2-4 Hz wavelet coefficients as the best representing features of IED signals. The advantages of wavelet cross-correlation over Fourier cross-correlation has been cited in [215]. Moreover low frequency trends of the EEG data, which have unwanted effects on correlation estimation, can be easily removed by wavelet decomposition.

2. **Graph computation and characterization**:

   MODWT Correlation estimation: The MODWT cross-correlation based connectivity measure [89] is used to calculate the correlation coefficients between each pair channels of IED (Non-IED) segments denoted as $\mathbf{c}_l^{\text{IED}}[k]$ ($\mathbf{c}_l^{\text{Non-IED}}[k]$)$\in \mathbb{R}^{1\times N_C}$, $k = \{1, 2, \ldots, N_C\}$, $l = \{1, 2, \ldots, N_L\}$ where N_C, and N_L are respectively number of possible connections, and number of IED (Non-IED) labels. The matrix of row concatenation of $\mathbf{c}_l^{\text{IED}}[k]$ ($\mathbf{c}_l^{\text{Non-IED}}[k]$) vectors, i.e. wavelet correlation of all the connections and all the time IED (Non-IED) labels are denoted as $\mathbf{I}^{\text{IED}}[k]$ ($\mathbf{I}^{\text{Non-IED}}[k]$).

   Reference sensitivity reduction and multiple hypothesis t-test (MHT): By thresholding matrix of wavelet correlation, $\mathbf{I}^{\text{IED}}[k]$ ($\mathbf{I}^{\text{Non-IED}}[k]$) the common connectivity graph [20] for IED (Non-IED) state will be obtained (two separate graphs for each state). These common graphs have several problems in our application: (1) To obtain a sparse graph, thresholding is needed, and the graph depends on the thresholding; (2) The neighborhood nodes are connected due to the volume conduction effect; (3) Comparing two separate connectivity graphs of two brain states to determine the distinguished connections is rather challenging. These problems are solved by applying multiple hypothesis t-test between $\mathbf{I}^{\text{IED}}[k]$ and $\mathbf{I}^{\text{Non-IED}}[k]$. For each connection a t-test upon the following hypothesis is carried out:

$$\begin{cases} H_0^t: & \mu_1^t = \mu_2^t \\ H_1^t: & \mu_1^t \neq \mu_2^t \end{cases} \tag{A.2}$$

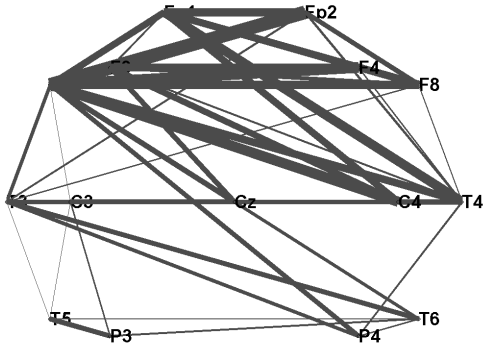
   where $t = 1, \ldots, N_C$, and μ_i is the mean of IED ($i = 1$) and Non-IED ($i = 2$) groups. The non-zero $t-values$ construct the SDCG. The positive and negative $t-values$ are separated for better analysis. The positive (negative) $t-values$ construct positive (negative) $t-value$ graph. A connection in positive (negative) $t - value$ graph shows the increase of wavelet correlations in IED (Non-IED) time intervals.
   A problem regarding the effect of the EEG reference (EEG data of a specific reference is

the subtraction of all the channels from that reference) is the sensitivity of the connectivity graph to the reference. If we calculate the $\mathbf{I}^{\text{IED}}[k]$ ($\mathbf{I}^{\text{Non-IED}}[k]$) matrix for two different references, the resulted graphs are not exactly the same, but not completely different. To solve this problem, the $\mathbf{I}^{\text{IED}}[k]$ ($\mathbf{I}^{\text{Non-IED}}[k]$) matrix is calculated for all possible references. Row concatenation of these matrices is denoted as $\mathbf{R}^{\text{IED}}[k]$ ($\mathbf{R}^{\text{Non-IED}}[k]$). Then the MHT is applied between $\mathbf{R}^{\text{IED}}[k]$ and $\mathbf{R}^{\text{Non-IED}}[k]$. The resulted graph gives the significant robust connections between the two brain states (IED and Non-IED) by considering the temporal and spatial information.

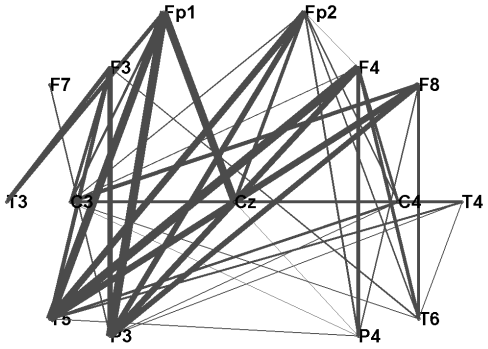
Feature extraction and clustering: The nodes of the resulted SDCG are quantified by global and local efficiency (GE and LE) [5,81]. High global efficiency of one node shows that the node is connected to many nodes of the graph. Local efficiency of one node is high when the neighbors of this node are highly connected. GE and LE are calculated for all the nodes of positive and negative $t-value$ graphs. The k-means method is utilized to cluster the nodes (EEG electrodes) of the SDCG based on five features including GE, LE of positive and negative $t-value$ graphs and power $t-values$. We can calculate the power of each electrode in IED and Non-IED time intervals in addition to correlations between the electrodes. Power $t-values$ are results of MHT between the powers of IED and Non-IED time intervals in the related frequency band. The source cluster (cluster including the electrodes close to the source) is labeled due to the physiological information about the patient.

A.4 Results and discussion

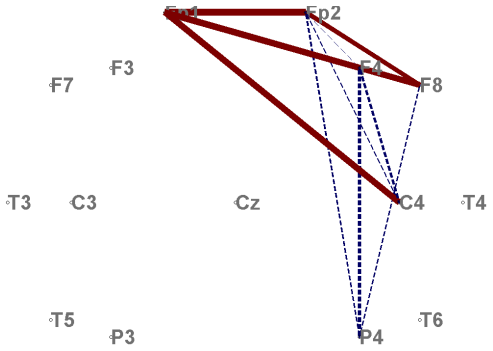
The proposed method is applied on the real and simulated [217, 218] EEG data of epileptic patients. Please refer to [216] for the protocol of the real data. In the SDCG of the simulated data, the electrodes near the source receive IED signals (volume conduction effect related to IED sources), hence the source electrodes (electrodes close to the source) have high LE in positive $t-value$ graph. Moreover, the GE of the source electrodes is high in negative $t-value$ graph, since the correlation between electrodes far from the source and source electrodes decreases during IED time interval. Similar properties is experimented in the real data. Therefore, GE, LE of positive and negative $t-value$ graphs and power $t-values$ (Section A.3) are selected to cluster the electrodes. The source cluster obtained by clustering (k-means) in this 5-D feature space for a real right frontal epilepsy data is shown in black in part (c) of Fig. A.2. The positive, negative $t-value$ graph connections, and the connections between the source cluster electrodes are depicted in parts (a)-(c) of this figure, respectively. Solid (dashed) lines show the positive (negative) $t-value$ connections in part (c). Each positive (negative) $t-value$ connection indicates the increase of wavelet correlations during IED (Non-IED) time intervals. The thickness of the connections is proportional to the absolute of $t-values$. The validity of the results is proved in the simulated data in which the electrodes close to the true source are detected. This result in real data is in accordance with the expert neurologist witness. Since



(a) Positive $t-value$ graph



(b) Negative $t-value$ graph



(c) Source electrode connections

Figure A.2: Sparse differential connectivity graph (SDCG) for right frontal epileptic patient. (a) Positive ((b) negative) $t-value$ graph connections show that the wavelet correlations of IED (Non-IED) time intervals are significantly higher than Non-IED (IED) time intervals. The thickness of the connections are proportional to the absolute of $t-values$. (c) The positive (negative) $t-value$ connections between source cluster electrodes are depicted in solid (dashed) lines. The source electrodes are shown in black.

SDCG indicates the connections whose wavelet correlations changing during IED and Non-IED time intervals, the noise, artifacts, background activity and their volume conduction effect are removed. But the volume conduction of the IED sources exists which promotes clustering the source electrodes as described above. Previous EEG connectivity studies suggested connectivity analysis on EEG activities (results of applying linear inverse problem). However any inverse problem method is based on some assumptions about the sources. The validity of the sources are dependent on these assumptions and the problem of volume conduction exists. To avoid these problems we applied connectivity analysis on the scalp EEG directly. However our aim is to study the brain functionality during epileptic and non-epileptic states of the brain to determine the related electrodes to the epileptiform sources from non-invasive EEG. This information is valuable for the intracranial electrode insertion. For precise seizure focus localization, we will apply the connectivity analysis on the intracranial EEG.

A.5 Conclusion

The proposed connectivity graph indicates the significant distinguished connections between two different brain states. By integrating complicated temporal information of EEG signal of the epileptic patient into a sparse differential connectivity graph and clustering the extracted features from the graph, we could determine the closer electrodes to the epileptiform sources.

Appendix B

MR artifact removal

B.1 Abstract

Integrating high spatial resolution of functional magnetic resonance imaging (fMRI) and high temporal resolution of electroencephalogram (EEG) is promising in simultaneous EEG and fMRI analysis, especially for epileptic patients. The EEG recorded inside a MR scanner is interfered with MR artifacts. In this article, we propose new artifact reduction approaches and compare them with the conventional image artifact reduction methods. Our proposed approaches are based on generalized eigenvalue decomposition (GEVD) and median filtering. The proposed methods are applied on experimental simultaneous EEG and fMRI recordings of an epileptic patient. The results show considerable improvement over conventional MR artifact reduction methods.

B.2 Introduction

Although simultaneous acquisition of EEG and fMRI may be essential for exploring the dynamics and localization of neural activity, several artifacts interfere the resulting EEG. These artifacts include: 1) MR artifacts, 2) RF pulse artifacts, and 3) Balistocardiogram (BCG) artifacts. The MR artifact is caused by the switching of the magnetic filed gradients used in the image acquisition, which induces unwanted voltages on the EEG recordings. The RF pulse artifact is due to the time varying electromagnetic field pulses (RF pulses) used for excitation in MRI recordings [219]. Although, the RF pulse has a very high frequency (in the order of several megahertz); but it is nonlinearly rectified to low frequencies (below 100Hz), within the EEG bandwidth. BCG artifact is caused by the micro motions of head EEG leads and wires within the static magnetic field. These motions are related to the pulsatile blood-flow in the head.

Different methods have been proposed for removing MR artifacts. Sparse component decomposition on the wavelet and discrete cosine basis [220], blind source extraction (BSE) followed by averaging-and-subtraction [221], and adaptive finite impulse response (FIR) filtering [222], are among these methods. Grouiller *et al.* [6] have compared some common methods such as image artifact reduction (IAR) [7], independent component analysis (ICA), fMRI artifact slice

template removal (FASTR), and filtering in the frequency domain using Fourier transform on simulated and real data. IAR [7] is a conventional method based on subtraction of the averaging artifact waveform followed by adaptive noise cancellation (ANC). FASTR is the combination of principal component analysis (PCA) and IAR. Among the compared methods of [6], IAR was shown to be the most effective method. This method is widely used for MR artifact removal in the literature.

In this article, we propose new approaches for MR artifact reduction. The proposed methods are based on generalized eigenvalue decomposition (GEVD) [223], [224], [225], [226], and median filtering. In fact, EEG signals having MR artifacts have both stationary and non-stationary properties at different time scales. MR artifacts include temporal structures that can be considered stationary. On the other hand, due to the switching of the magnetic field gradient, EEG contaminated with MR artifact is non-stationary in the time domain. In this work, two representations of MR artifact using this prior information have been utilized in an iterative GEVD scheme, together with a median and low-pass filter to extract the MR artifact. The results of the proposed method are compared with the IAR method over real simultaneous EEG and fMRI of epileptic patients.

The organization of the paper is as follows. Section B.3, explains the GEVD and its application in source separation. The proposed MR artifact reduction approaches are presented in Section B.4. Section B.5 is devoted to the experimental results by reporting the results and quantitative comparison of the conventional IAR method and the proposed methods. Concluding remarks are presented in Section B.6.

B.3 Review of Generalized Eigenvalue Decomposition

We assume zero-mean N-dimensional non-stationary observations $\mathbf{x}(t)$ that are mutually dependent in different N channels[1]. We are interested in linear mixtures of the form $y(t) = \mathbf{w}^T\mathbf{x}(t)$ that satisfy some measure of signal separability. For this, we define the following cost function:

$$J(\mathbf{w}) \doteq \frac{E_\theta\{y(\theta)y(\theta+\tau_\theta)\}}{E_t\{y^2(t)\}} \tag{B.1}$$

where $E_t\{\cdot\}$ represents averaging over t, and τ and τ_θ are respectively the time intervals and the time-varying lags that are found from our *a priori* knowledge of the signal's structure and non-stationarity. By maximizing this cost function, we are looking for the $y(t)$ with bounded energy and with a maximal lagged correlation over specific time intervals (θ) and time-lags (τ_θ).

Equation (B.1) can be rewritten as follows:

$$J(\mathbf{w}) = \frac{\mathbf{w}^T B_x \mathbf{w}}{\mathbf{w}^T C_x \mathbf{w}} \tag{B.2}$$

where

$$C_x \doteq E_t\{\mathbf{x}(t)\mathbf{x}(t)^T\} \tag{B.3}$$

[1]Such signals can for example be the output of a linear process of the form $\mathbf{x}(t) = A\mathbf{s}(t) + \mathbf{n}$ where $A \in \mathbb{R}^{N\times M}$

and

$$B_x \doteq E_\theta\{\mathbf{x}(\theta)\mathbf{x}(\theta+\tau_\theta)^T\} \tag{B.4}$$

The matrix C_x is the covariance matrix of $\mathbf{x}(t)$, which is known to be symmetric and positive definite. On the other hand, the matrix B_x is not generally symmetric and we need to make it symmetric ($B_x \leftarrow (B_x + B_x^T)/2$) for further use [2].

Equation (B.2) is in the form of the *Rayleight quotient* [226], and its maximum value is achieved through the joint diagonalization of the matrix pair (B_x, C_x):

$$\begin{cases} WB_xW^T = \Lambda \\ WC_xW^T = I \end{cases} \tag{B.5}$$

where Λ is a diagonal matrix containing real generalized eigenvalues on its diagonal (in descending order), and $W = [\mathbf{w}_1, ...\mathbf{w}_N]$ is the matrix containing the corresponding generalized eigenvectors as its columns. The input signals may next be decomposed as follows:

$$\mathbf{y}(t) = W^T\mathbf{x}(t) \tag{B.6}$$

where $\mathbf{y}(t) = [y_1(t), ..., y_N(t)]^T$ has decorrelated components, with the first component maximizing defined quotient in (B.2).

B.4 Proposed MR Artifact Reduction Approaches

In Fig.B.1, a typical segment of EEG contaminated with MR artifacts of two channels is depicted. MR artifacts are MR scanner dependent. MR artifact of the MR scanner used for our data are repeated with a period of three seconds, with two seconds of activation in each period (magnetic field gradient was on). A closer look at the MR artifacts of each period, shows sharp regular peaks with dominant high amplitude that are repeated every 62-63 samples (indicated with circles in Fig.B.1). Each peak is also surrounded by several smaller peaks.

In the previous Section, we noticed that the components extracted by GEVD are ranked according to an order that depends on the statistics carried by the matrix B_x. The spatial whiteness of the extracted components is guaranteed by the diagonalization of C_x. Here, our objective is to use the above mentioned properties of the MR artifact structure to form the matrix B_x, such that it contains the statistical properties of these artifacts. Therefore, since the $y_i(t)$ calculated from (B.6) are sorted in descending order of their corresponding eigenvalues, the MR artifacts are expected to be most concentrated in the first few components. In the following, two models of MR artifact (B_x) using this prior information are presented.

a. *First Statistical Measure*

In the first approach, we first detect the dominant MR peaks from an arbitrary channel. Due to

[2]Note that here we are not concerned by the non-positive definiteness of B_x; since it is verified experimentally that most of the eigenvalues has positive values and negative eigenvalues have significantly smaller absolute values than the positive ones.

the stationarity of these peaks, a robust method for detecting the peaks is a *matched filter* that uses a typical MR artifact segment as its template filter. A window is considered around each sharp peak of MR artifact. The resulted windowed segment is assumed to be stationary. Hence the average of the correlation matrices of windowed segments of two proceeding peaks can be a good model of statistical properties of MR artifact. In other words since the MR artifacts of different channels are rather regular, the samples that are equally distanced from each peak are highly correlated with one another. We can therefore, calculate the correlation between the windows surrounding each peak as a statistical property of the MR artifacts. Carrying out this procedure over the different channels, i.e., by considering each time sample of the different channels as an N-dimensional vector, we can redefine the matrix B_x defined in (B.4) as follows:

$$B_{x1} = E_\theta\{\mathbf{x}(\theta)\mathbf{x}(\theta + \Delta p_i)^T\} \tag{B.7}$$

where $\Delta p_i = p_{i+1} - p_i$ and $\theta \in \{\text{MR artifact time interval}\}$. p_i is the location of the i-th sharp peak of the MR artifact in the original data.

b. *Second Statistical Measure*

In the first model, the covariance matrix B_{x1} was calculated by direct vector-by-vector multiplication of the samples of short windows of $\mathbf{x}(\theta)$. According to Fig. B.1, each of these windows, consists of a high-amplitude MR artifact peak and a few smaller peaks around it. Therefore, in the statistical measure B_{x1}, each sample will have a contribution as strong as its energy. In other words, the smaller peaks around the dominant peaks of MR artifact will have less influence on the statistics of B_{x1} and the GEVD procedure will make less attempt to remove such peaks from the input signal. However, the strong and weak MR artifacts are somehow equally destructive in EEG analysis. We can therefore attempt in equalizing the peaks before the calculation of our statistical measure. This idea leads us to the following definition of the MR artifact covariance matrix:

$$B_{x2} = E_\theta\{[\mathbf{F}(\theta) \otimes \mathbf{x}(\theta)][\mathbf{F}(\theta) \otimes \mathbf{x}(\theta)]^T\} \tag{B.8}$$

where the operator $\otimes$ represents sample-by-sample multiplication, and $\theta \in \{\text{MR artifact time interval}\}$. $\mathbf{F}(\theta)$ is a column of the matrix including weighting function in each row. The weighting function equalizes the peak amplitudes of the windowed segments. B_{x2} is the average of weighted windowed segments energies.

B.4.1 Nonlinear Median Filtering

The GEVD procedure provides a means of transforming the input data, into components that are ranked according to their similarity with the MR artifacts. However, due to the linearity of the transformation (B.6), the performance of this primary stage is limited and it can not fully separate the MR artifacts from other components (including the EEG) except if the contaminated EEG signal perfectly satisfies a linear mixture ($\mathbf{x} = A\mathbf{s} + \mathbf{n}$), with a moderately low noise and a sufficient number of observation channels. Here the model of contaminated EEG signal

is rather complex as proved by our obtained results of applying linear ICA. Here we propose the GEVD followed by a nonlinear denoising procedure that is applied to the first D ($D < N$) components of the GEVD outputs while keeping the rest of $(N - D)$ components unchanged. This nonlinear denoising may indeed be applied to the original channels too (without the GEVD procedure) that is our third proposed approach (Fig. B.2.c); but in this case we would not benefit from the mutual information of the different channels. In fact, since the GEVD procedure concentrates the energy of the MR artifacts in the first few components, the EEG (non-MR artifact) components are less influenced by the nonlinear filtering when we apply this filter at the output of the GEVD procedure.

In this study, a two-step *moving window median* (MWM) filter was used for nonlinear filtering. Peaks such as the MR artifacts are less likely to pass the MWM filter and by subtracting the output of the second MWM filter from the input signal, we can assure that the MR artifact peaks are effectively removed, while the EEG components are least influenced. A theoretical study of the MWM filter and its extensions may be found in [227]. As a rule of thumb, the median filter window lengths should be wider than the width of the the MR artifacts. Therefore, the first median filter highly attenuates the sharp peaks, while the second one (with a wider window length) reduces the remaining peaks and smoothes the results. The MWM filter is further followed by a low-pass filter, to suppress the possible out of band components introduced by the nonlinear MWM filter and higher frequencies that are not interested in EEG analysis. After applying the filtering procedure denoted above (Median-LPF) to the first D components of GEVD, the D filtered and $(N - D)$ unchanged components are back-projected to the input space using the inverse of the decomposing matrix W^T (Fig. B.2.a). The preprocessed data, denoted by X in Fig. B.2, is the matrix including preprocessed EEG channels in each row. Preprocessed data is contaminated EEG with MR artifact without dc and linear trend. To detrend, the best straight-line fit is removed from each processed channel segment.

B.4.2 Iterative Denoising

Up to this point, we have improved the input signal quality by applying GEVD (using either B_{x1} or B_{x2}), the nonlinear filtering procedure, and the back-projection. As proposed in [228], this procedure may be repeated in several iterations, each time over the output of the previous run. This leads to a deflation procedure that removes one (or more) dimensions of the MR artifacts in each iteration and is repeated until the residual signals are 'sufficiently clean'. For this, we need some measure of signal cleanness that can be used as the stopping criterion of the iterative procedure. Due to the switching magnetic field gradient, we propose to use the ratio between the signal energy of the MR artifact time intervals (θ) over the rest of data, i.e. when the magnetic field gradient is off ($\tilde{\theta}$)):

$$C = E_{ch}\{\frac{E_{\theta}\{\hat{x}_{ch}^2(\theta)\}}{E_{\tilde{\theta}}\{\hat{x}_{ch}^2(\tilde{\theta})\}}\} \tag{B.9}$$

where $\hat{x}(t)$ is the cleaned data of each channel in each iteration and $E_{ch}\{\cdot\}$ is averaging over channels. The iterative procedure is stopped when the normalized difference of cleaning criteria between two iterations is less than a predefined threshold.

We hereby refer to this algorithm as Iterative GEVD. The overall flowchart of this algorithm is depicted in Fig. B.2.a. Following, the two statistical measures defined in (B.7) and (B.8), the iterative procedure can either be based on B_{x1}, which we call Iterative GEVD-Similarities (IGS), or on B_{x2}, which we call Iterative GEVD-Weighted (IGW). The flowcharts of IGS and IGW are depicted in Fig. B.2.b.

B.5 Experimental Results

B.5.1 Data

The recordings were made in the 3T scanner (3T Bruker BioSpin, Bruker Medizintechnik GmbH, Ettlingen, Germany) utilizing an MR compatible EEG amplifier (SD32, Micromed, Treviso, Italy) with 17 c-shaped electrodes positioned according to the 10/20 system (O1 and O2 were not used for subjects' comfort). The reference electrode was at Oz. A Gradient-Echo Echo Planar Imaging (GE-EPI) sequence used for MR image acquisition [6]. The sampling rate of EEG acquisition was 1024Hz. An anti-aliasing hardware low-pass filter with cut-off frequency of 286.8Hz was applied. EEG signals were calibrated with a square wave of 100μV utilizing an external calibrator plugged on all inputs. The epileptic patient was required to keep eyes closed and relax during the experiment.

B.5.2 Results

Initially, the data were preprocessed for DC and trend removal. The three proposed approaches were applied on the preprocessed data. For the first two approaches (IGS and IGW), as explained before, the sharp peaks of the MR artifacts were detected using a matched filter (Fig. B.1). Next, C_x, B_{x1} and B_{x2} were calculated according to the procedures explained in previous Sections. In the IGS method B_{x1} and C_x were jointly diagonalized, while in the IGW approach B_{x2} and C_x were jointly diagonalized. In each iteration of the proposed procedure, the first $D = 3$ of the $N = 15$ components extracted by GEVD, were denoised by the nonlinear MWM filter. The choice of D was based on empirical study of the output of the GEVD stage; the final results are not very sensitive to the value of $D > 3$. In the nonlinear filtering step, the first and second MWM window lengths used in the denoising were 20 and 40 samples (about 20ms and 40ms), respectively, since the typical width of sharp MR artifact peaks are about 20ms. The low-pass filter following the MWM filter was a first-order filter with a cut-off frequency of 45Hz, which is close to the effective bandwidth of the EEG. Next, the filtered and unchanged components were back-projected using the inverse of the decomposing matrix. This procedure was repeated until the normalized difference of the cleaning criteria defined in (B.9) between two consecutive iterations became less than 0.001. This limit was typically reached around 16 iterations. In Fig.

B.3, a typical segment of contaminated EEG and the results of IGS and IGW are shown. For comparison, the results of the Median-LPF and IAR [6, 7] are added in Fig. B.3.

B.5.3 Evaluation Methods

By visual inspection (Fig. B.3), the proposed method outperforms the IAR method presented in [6,7]. It is necessary to define an objective criterion for comparing the methods quantitatively. Since there is no ground truth, measuring the performance of each method is challenging. A first idea is to compare the signal with a reference which could be the signal acquired out of scanner or the time segment without MR artifact. The EEG signal inside and outside of MR scanner of a patient even with the same electrode locations are not necessarily similar. The EEG itself is very non-stationary in the time domain, especially for epileptic patients, since the IEDs (Interictal Epileptiform Discharges) appear spontaneously and the shape of IEDs are different. Due to this problem, the EEG contaminated with MR artifact and cleaned signal were compared. Comparable methods in the time domain were evaluated. The cleaned data were segmented around each peak of the MR artifact. The similarity between each two proceeding segments or all the segments was calculated. In other words, the periodicity of the cleaned data around each peak was measured. The cleaned data preserves the periodicity under each peak if the artifact removal method has not reduced MR artifact perfectly. In another method, the similarity between cleaned and original data under each sharp peak was measured. The cleaned data has some overshoots under each peak when the artifact removal method has not worked efficiently. However, since these methods of evaluation were not consistent for different MR artifact removal approaches, we compared spectral indexes of the EEG that are known to be more consistent. The following three criteria were defined for this purpose:

$$\begin{aligned} F_1 &= 10log_{10}E_{ch}\{\tfrac{E_i\{\hat{P}(fp_i)\}}{E_i\{P(fp_i)\}}\} \\ F_2 &= 10log_{10}E_{ch}\{E_i\{\tfrac{\hat{P}(fp_i)}{P(fp_i)}\}\} \\ F_3 &= 10log_{10}E_{ch}\{\tfrac{E_{i=17,34}\{\hat{P}(fp_i)\}}{E_{i=17,34}\{P(fp_i)\}}\} \end{aligned} \tag{B.10}$$

where $E_{ch}\{\cdot\}$ and $E_i\{\cdot\}$ are averages over channels and original data PSD peaks, respectively. P and $\hat{P}$ are the power spectral density (PSD) of the original and cleaned data, respectively. fp_i is an interval of 4Hz around the peak of the original data PSD in frequencies higher than 14Hz. This range is chosen since the interested spikes or IEDs appear below 14Hz. Since the sampling rate is not a multiple of the MR artifact frequencies, the energy of MR artifact frequencies can be distributed among a few frequency bins of the PSD. Therefore, an interval of 4Hz is considered around each PSD peak.

F_1 and F_2 calculate the PSD attenuation (in dB) of cleaned data around each peak of the original data PSD (first row of Fig. B.4). The method that most attenuates the PSD peaks of the MR artifact is preferred. Since sharp peaks are repeated in MR artifact every 62 or 63 samples and the sampling rate is 1024Hz, peaks around 16-17 Hz and their harmonics in the PSD of the original data is related to MR artifact. Therefore, F_3 measures the attenuation around

Table B.1: Frequency domain evaluation results (in dB) at low and high frequency resolutions. Orig-Out, IGS, IGW, Med-LPF, and IAR stand for original signal outside the MR scanner, Iterative GEVD-similarities, Iterative GEVD-weighted, Median-low-pass filter and Image Artifact Reduction [6], [7]

Frequency	Low Resolution			High Resolution		
Method(dB)	F_1	F_2	F_3	F_1	F_2	F_3
Orig-Out	-34	-33	-23	-35	-30	-25
IGS	-28	-27	-16	-28	-24	-17
IGW	-27	-25	-14	-26	-23	-15
Med-LPF	-26	-25	-13	-25	-22	-14
IAR	-21	-23	-11	-20	-21	-12

peaks of 16Hz and 32Hz that are shown in Fig. B.4. The peaks in higher harmonics are not considered in F_3, because frequencies higher than 45Hz are not of interest in EEG analysis. PSD of contaminated EEG, EEG signal out of MR scanner, and the processed signal by IGS, IGW, Median-LPF, and IAR are depicted respectively from up to bottom in Fig. B.4. To compute the PSD, the Discrete Fourier Transforms (DFT) of 0.5s windowed segments are calculated. To compare the results in high frequency resolution, the same criteria (F_1, F_2, and F_3) have been utilized, but the PSD is computed by the DFT of 9s windowed segments. The window length is chosen 9s since the gradient is switched on every 3 seconds. To measure how the cleaned data is following this period, several periods (e.g., 3 periods) are needed.

F_1, F_2 and F_3 in low and high frequency resolutions were calculated for IAR and the three proposed methods. It is important to mention that the low-pass filter applied on the results of the proposed methods utilized on the results of IAR method before F_1-F_3 calculation for comparing all the methods equally. The calculations of F_1-F_3 were carried out over 50 different 30s EEG segments for an epileptic patient. The average of 50 calculations of F_1-F_3 are reported in Table B.1. The method that better reduces the MR artifact has greater absolute values of F_1-F_3 that is a necessary condition but not sufficient. Since the outer original signal has no MR artifact (no PSD peaks after 14Hz), F_1-F_3 for the outer scanner signal has the maximum absolute values that is the minimum resemblance to the contaminated EEG with MR artifact. Hence the F_1-F_3 of the outer signal shows the minimum value in each column of Table B.1.

By comparing the results of Table B.1, it can be noticed that the proposed methods have outperformed the conventional IAR method. Among, the proposed methods first IGS, second IGW and third Median-LPF have better reduced the MR artifacts (using the defined criteria). In the IGS method, components of GEVD are well amplified in the first few components. But in the IGW the MR artifact is not extracted as good as the IGS method. The Median-LPF method is very simple, but it does not consider the information between channels. It also does not use the prior information as much as the first two methods. The PSD attenuation in MR artifact peaks of the IGS method is comparable with the EEG acquired outside of the scanner. This method has effectively mitigated MR artifacts from contaminated EEG.

B.6 Conclusion

In simultaneous EEG and fMRI recording, MR artifact reduction is an important issue and the development of an efficient artifact removal algorithm has significant effect on further processing. We proposed three artifact removal methods. The performance of our proposed artifact removal methods was evaluated on real epileptic patient data and appeared to outperform the conventional IAR method. The IGS method has shown better performance in MR artifact reduction compared to the other proposed methods. It is important to mention that as MR artifacts are MR scanner dependent, our method has been customized to the MR scanner used for our data. The presented algorithms were based on a linear model. Convolutive models will be studied in the future work for this purpose.

B.7 Acknowledgments

We gratefully acknowledge Olivier David and Frédéric Grouiller for acquiring the data and the assistance at the beginning of this study.

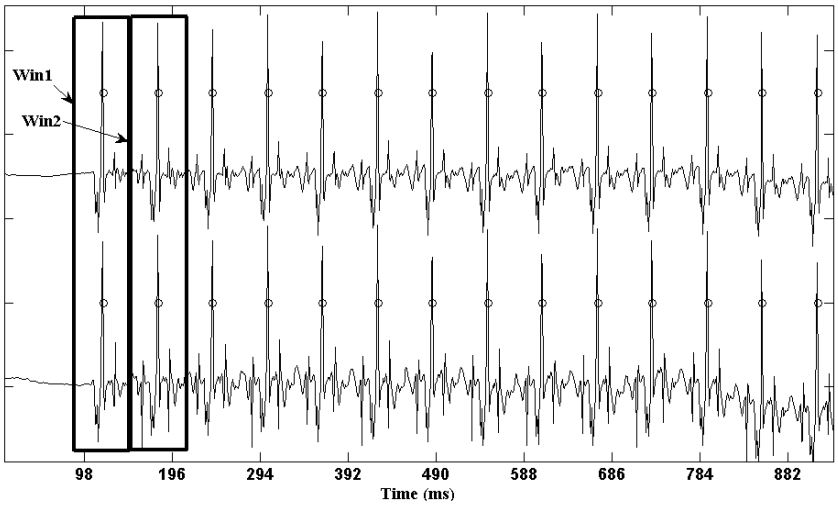


Figure B.1: Peak detection and windowing on MR Artifact of EEG signal. Peaks are shown with circles.

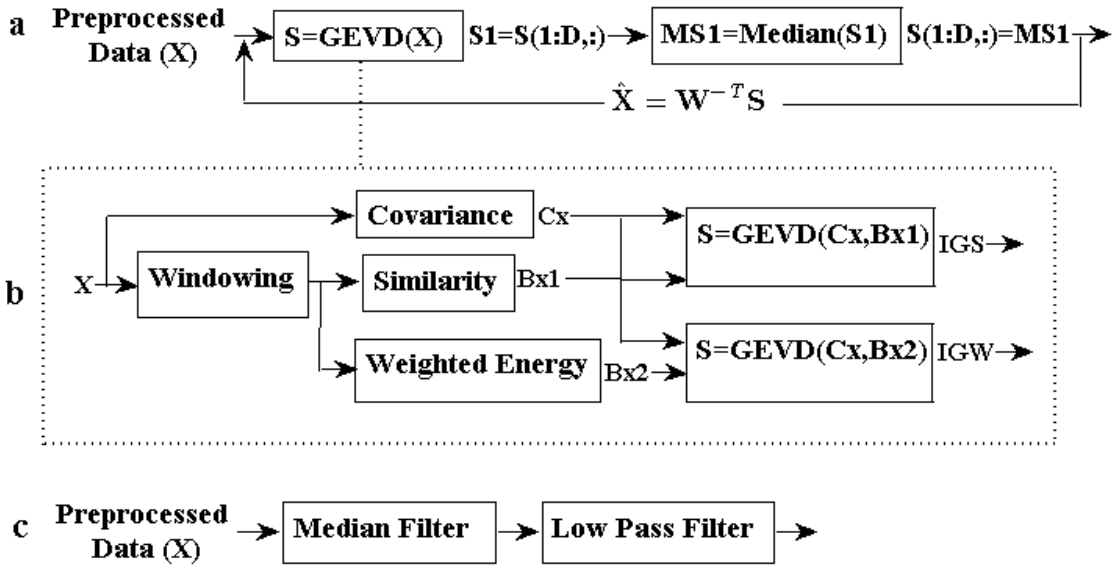


Figure B.2: Flowchart of proposed MR artifact reduction algorithms. (a) Iterative GEVD; (b) IGS and IGW, and (c) Median-low-pass filtering. D is the number of first few components in which the MR artifact is most amplified.

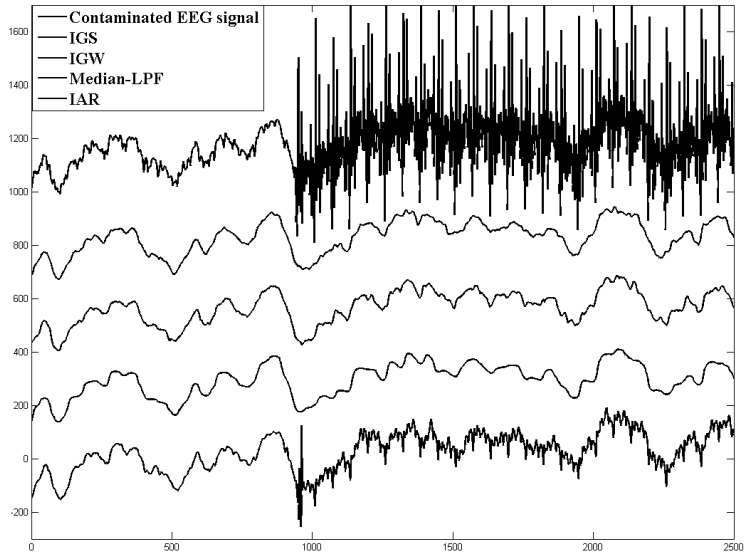


Figure B.3: Comparison of methods in reducing MR artifact from EEG signal. From up to bottom, contaminated EEG signal with MR artifact, the processed signal by IGS, IGW, Med-LPF, and IAR.

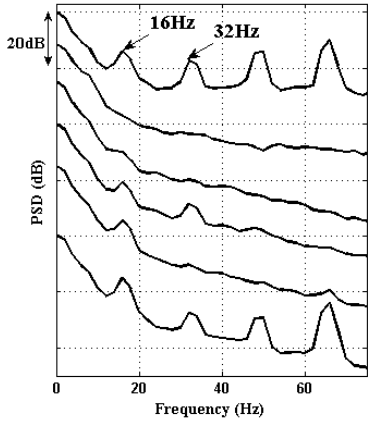


Figure B.4: Comparison of the PSDs. From up to bottom, PSD of contaminated EEG signal, outer scanner EEG signal, and the processed signal by IGS, IGW, Med-LPF, and IAR.

Appendix C

Types of graphs

There are different types of graphs which we explain some of them in this appendix.

- *Simple graph:* in a simple graph between two nodes no more than one edge is allowed. The edges are undirected and no loop (an edge with common ends) is permitted.
- *Multi-graph:* a graph in which more than one (directed or undirected) edge is allowed between nodes. Moreover loops are allowed in multi-graphs.
- *Undirected graph:* a graph in which edges have no orientation.
- *Directed graph (digraph):* the edges of a directed graph have directions. Oriented graph is a directed graph in which each edge is permitted to have a unique direction, i.e. between nodes i and j only one direction exist.
- *Cyclic and acyclic graph:* cyclic graph is a digraph including closed loops of edges. Conversely, acyclic graph does not contain the loops of edges.
- *Weighted graph:* in a weighted graph all of the edges do not have the same importance. To quantify this importance a weight (a number) is assigned to each edge.
- *Dynamic graphs:* these graphs may evolve over time which nodes and edges appearing or disappearing.

Types of graphs in terms of connectivity

- *Complete graph:* A complete graph is a fully connected graph which includes the maximum number of the possible edges.
- *Connected and disconnected graphs:* A path between nodes i and j is an ordered sequence of non-repeated edges and nodes connecting node i to node j. A graph is called connected if there exist at least one path connecting any node pairs, otherwise it is called disconnected. A k-node connected (k-connected) or k-edge connected graph is a graph in which the removal of some set of k nodes or of k edges makes it disconnected.

- *Strongly and weakly connected digraphs:* For the digraphs if by replacing directed edges with undirected ones the graph becomes connected then the digraph is called weakly connected. If there exist one path from i to j and from j to i for every node pairs, the graph is called strongly connected. Therefore an oriented graph cannot be strongly connected, whereas it can be weakly connected or disconnected.

Appendix D

Clustering coefficient and path length

In this appendix, first we define the general definitions for graphs including adjacency matrix, subgraph, and path. Next we explain clustering coefficient and path length for the graphs.

D.1 Adjacency matrix

Let us assume graph (undirected) $\mathbf{U} = (\mathbf{V_U}, \mathbf{E_U})$, where $\mathbf{V_U}$ and $\mathbf{E_U}$ are the set of vertices (or nodes) and the set of edges (or connections). The adjacency matrix, $\mathbf{A_U}$ is symmetric as if there is an edge between node pairs (i, j) then $a_{ij} = a_{ji} = 1$, otherwise $a_{ij} = a_{ji} = 0$. The complete graph includes $(N^2 - N)/2$ edges.

D.2 Subgraph

The subgraph of node i in graph $\mathbf{U}$ is defined as the graph of adjacent nodes of node i or neighbors of node i, i.e. $\{j \neq i \in \mathbf{V} |\ a_{ij} = a_{ji} = 1\}$ denoted as $\mathbf{U}_{i^+}$. When i is removed, the subgraph is denoted as $\mathbf{U}_{i^-}$.

D.3 Path

A path from node i to node j is an ordered sequence of non-repeated edges and nodes connecting node i to node j [17]. The shortest path length between nodes i and j is the minimum number of edges traversed to get from i to j.

We assume the shortest path length matrix denoted as $\mathbf{L_U} = [l_{ij}]$, where l_{ij} is the shortest path length between nodes i and j. $\mathbf{L_U}$ is symmetric as the path from i to j is the same path from j to i in graph $\mathbf{U}$. If any path does not exist from i to j, then l_{ij} is equal to ∞.

D.4 Clustering coefficient

Clustering coefficient is defined for connected graphs or strongly connected digraphs. The clustering coefficient of node i in graph $\mathbf{U}$ is defined as:

$$C[\mathbf{U}_{i^-}] = \frac{\text{order of } \mathbf{U}_{i^-}}{\text{maximum possible number of edges in } \mathbf{U}_{i^-}} = \frac{\text{order of } \mathbf{U}_{i^-}}{\frac{N_i \times (N_i - 1)}{2}} \tag{D.1}$$

where $\mathbf{U}_{i^-}$ is the subgraph of node i, i.e. the graph of the adjacent nodes of i if we remove i. N_i is the number of nodes of $\mathbf{U}_{i^-}$. If $\mathbf{U}_{i^-}$ is a complete graph then $C[\mathbf{U}_{i^-}]$ is maximum and equal to 1. The clustering coefficient of graph $\mathbf{U}$ is defined as:

$$C[\mathbf{U}] = \frac{1}{N} \sum_{i \in \mathbf{V_U}} C[\mathbf{U}_{i^-}]. \tag{D.2}$$

D.5 Path length

Similar to clustering coefficient, path length is also defined for connected graphs or strongly connected digraphs. The path length $P[\mathbf{U}]$ of graph $\mathbf{U}$ is defined as the average of the shortest path length l_{ij} between two arbitrary nodes i and j:

$$P[\mathbf{U}] = \frac{1}{N(N-1)} \sum_{i \in \mathbf{V_U}, i \neq j} l_{ij} \tag{D.3}$$

$P[\mathbf{U}]$ can be divergent for a disconnected graph and oriented graphs. Latora et al. [81] proposed local and global efficiency measures in terms of $1/l_{ij}$ in order to refine the problem of divergence. The concept of local and global efficiency is related to clustering coefficient and path length, respectively.

Appendix E

Graph classes

There are several graph classes which we explain some of them here. To define the graph classes, we start with the definition of degree distribution.

E.1 Degree distribution

Distribution function $p[k]$ [75, 77, 78] give the probability that a randomly selected node has exactly k edges. $p[k]$, $k = 0, 1, \ldots$ measures the proportions of nodes of the graph having degree equal to k. $p[k]$ is defined as:

$$p[k] = \frac{n_k}{N} \tag{E.1}$$

where n_k and N are the number of nodes of the graph having degree of k and the size of the graph, respectively. The n-moment of $p[k]$ [75, 77, 78] can be defined as:

$$\langle k \rangle^n = \sum_k k^n p[k] \tag{E.2}$$

The first moment $\langle k \rangle$ is the average degree of graph.

E.2 Regular graph

In a regular graph all of the nodes have the same degree denoted as k. In a k-regular graph, the degree of all of the nodes is equal to k. Similarly k_{out}-regular and k_{in}-regular digraphs can also be defined.

E.3 Random graph

A random graph is a graph in which edges are distributed randomly. A simple random graph has a bell-shaped Poisson degree distribution [74, 76] as:

$$p[k] = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \tag{E.3}$$

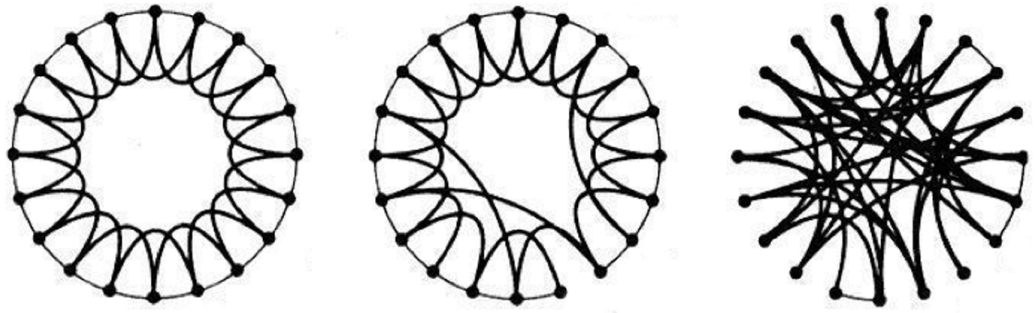


Figure E.1: From left to right, an example of a 4-regular, small world and random graph with the size of 20 are demonstrated. Rewiring with probability p is to change with probability p the end of an edge from its neighborhood to another node in the graph chosen uniformly at random with avoiding duplicate edges. By adding a few shortcuts to the regular graph (increasing the rewiring probability p), its path length will be decreased, while keeping its clustering coefficient high (small world network). By increasing the shortcuts or increasing the rewiring probability up to $p = 1$, we decrease both the path length and clustering coefficient, which leads to a random graph. This figure is taken from [5].

E.4 Small world network

In small world network there is a relatively short path between two nodes. A small world network is a network with "small" path length and "large" clustering coefficient [5]. The regular graph has high path length and high clustering coefficient. Because the neighbors of each node are well connected while reaching a node from another can be through a long path. On the other hand the random graph has low path length and low clustering coefficient. In random graphs there are lots of shortcuts which shortens the paths to reach a node from another whereas the neighbors of a node are not well connected. A small world network has low path length and besides it has high clustering coefficient. An example of regular, small world and random graph are demonstrated in Figure E.1 (from [5]).

E.5 Scale-free network

The degree distribution of the scale-free networks is no more Poisson, i.e. degrees are not clustered around one characteristic average degree (scale). These networks have power-law tail degree-distribution [17, 75, 229, 230] as:

$$p[k] \sim k^{-\gamma}, \gamma > 1. \tag{E.4}$$

where the degrees can span a wide range of values. Scale-free networks are a sub-class of small world networks.

Appendix F

Multiple testing

Here we explain general concepts in multiple testing including the statistical hypothesis, test statistic, significance level, p-value, and multiple test corection.

F.1 Statistical hypothesis and test statistic

Multiple testing methods are procedures to test $N_v > 1$ hypotheses simultaneously while controlling an error rate [190, 231].

We assume N_v dimensional observations $\mathbf{X} = [x_{mn}]$, $m = 1, \ldots, N_o$, $n = 1, \ldots, N_v$, where $\mathbf{x}_{m:}$, the mth row of matrix $\mathbf{X} \in \mathbb{R}^{N_o \times N_v}$, contains N_v samples of mth observation. N_o and N_v are the number of observations and variables, respectively. Let's assume $\mathbf{x}_{1:}$, $\ldots$, $\mathbf{x}_{N_o:}$ be independent and identically distributed (i.i.d.).

In a one-sample multiple testing, we test:

$$\begin{cases} H_0^n : & \mu_n = \mu_n^0 \\ H_1^n : & \mu_n \neq \mu_n^0 \end{cases} \tag{F.1}$$

where $n = 1 \ \ldots \ N_v$ and N_v is the number of tests. μ_n is the mean of nth column of matrix $\mathbf{X}$ or mean of $\mathbf{x}_{:n}$. μ_n^0 is the hypothesized null value. For example we can test if the coupling measure values for N_o time intervals of N_v connections are significantly non-equal to zero ($\mu_n^0 = 0$). H_0 and H_1 are called null and alternative hypotheses, respectively.

In two-sample multiple testing, we consider two populations $l = 1$ and $l = 2$. We assume N_v dimensional observations $\mathbf{X}^l = [x_{mn}^l]$, $m = 1, \ldots, N_o^l$, $n = 1, \ldots, N_v$, where N_o^l is the number of observations. $\mathbf{x}_{m:}^l$ is the mth row of matrix $\mathbf{X}^l \in \mathbb{R}^{N_o^l \times N_v}$.

For the two-sample multiple testing we are interested to test:

$$\begin{cases} H_0^n : & \mu_n^1 = \mu_n^2 \\ H_1^n : & \mu_n^1 \neq \mu_n^2 \end{cases} \tag{F.2}$$

where $n = 1 \ \ldots \ N_v$. μ_n^l is the mean of nth column of matrix $\mathbf{X}^l$ or mean of $\mathbf{x}_{:n}^l$. A test statistic (standardized or non standardized) is chosen depending on the probability model and the test hypothesis. The test statistic is a quantity calculated from the data to decide about the rejection

of H_0. For example a standardized test statistic can be defined as:

$$t_n = \frac{\widehat{\mu}_n^1 - \widehat{\mu}_n^2}{\sqrt{\frac{(\widehat{\sigma}_n^1)^2}{N_o^1} + \frac{(\widehat{\sigma}_n^2)^2}{N_o^2}}} \tag{F.3}$$

where $\widehat{\mu}_n^l$ and $\left(\widehat{\sigma}_n^l\right)^2$ are the estimated mean and variance of $\mathbf{x}_{:n}^l$.

F.2 Significance level and p-value

Based on the distribution of test statistic (using predefined statistical distributions or estimated by resampling methods) the p-values are calculated for each test.

The probability of falsely rejecting null hypothesis is Type I error rate or false alarm rate or significance level denoted as α.

The p-value is the probability value of getting a test statistic value equal or greater than one observed in the data assuming the null hypothesis is true. The p-value is compared with the α of the test and if it is smaller than α then the null hypothesis of the test is rejected. p-value provides how significant a test is. In other words p-value shows how likely the null hypothesis is false. The smaller the p-value the more convincing the rejection of the null hypothesis. For example for a test which its p-value is compared with $\alpha = 0.05$, if the p-value is smaller than 0.05 then H_0 is rejected at the 5% significance level.

F.3 Multiple test correction

F.3.1 Familywise error rate (FWER)

If we do one test, the probability of the event of "making a Type I error" is equal to α. If we do N_v tests, assuming the tests are independent, the probability of "making a Type I error" during the N_v tests will be α^{N_v}. Similarly the probability of the event "not making a Type I error" during N_v tests or family tests is $(1-\alpha)^{N_v}$. Therefore the probability of "making one or more Type I errors" which is the complement of "not making Type I error" is:

$$\alpha_{fw} = 1 - (1-\alpha)^{N_v} \tag{F.4}$$

where α_{fw} is $familywise$ α and α is the per test α. For example for $N_v = 10$ tests, if the probability of making a Type I error (α) is 0.05 then the probability of making at least one Type I error for the whole family of N_v tests is $\alpha_{fw} = 0.4$. This is called α inflation in multiple testing [232].

There are several methods for controlling the α inflation based on familywise error rate. Here we explain a few methods of Šidàk and Bonferonni [232, 233]. These methods introduce a correction on α values in order to keep α_{fw} fixed.

1. *Šidàk and Bonferonni corrections for independent tests:*

Equation (F.4) can be rewritten as:

$$\alpha = 1 - (1 - \alpha_{fw})^{1/N_v} \tag{F.5}$$

This equation is called Šidàk equation [232, 233] which shows that in order to obtain α_{fw} level, we need to adapt the α values used for each test. Since there is a fractional power in (F.5), Bonferonni tried to simplify this equation by approximating the Šidàk equation with the first term of its Taylor expansion which is:

$$\alpha \approx \frac{\alpha_{fw}}{N_v} \tag{F.6}$$

Šidàk and Bonferonni equations are linked as

$$\alpha = 1 - (1 - \alpha_{fw})^{1/N_v} \geqslant \frac{\alpha_{fw}}{N_v} \tag{F.7}$$

which shows that Bonferonni correction is more conservative and less powerful compared to Šidàk method. However (F.6) can be computed simpler than (F.5).

2. *Šidàk and Bonferonni corrections for non-independent tests:* For non-independent tests, the Šidàk correction gives a lower bound [232, 233]:

$$\alpha_{fw} \leqslant 1 - (1 - \alpha)^{N_v} \tag{F.8}$$

which can be rewritten as:

$$\alpha \geqslant 1 - (1 - \alpha_{fw})^{1/N_v} \tag{F.9}$$

Based on this inequality the Šidàk correction for non-independent tests can be approximated as:

$$\alpha \approx 1 - (1 - \alpha_{fw})^{1/N_v} \tag{F.10}$$

which is conservative according to the inequality in (F.9). Sidàk and Bonferonni equations are linked using previous Bonferonni approximation (F.6) as:

$$\alpha_{fw} \leqslant 1 - (1 - \alpha)^{N_v} < N_v\alpha \tag{F.11}$$

The Bonferonni correction for non-independent test is approximated as (F.6).

Adjusted p-value based on FWER corrections

The p-values calculated based on the test statistic distribution (as we described in section F.2) are termed as raw p-values and are denoted as $p[n]$, $n = 1, \ldots, N_v$. For a single test, the p-value of each test is compared with α, if p-value is smaller than α then H_0 is rejected. For multiple test correction if we replace α with (F.10) assuming non-independent tests, then we have:

$$p[n] \leqslant 1 - (1 - \alpha_{fw})^{1/N_v}\,. \tag{F.12}$$

If we rewrite this equation we obtain:

$$a[n] = 1 - (1 - p[n])^{N_v} \leqslant \alpha_{fw}, \tag{F.13}$$

where $a[n]$ is the adjusted p-value of each test based on Sidàk correction. For multiple testing if $a[n]$ is smaller than α_{fw} then the null hypothesis of the nth test is rejected. Now the Type I error rate for all of the tests are kept fixed equal to α_{fw}.

Similarly the adjusted p-values can be obtained for Bonferonni correction method as:

$$a[n] = N_v p[n]. \tag{F.14}$$

Appendix G

Cross-correlation coefficient of non-stationary signals

G.1 Introduction

In this appendix, firstly we discuss about the choice of maximum time lag in the MMCC estimation for simulated non-stationary signals. Secondly, we discuss about the choice of this parameter for MMCC and time causality estimation for real non-stationary signals. Thirdly we explain about the effect of byphasic signals on MMCC estimation. Finally, the concluding remarks are brought.

G.2 The choice of maximum time lag

Here we discuss about MODWT cross-correlation and time causality estimation between non-stationary simulated and real signal pairs.

G.2.1 Simulated data

Let's consider two non-stationary time series $\mathbf{x}(t)$ and $\mathbf{y}(t)$ as following:

$$\begin{cases} \mathbf{s}(t) = \sin(2\pi \frac{f}{f_s} t) \\ \mathbf{w}(t) = c_1 + \exp^{-\frac{(t-\tau_1)^2}{2\sigma^2}} \\ \mathbf{x}(t) = \mathbf{s}(t)\mathbf{w}(t) + c_2\mathbf{n}(t) \\ \mathbf{y}(t) = c_3\mathbf{x}(t-\tau^*) + c_4\mathbf{n}_1(t) + c_5 \end{cases} \tag{G.1}$$

where $\mathbf{w}(t)$ is a Gaussian window with mean and variance equal to τ_1 and σ^2, respectively. τ^* is the time causality between time series $\mathbf{x}$ and $\mathbf{y}$. $\mathbf{n}$ and $\mathbf{n}_1$ are standard Gaussian noise. $c_{\{1,\dots,5\}}$ are positive constants. f and f_s are the frequency and sampling rate, respectively. The time duration of $\mathbf{s}(t)\mathbf{w}(t)$ is proportional to σ^2. We assume $\mathbf{d}_x^j$ and $\mathbf{d}_y^j$ as the MODWT coefficients of $\mathbf{x}$ and $\mathbf{y}$ at frequency level j, respectively. The cross-correlation coefficient ($\hat{\rho}(\tau)$) is estimated (6.5) between $\mathbf{d}_x^j$ and $\mathbf{d}_y^j$ for different time shifts τ ($|\tau| \leqslant \tau^{\max}$). $\tau^{\max}$ is the maximum time

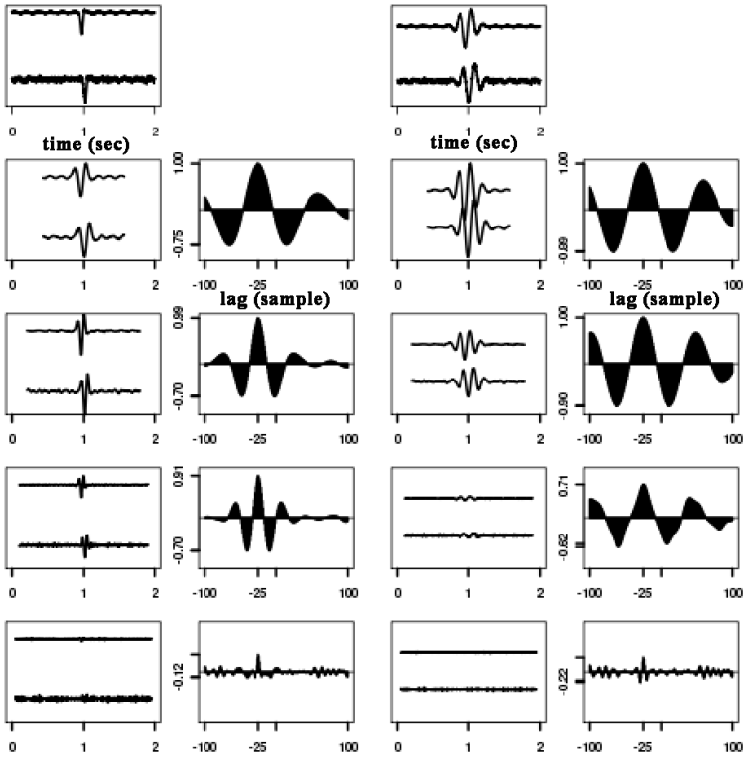


Figure G.1: Cross-correlation between MODWT coefficients of simulated signal pair for different frequency bands are depicted in second (σ=10) and forth ($\sigma = 50$) columns. From second to last row the frequency bands are 4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz, respectively. Temporal signal pairs are indicated in the first row for $\sigma = 10$ (first column) and 50 (third column). Their respective MODWT coefficients are demonstrated in odd columns.

shift. The values of simulated data parameters are reported in Table G.1.

Table G.1: Simulated data parameters.

Parameter	f(Hz)	f_s(Hz)	t(sec)	τ^*,τ_1 (sample)	$c_{1:5}$
value	6	512	0:2	25,500	0.025,0.02,2,0.05,0.3

Figure G.1, shows an example of signals in the form of $\mathbf{x}$ and $\mathbf{y}$. The first row demonstrates the temporal $\mathbf{x}$ and $\mathbf{y}$ and first and third columns correspond to σ equal to 10 and 50 samples, respectively. $\mathbf{d}_x^j$ and $\mathbf{d}_y^j$ of frequency levels j=6:3 (4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz) are shown in rows 2:5, respectively. Their respective cross-correlation functions (CCF) are depicted in the second and forth columns. The CCF of $\mathbf{d}_x^j$ and $\mathbf{d}_y^j$ with $\sigma = 10$ are depicted for different values of $\tau^{\max}$ in columns of Figure G.2. From first to forth columns, the CCF is depicted for $\tau^{\max}$ less than τ^*=25 samples (20 samples), greater than τ^* (27 and 100 samples) and adaptively selection of $\tau^{\max}$ (this method is explained later in this section), respectively. The lag at which the maximum of absolute CCF occurs is marked on the horizontal axis of images in Figures G.1 and G.2.

In order to provide proper MMCC estimation and true lag estimation performance, the correct choice of $\tau^{\max}$ is important. The true lag (τ^*) between non-stationary signals, $\mathbf{d}_{\mathbf{x}}^j$ and $\mathbf{d}_{\mathbf{y}}^j$, can

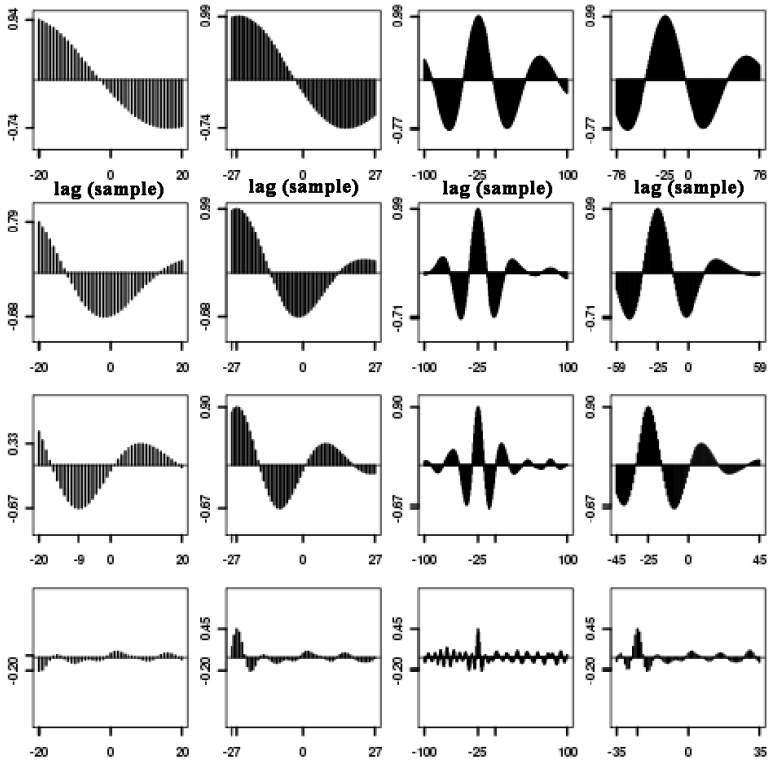


Figure G.2: Cross-correlation between MODWT coefficients of simulated signal pair (Figure G.1) are shown for $\tau^{\max} = 20$, 27, 100, and adaptive selection of $\tau^{\max}$ from first to last columns, respectively. From first to forth row the related frequency bands are 4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz, respectively.

be greater or smaller than $\frac{1}{f}$ (sec) as these signals are not periodic. When τ^* is greater than $\frac{f_s}{f}$ (sample), the true lag can be missed by confining $\tau^{\max}$ to less than $\frac{f_s}{f}$ (sample). In the first column of Figure G.2, it can be seen that the MMCC may not be properly estimated since $\tau^{\max}$ is not large enough. Therefore, $\tau^{\max}$ is suggested to be big enough to include the true lag. Contrarily, increasing the number of time shifts ($\tau^{\max}$), may increase variance and bias of cross-correlation for particular signals (Please see Appendices H and I).

In second ($\tau^{\max} = 27$ samples) and third ($\tau^{\max} = 100$ samples) columns of Figure G.2, the true lag is obtained correctly which $\tau^{\max}$ includes the true lag. However $\tau^{\max} = 100$ is much greater than it is needed ($\tau^* = 25$ samples), which may increase the variance and bias of MMCC estimation. To end with, the smallest maximum number of time lags examined in cross-correlation analysis including the true lag (τ^*) can be a proper selection. Satisfying this condition, requires the physiological knowledge about the time lags between underlying signals.

Here we explain an algorithm for an adaptive selection of τ^{max} to deal with a physiological range of time lags as following. Let's assume the physiological constraint be the range $[a, b]$. We initially start with $\tau^{\max}$ equal to lower limit of physiological range (a) and increase it iteratively up to its upper limit (b). At each iteration we test if $\tau^{\max}$ is sufficiently large to obtain the MMCC and the two nearest neighborhood extrema, otherwise $\tau^{\max}$ is increased. This procedure continues till $\tau^{\max}$ reaches to b. The results of this adaptively selection of $\tau^{\max}$ is shown in

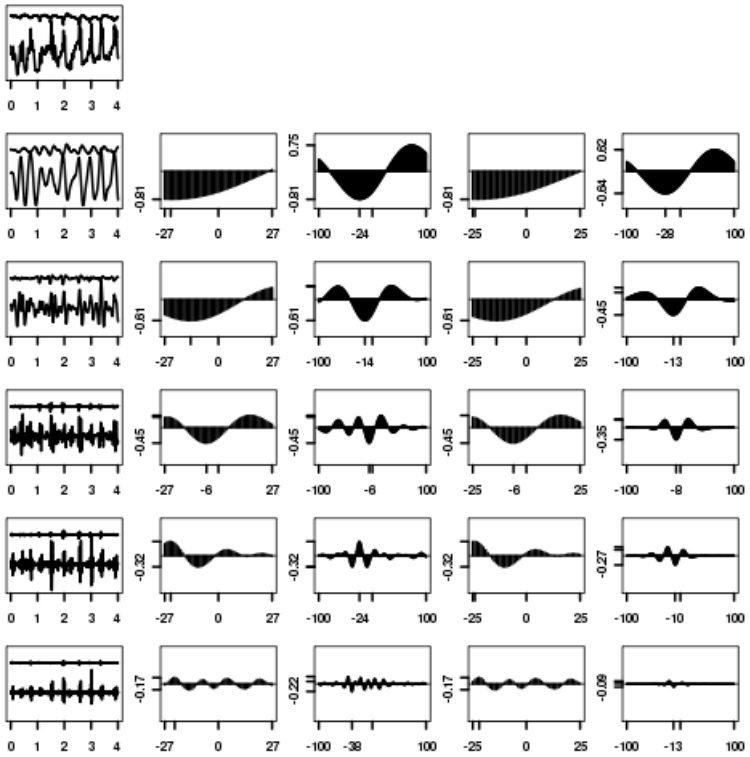


Figure G.3: Cross-correlation between MODWT coefficients of a typical real signal pair for different frequency bands. From second to last row the frequency bands are 2-4 Hz, 4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz, respectively. The temporal signal pair (first row) in a typical IED time interval and wavelet coefficients (first column) of different frequency bands are depicted. CCF for $\tau^{\max} = 27$, 100 samples, and adaptive selection of $\tau^{\max}$ for the demonstrated IED time window are shown in second to forth columns, respectively. Last column shows the CCF for the same signal pair during the whole selected data for processing.

the last column of Figure G.2 for $[a, b] = [25, 100]$ samples. It can be seen that using adaptive method, the true lag is estimated correctly with an adaptive $\tau^{\max}$ associated to each frequency level. However this method is computationally expensive and time consuming.

G.2.2 Real data

Here we demonstrate the effect of $\tau^{\max}$ on MMCC estimation for a typical real signal pair. This signal pair belongs to two nodes located in left posterior hippocampus (node 40) and left amygdala (node 13) for P1. In Figure G.3, the first row shows the temporal signal pair related to nodes 40 (upper) and 13 (lower) during a typical IED time window. The first column demonstrates the MODWT coefficients of this signal pair. Rows from 2 to 6 are related to the frequency ranges 2-4 Hz, 4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz, respectively. The second to forth columns show the CCF between the MODWT coefficients for $\tau^{\max}$ equal to 27 and 100 samples, and adaptively selection of τ^{max} respectively. The fifth column is the CCF for the whole length of selected data for processing. The physiological propagation delay (time causality) for our application is reported as $[25, 100]$ samples in [22]. For the demonstrated signal pair the time causality (section 6.3, page 61) is almost the same for the different selection of τ^{max} in the

second to forth columns and second to fifth rows. The cross-correlation coefficients between the signal pair is low in the 32-64 HZ frequency band (last row). This signal pair are connected in the DCGs related to 2-4 Hz and 8-16 Hz and in the related dDCGs, node 13 causes node 40 according to the estimated time causalities in these frequency bands (second and forth rows of last column).

Empirical variance of MMCC and time causality for different choices of τ^{max}

The empirical variance of MMCC estimation is approximated as the standard deviation of estimated MMCC in different IED time intervals. Increase of τ^{max} may increase the empirical variance of MMCC estimation. The box plot of MMCC for the same signal pair of Figure G.3 for different IED time intervals are shown in Figure G.4(b). The box plot of MMCC is depicted for $\tau^{\max} = 27$ and 100 samples in purple and green colors, respectively. Columns one to five correspond to frequency bands 2-4 Hz, 4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz, respectively. Considering the box plot of different signal pairs for different patients, the empirical variance of MMCC is increased for $\tau^{\max} = 100$ compared to 27 samples specially for lower frequencies.

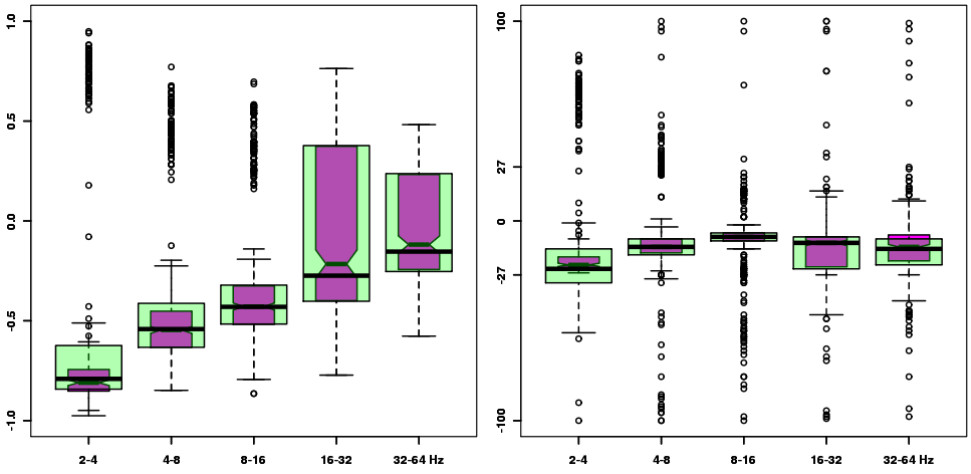
Figure G.4(a) shows the box plot of time causality. Purple and green box plots are associated with $\tau^{\max}$=27 and 100 samples, respectively. From first to last columns, the box plots are depicted for different frequency bands (2-4 Hz, 4-8 Hz, 8-16 Hz, 16-32 Hz, and 32-64 Hz), respectively. The median of time causalities for each frequency band is almost the same for $\tau^{\max} = 27$ ({-22,-13,-8,-10,-13}) and 100 ({-24,-13,-8,-11,-14}), while the empirical variance of time causality (standard deviation of estimated time causality in different IED time intervals) is increased for lower frequencies.

For $\tau^{\max} = 27$ samples, if the MMCC does not occur in $|\tau| < \tau^{\max}$ (the maximum of CCF occurs at a time lag greater than 27 samples), no number is considered for the MMCC, i.e. the true lag between underlying signals in the related frequency band is greater than $\tau^{\max}$.

G.3 MMCC estimation of byphasic signals

The MODWT cross-correlation function between non-stationary time series including byphasic signals may alternate in sign and tend to damp out with increasing time shift. The rate of this damping is different in different time windows and frequency ranges, depending on the underlying signals. In the time interval the two non-stationary signals stay in high correlation, the CCF alternates in sign. By decreasing the time duration of the byphasic term, the damping rate of CCF may increase. This issue is tested in the simulated data explained in section G.2.1, page 151. The time duration of byphasic term ($\mathbf{s}(t)\mathbf{w}(t)$) is proportional to σ^2. The decrease of damping rate of CCF by increasing σ is shown in second row and even columns of Figure G.1. σ is increased from second ($\sigma = 10$) to forth ($\sigma = 50$) column. For example, in second row and even columns, the value of MMCC and neighborhood extrema are { 1,0.75 } and { 1,0.99 }, respectively, which shows high damping rate of CCF for $\sigma = 10$ rather than $\sigma = 50$.

In real data, the same alternation can be seen in CCF. One of the values of the alternating



(a) ρ^{max} (b) τ^*

Figure G.4: The box plot of (a) MMCC and (b) time causality (6.6) for $\tau^{\max} = 27$ (purple) and 100 (green) samples. First to fifth columns corresponds to 2-4 Hz, 4-8 Hz, 8-16 Hz, 16-32 Hz and 32-64 Hz, respectively. Within box plots, the solid central line is the median of (a) ρ^{max} and (b) τ^*, and the edges of the box are the lower and upper quartiles (25th and 75th percentiles). The whiskers extend from each end of the box to the extreme values within 1.5 times the interquartile range. Outliers are the points out of this range that are shown by 'o' sign.

peaks is numerically greater but to know the significance of the greater peak, the confidence interval of each cross-correlation value should be studied. The related topic for the MMCC and time causality estimation are discussed in section 9.2, page 87 and section 6.3.3, page 63, respectively. Here we briefly summarize this issue in the next section.

G.4 Conclusion

The variance of MMCC estimation can be due to the non-stationarity of the couplings between signal pair during different time intervals and the estimation error. Different parameters can increase the estimation error like the maximum number of time lags, damping rate of CCF, number of independent samples of wavelet coefficients (appendix H, page 157), significance of CCF values, and correct choice of related frequency bands. Since the estimation of the distribution of our test-statistic in terms of MMCC is difficult due to non-stationarity of signals, in DCG calculation we used permutation-based multiple testing [190] (section 6.2.3, page 58) which assumes less assumptions about the data while it is a powerful tool for identification of significant discriminated connections. In time causality estimation, we considered long temporal signals for a better estimation of MMCC and we tested the reliability of time causality using jackknife methods (section 6.3.3, page 63).

Appendix H

MODWT cross-correlation between non-stationary fractionally differenced signals

Let's assume $\mathbf{x}$ and $\mathbf{y}$ be the realizations of two stochastic processes whose r_xth and r_yth order backward differences [89, 160, 167] are stationary Gaussian processes with zero mean. If $N_f > 2r$, $r = max(r_x, r_y)$ then $\mathbf{d}_x^j$ and $\mathbf{d}_y^j$ are the realizations of stationary processes with zero mean where $\mathbf{d}_x^j$ and $\mathbf{d}_y^j$ are the MODWT coefficients of $\mathbf{x}$ and $\mathbf{y}$, respectively. N_f is the length of base wavelet filter. Whitcher [89] has shown that if $N_f > 2r$ and $\left\{\mathbf{d}_x^j, \mathbf{d}_y^j\right\}$ is the realization of a bivariate Gaussian stationary process with square integrable autospectra, then the MODWT cross correlation estimation $\hat{\rho}(\mathbf{d}_x^j, \mathbf{d}_y^j, \tau)$ (6.5) is asymptotically normally distributed with characterized confidence interval. The approximate 100 (1-2p)% confidence interval of $\hat{\rho}(\mathbf{d}_x^j, \mathbf{d}_y^j, \tau)$ is calculated as [89, 167]:

$$\left[\tanh\left\{\tanh^{-1}(\hat{\rho}(\mathbf{d}_x^j, \mathbf{d}_y^j, \tau)) - \frac{\Phi^{-1}(1-p)}{\sqrt{T^j - 3}}\right\}, \tanh\left\{\tanh^{-1}(\hat{\rho}(\mathbf{d}_x^j, \mathbf{d}_y^j, \tau)) + \frac{\Phi^{-1}(1-p)}{\sqrt{T^j - 3}}\right\}\right] \quad \text{(H.1)}$$

where p is set to 0.025 for the confidence interval at 95% confidence level. Φ^{-1} is the quantile function of Gaussian distribution. T^j is the length of independant samples of wavelet coefficients at level j. $T^j = \frac{T}{2^j} - \left[(N_f - 2)(1 - 2^{-j})\right]$ [89]. T is the length of temporal signal. Equation (H.1) provides the confidence interval for $\tau = 0$ which can be generalized [89] for arbitrary lag (τ) by replacing T^j with $T^j - \tau$. Increase of τ^{max} or increase of $\frac{1}{\sqrt{(T^j - \tau^{max} - 3)}}$ for constant T^j, may increase the confidence interval of MODWT cross correlation. This increase is greater for smaller number of T^j.

Increase of $\frac{1}{\sqrt{(T^{j_0} - \tau^{max} - 3)}}$ for different choices of τ^{max} and for different length of T^{j_0} is shown in Figure H.1. T^{j_0} is T^j at a given frequency level j_0. Since increase of τ^{max} decreases the number of samples of underlying signals in cross-correlation function analysis, therefore τ^{max} is generally preferred to be a function of length of underlying signals (T^{j_0}), like $\tau^{max} = \frac{T^{j_0}}{4}$, $\sqrt{T^{j_0}}$, or $10log_{10}(\frac{T^{j_0}}{2})$ [85,183]. It can be seen that $\frac{1}{\sqrt{(T^{j_0} - \tau^{max} - 3)}}$ is less for $\tau^{max} = 10log_{10}(\frac{T^{j_0}}{2})$

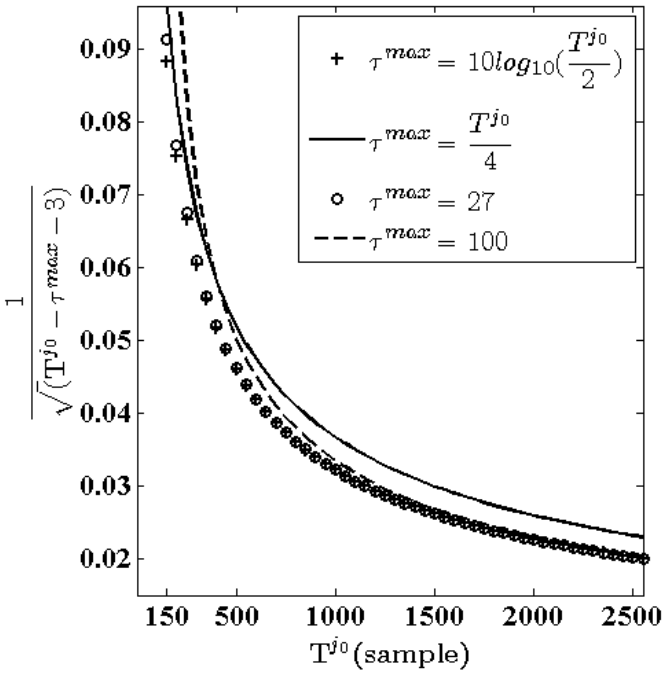


Figure H.1: Increase of confidence interval of cross correlation estimation by increment of maximum number of time lags. This confidence interval is proportional to $\frac{1}{\sqrt{(T^{j_0}-\tau^{max}-3)}}$ (H.1). T^{j_0} and τ^{max} are the number of independent samples at a given level j_0 and maximum time lag examined in cross correlation analysis. $\frac{1}{\sqrt{(T^{j_0}-\tau^{max}-3)}}$ is depicted for different number of T^{j_0} and different choices of τ^{max}.

rather than $\tau^{max} = \frac{T^{j_0}}{4}$ for shorter length signals. Two fixed choices ($\tau^{max} = 27$ and 100 sample) are compared with the two former length-dependent choices of τ^{max}. These choices correspond to some physiological knowledge.

Appendix I

Bias of maximum cross-correlation coefficient

Increasing τ^{max}, and eventually increasing the number of comparisons between variables for choosing the largest absolute cross-correlation between a pair of series out of $2\tau^{max}+1$ computed cross-correlation coefficients, may increase the bias of cross correlation estimation [203, 234].

Let $\{\rho(-\tau^{max}), \ldots, \rho(0), \ldots, \rho(\tau^{max})\}$ be the realizations of an independent, identically distributed (i.i.d.) zero-mean and unit variance Gaussian random variable $\rho(\tau)$. The bias of maximum cross-correlation estimation inflates by increasing the number of comparisons ($N_\tau = 2\tau^{max}+1$) as [235]:

$$\lim_{N_\tau \to \infty} P\{\max_{\tau \in [-\tau^{max}, \tau^{max}]} |\rho(\tau)| < \sqrt{2log(N_\tau)}\} \to 1 \tag{I.1}$$

where P is the probability.

Bibliography

[1] O. David, A. Wozniak, L. Minotti, and P. Kahane, "Preictal Short-Term Plasticity Induced By Intracerebral 1 Hz Stimulation," *NeuroImage*, vol. 39, no. 4, pp. 1633 – 1646, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/B6WNP-4R5VYTW-4/2/3fafd2facd2c6a2dd435d4204333acaa

[2] *Epilepsy Association of Scotland, http://www.epilepsiemuseum.de/alt/postergalerietr.html.*

[3] W. A. Hauser, J. F. Annegers, and W. A. Rocca, "Descriptive Epidemiology Of Epilepsy: Contributions Of Population-Based Studies From Rochester, Minnesota." *Mayo Clin Proc*, vol. 71, no. 6, pp. 576–586, Jun 1996.

[4] *M. R. Villarreal, LadyofHats, Gallery 2007-2009, biology, fr.wikipedia.org/wiki/Neurone.*

[5] D. J. Watts and S. H. Strogatz, "Collective Dynamics of Small-World Networks," *letters to nature*, vol. 393, pp. 440–442, 1998.

[6] F. Grouiller, L. Vercueil, A. Krainik, C. Segebarth, P. Kahane, and O. David, "A Comparative Study of Different Artefact Removal Algorithms for EEG Signals Acquired During Functional MRI," *Neuroimage*, vol. 38, pp. 124–137, 2007.

[7] P. Allen, O. Josephs, and R. Turner, "A Method for Removing Imaging Artifact from Continuous EEG Recorded During Functional MRI," *Neuro-Image*, vol. 12, pp. 230–239, 2000.

[8] M. Le Van Quyen, C. Adam, M. Baulac, J. Martinerie, and F. J. Varela, "Nonlinear Interdependencies Of EEG Signals In Human Intracranially Recorded Temporal Lobe Seizures," *Brain Res*, vol. 792, no. 1, pp. 24–40, May 1998.

[9] S. Monto, S. Vanhatalo, M. D. Holmes, and J. M. Palva, "Epileptogenic Neocortical Networks Are Revealed By Abnormal Temporal Dynamics In Seizure-Free Subdural EEG," *Cereb Cortex*, vol. 17, no. 6, pp. 1386–1393, Jun 2007.

[10] G. J. Ortega, R. G. Sola, and J. Pastor, "Complex Network Analysis Of Human ECoG Data," *Neurosci Lett*, vol. 447, no. 2-3, pp. 129–133, Dec 2008.

[11] M. A. Kramer, U. T. Eden, S. S. Cash, and E. D. Kolaczyk, "Network Inference With Confidence From Multivariate Time Series," *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 79, no. 6 Pt 1, p. 061916, Jun 2009.

[12] C. Wilke, W. van Drongelen, M. Kohrman, and B. He, "Identification Of Epileptogenic Foci From Causal Analysis Of ECoG Interictal Spike Activity," *Clin Neurophysiol*, vol. 120, no. 8, pp. 1449–1456, Aug 2009.

[13] F. Wendling, F. Bartolomei, and L. Senhadji, "Spatial Analysis of Intracerebral Electroencephalographic Signals in the Time and Frequency Domain: Identification of Epileptogenic Networks in Partial Epilepsy," *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.*, vol. 367, pp. 297–316, 2009.

[14] F. Rosenow and H. Lüders, "Presurgical Evaluation Of Epilepsy," *Brain*, vol. 124, no. 9, pp. 1683–1700, 2001. [Online]. Available: http://brain.oxfordjournals.org/cgi/content/abstract/124/9/1683

[15] P. Kahane, L. Minotti, D. Hoffmann, J. Lachaux, and P. Ryvlin, "Invasive EEG in the Definition of the Seizure Onset Zone: Depth Electrodes," *Handbook of Clin. Neurophysiol. Pre-surgical assessment of the epilepsies with clinical neurophysiology and functional neuroimaging (F. Rosenow and H. Lüders, eds.): Elsevier Science*, 2004.

[16] B. Horwitz, "The Elusive Concept Of Brain Connectivity," *NeuroImage*, vol. 19, no. 2, pp. 466 – 470, 2003.

[17] O. Sporns, D. R. Chialvo, M. Kaiser, and C. C. Hilgetag, "Organization, Development and Function of Complex Brain Networks," *Trends in Cognitive Science*, vol. 8, pp. 418–425, 2004.

[18] C. J. Stam, "Functional Connectivity Patterns Of Human Magnetoencephalographic Recordings: A 'Small-World' Network?" *Neuroscience Letters*, vol. 355, no. 1-2, pp. 25 – 28, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/B6T0G-4B2CRKJ-8/2/b7d003a81cc293afdeafb521e9c4599c

[19] M. A. Kramer, E. D. Kolaczyk, and H. E. Kirsch, "Emergent Network Topology At Seizure Onset In Humans," *Epilepsy Res*, vol. 79, no. 2-3, pp. 173–186, May 2008. [Online]. Available: http://dx.doi.org/10.1016/j.eplepsyres.2008.02.002

[20] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs," *The Journal of Neuroscience*, vol. 26, pp. 63–72, 2006.

[21] M. Valencia, J. Martinerie, S. Dupont, and M. Chavez, "Dynamic Small-World Behavior In Functional Brain Networks Unveiled By An Event-Related Networks Approach," *Phys. Rev. E*, vol. 77, no. 5, p. 050905, May 2008.

[22] G. Alarcon, J. J. G. Seoane, C. D. Binnie, M. C. M. Miguel, J. Juler, C. E. Polkey, R. D. Elwes, and J. M. O. Blasco, "Origin And Propagation Of Interictal Discharges In The Acute Electrocorticogram. Implications For Pathophysiology And Surgical Treatment Of Temporal Lobe Epilepsy," *Brain*, vol. 120 (Pt 12), pp. 2259–2282, Dec 1997.

[23] A. Hufnagel, M. Dumpelmann, J. Zentner, O. Schijns, and C. E. Elger, "Clinical Relevance Of Quantified Intracranial Interictal Spike Activity In Presurgical Evaluation Of Epilepsy," *Epilepsia*, vol. 41, no. 4, pp. 467–478, Apr 2000.

[24] G. J. Ortega, L. Menendez de la Prida, R. G. Sola, and J. Pastor, "Synchronization Clusters Of Interictal Activity In The Lateral Temporal Cortex Of Epileptic Patients: Intraoperative Electrocorticographic Analysis," *Epilepsia*, vol. 49, no. 2, pp. 269–280, Feb 2008.

[25] N. J. Mars, P. M. Thompson, and R. J. Wilkus, "Spread Of Epileptic Seizure Activity In Humans," *Epilepsia*, vol. 26, no. 1, pp. 85–94, 1985.

[26] M. Saab and J. Gotman, "A System To Detect The Onset Of Epileptic Seizures In Scalp EEG," *Clinical Neurophysiology*, vol. 116, no. 2, pp. 427 – 442, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/B6VNP-4DBT0WY-2/2/a9e26ed2e28c7ddcaeefa00167966249

[27] A. Ozkaya and M. Korürek, "Estimating Short-Run And Long-Run Interaction Mechanisms In Interictal State," *Journal of Computational Neuroscience*, vol. 28, pp. 177–192, 2010, 10.1007/s10827-009-0198-7. [Online]. Available: http://dx.doi.org/10.1007/s10827-009-0198-7

[28] G. E. Chatrian, L. Bergamini, M. Dondey, D. W. Klass, M. Lennox-Buchthal, and I. Petersén, "A Glossary Of Terms Most Commonly Used By Clinical Electroencephalographers," *Electroencephalography and Clinical Neurophysiology*, vol. 37, no. 5, pp. 538 – 548, 1974. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-482XHF9-167/2/bf55c776fc11a44dc87e7c9bf703d747

[29] IFSECN, "A glossary of terms most commonly used by clinical electroencephalographers," *ElectroencephalogrClinNeurophysiol*, vol. 37, pp. 538–53, 1974.

[30] H. Lüders and W. Bingaman, *Textbook Of Epilepsy Surgery*. Informa UK Ltd, 2008.

[31] J. Gonzàlez-Martìnez, T. Srikijvilaikul, D. Nair, and W. E. Bingaman, "Long-Term Seizure Outcome In Reoperation After Failure Of Epilepsy Surgery," *J Neurosurgery*, vol. 60, pp. 873–880, 2007.

[32] J. Gotman, "Automatic Detection Of Seizures And Spikes," *Journal of Clinical Neurophysiology. Inpatient/Outpatient EEG Monitoring*, vol. 16, pp. 130–140, 1999.

[33] F. Bartolomei, P. Chauvel, and F. Wendling, "Epileptogenicity Of Brain Structures In Human Temporal Lobe Epilepsy: A Quantified Study From Intracerebral EEG," *Brain*, vol. 131, no. Pt 7, pp. 1818–1830, Jul 2008. [Online]. Available: http://dx.doi.org/10.1093/brain/awn111

[34] A. K. Roopun, R. D. Traub, T. Baldeweg, M. O. Cunningham, R. G. Whittaker, A. Trevelyan, R. Duncan, A. J. C. Russell, and M. A. Whittington, "Detecting Seizure Origin Using Basic, Multiscale Population Dynamic Measures: Preliminary Findings," *Epilepsy Behav*, vol. 14 Suppl 1, pp. 39–46, Jan 2009. [Online]. Available: http://dx.doi.org/10.1016/j.yebeh.2008.09.008

[35] K. A. Schindler, S. Bialonski, M. T. Horstmann, C. E. Elger, and K. Lehnertz, "Evolving Functional Network Properties And Synchronizability During Human Epileptic Seizures," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 18, no. 3, pp. 0 331 191–6, 2008. [Online]. Available: http://link.aip.org/link/?CHA/18/033119/1

[36] L. Alonso-Nanclares, J. Gonzalez-Soriano, J. R. Rodriguez, and J. DeFelipe, "Gender Differences In Human Cortical Synaptic Density," *Proc Natl Acad Sci U S A*, vol. 105, no. 38, pp. 14 615–14 619, Sep 2008. [Online]. Available: http://dx.doi.org/10.1073/pnas.0803652105

[37] A. Kaplan and S. L. Shishkin, "Application Of The Change-Point Analysis To The Investigation Of The Brain's Electrical Activity," *Chapter 7 in: B.E.Brodsky, B.S.Darkhovsky. Nonparametric Statistcal Diagnosis: Problems and Methods. Kluwer Academic Publishers, Dordrecht (the Netherlands)*, pp. 333–388, 2000, http://brain.bio.msu.ru/papers/chp2000/.

[38] H. Berger, "Über Das Elektroenkephalogramm Des Menschen," *Archiv für Psychiatrie und Nervenkrankheiten*, vol. 87, pp. 527–570, 1929.

[39] B. N. Cuffin, "Effects Of Local Variations In Skull And Scalp Thickness On EEG's And MEG's," *TBME*, vol. 40, no. 1, pp. 42 –48, jan. 1993.

[40] O. Foerster and H. Altenburger, "Elektrobiologische Vorgänge An Der Menschlichen." *Dtsch Z Nervenheilkd*, vol. 135, pp. 277–88, 1935.

[41] C. Ajmone-Marsan and J. Van Buhen, "Epileptiform Activity In Cortical And Subcortical Structures In The Temporal Lobe Of Man." *M. Baldwin and P. Bailey, Ed. Temporal lobe epilepsy. Springfield, (IL): C. T. Charles*, pp. 78–108, 1958.

[42] D. R. Nair, R. Burgess, C. C. McIntyre, and H. Lüders, "chronic Subdural Electrodes In The Management Of Epilepsy," *Clin Neurophysiol*, vol. 119, no. 1, pp. 11–28, Jan 2008.

[43] J. Talairach and J. Bancaud, "Stereotaxic Approach To Epilepsy. Methodology Of Anatomo-Functional Stereotaxic Investigations," *Progr Neurol Surg*, vol. 5, pp. 297–354, 1973.

[44] J. Bancaud and J. Talairach, "Methodology Of Stereo EEG Exploration And Surgical Intervention In Epilepsy," *Rev Otoneuroophtalmol*, vol. 45, no. 4, pp. 315–328, 1973.

[45] J. Talairach, J. Bancaud, and G. Szikla, "Approche Nouvelle De La Neurochirugie De L'Epilepsie. Méthodologie Stérérotaxique Et Résultats Thérapeutiques," *Neurochirurgie*, vol. 20, pp. 1–240, 1974.

[46] J. Talairach, J. Bancaud, G. Szikla, A. Bonis, S. Geier, and C. Vedrenne, "New Approach To The Neurosurgery Of Epilepsy. Stereotaxic Methodology And Therapeutic Results. 1. Introduction And History," *Neurochirurgie*, vol. 20 Suppl 1, pp. 1–240, Jun 1974.

[47] F. Bartolomei, M. Guye, M. Gavaret, J. Régis, F. Wendling, C. Raybaud, and P. Chauvel, "The Presurgical Evaluation Of Epilepsies," *Rev Neurol (Paris)*, vol. 158, no. 5 Pt 2, pp. 4S55–4S64, May 2002.

[48] M. Cossu, F. Cardinale, L. Castana, A. Citterio, S. Francione, L. Tassi, A. L. Benabid, and G. L. Russo, "Stereoelectroencephalography In The Presurgical Evaluation Of Focal Epilepsy: A Retrospective Analysis Of 215 Procedures," *Neurosurgery*, vol. 57, no. 4, pp. 706–18, Oct 2005.

[49] M. Cossu, F. Cardinale, N. Colombo, R. Mai, L. Nobili, I. Sartori, and G. L. Russo, "Stereoelectroencephalography In The Presurgical Evaluation Of Children With Drug-Resistant Focal Epilepsy," *J Neurosurg*, vol. 103, no. 4 Suppl, pp. 333–343, Oct 2005. [Online]. Available: http://dx.doi.org/10.3171/ped.2005.103.4.0333

[50] D. Cosandier-Rimélé, J.-M. Badier, P. Chauvel, and F. Wendling, "A Physiologically Plausible Spatio-Temporal Model For EEG Signals Recorded With Intracerebral Electrodes In Human Partial Epilepsy," *IEEE Trans Biomed Eng*, vol. 54, no. 3, pp. 380–388, Mar 2007. [Online]. Available: http://dx.doi.org/10.1109/TBME.2006.890489

[51] A. R. McIntosh, C. L. Grady, L. G. Ungerleider, J. V. Haxby, S. I. Rapoport, and B. Horwitz, "Network Analysis Of Cortical Visual Pathways Mapped With Pet," *J Neurosci*, vol. 14, no. 2, pp. 655–666, Feb 1994.

[52] K. J. Friston, L. Harrison, and W. Penny, "Dynamic Causal Modelling," *Neuroimage*, vol. 19, no. 4, pp. 1273–1302, Aug 2003.

[53] W. D. Penny, K. E. Stephan, A. Mechelli, and K. J. Friston, "Modelling Functional Integration: A Comparison Of Structural Equation And Dynamic Causal Models," *Neuroimage*, vol. 23 Suppl 1, pp. S264–S274, 2004. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2004.07.041

[54] L. Lee, K. Friston, and B. Horwitz, "Large-Scale Neural Models And Dynamic Causal Modelling," *Neuroimage*, vol. 30, no. 4, pp. 1243–1254, May 2006. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2005.11.007

[55] O. David, "Dynamic Causal Models And Autopoietic Systems," *Biol Res*, vol. 40, no. 4, pp. 487–502, 2007. [Online]. Available: http://dx.doi.org//S0716-97602007000500010

[56] O. David, I. Guillemain, S. Saillet, S. Reyt, C. Deransart, C. Segebarth, and A. Depaulis, "Identifying Neural Drivers With Functional Mri: An Electrophysiological Validation," *PLoS Biol*, vol. 6, no. 12, pp. 2683–2697, Dec 2008. [Online]. Available: http://dx.doi.org/10.1371/journal.pbio.0060315

[57] Y. Saito and H. Harashima, "Tracking Of Information Within Multichannel EEG Record," *In: N. Yamaguchi and K. Fujisawa (Eds) Recent Advances In EEG And EMG Data Processing. Elsevier, Amsterdam*, pp. 133–146, 1981.

[58] R. Cohn and H. S. Leader, "Synchronization Characteristics Of Paroxysmal EEG Activity," *Electroencephalography and Clinical Neurophysiology*, vol. 22, no. 5, pp. 421 – 428, 1967.

[59] G. Carter, "Time Delay Estimation For Passive Sonar Signal Processing," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 29, no. 3, pp. 463 – 470, jun 1981.

[60] J. Ianniello, "Time Delay Estimation Via Cross-Correlation In The Presence Of Large Estimation Errors," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 30, no. 6, pp. 998 – 1003, dec 1982.

[61] J. Gotman, "Measurement Of Small Time Differences Between EEG Channels: Method And Application To Epileptic Seizure Propagation," *Electroencephalogr Clin Neurophysiol*, vol. 56, no. 5, pp. 501–514, Nov 1983.

[62] V. M. Eguíluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian, "Scale-Free Brain Functional Networks," *Phys. Rev. Lett.*, vol. 94, no. 1, p. 018102, Jan 2005.

[63] L. Amini, C. Jutten, S. Achard, O. David, H. Soltanian-Zadeh, G. A. Hossein-Zadeh, P. Kahane, L. Minotti, and L. Vercueil, "Directed Epileptic Network From Scalp And Intracranial EEG Of Epileptic Patients," in *IEEE International Workshop, MLSP09, Grenoble, France*, Sep. 2-4 2009.

[64] K. Friston, "Functional And Effective Connectivity In Neuroimaging: A Synthesis," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 56–78, 1994.

[65] A. A. Fingelkurts, A. A. Fingelkurts, and S. Kähkönen, "Functional Connectivity In The Brain–Is It An Elusive Concept?" *Neuroscience & Biobehavioral Reviews*, vol. 28, no. 8, pp. 827 – 836, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/B6T0J-4F6MKRP-3/2/5d8197d6a663083b67cff9d97f3e66bf

[66] B. Bollobas, *Advances In Graph Theory*. North-Holland, 1978.

[67] ——, *Graph Theory (Mathematics Studies)*. Elsevier, 1982.

[68] J. Akiyama, Y. Egawa, and H. Enomoto, *Graph Theory and Applications (Annals of discrete mathematics)*. Elsevier, 1988.

[69] B. Bollobas, *Random Graphs*. Cambridge University Press, 2001.

[70] J. L. Gross and J. Yellen, *Graph Theory.* Discrete Mathematics and Its Applications, 2003.

[71] J. Xu, *Theory and Application of Graphs (Network Theory and Applications).* Springer, 2003.

[72] J. Bondy and U. Murty, *Graph Theory.* Springer, 2008.

[73] M. E. J. Newman and J. Park, "Why social networks are different from other types of networks," *Phys. Rev. E*, vol. 68, no. 3, p. 036122, Sep 2003.

[74] S. H. Strogatz, "Exploring Complex Networks," *Nature*, vol. 410, no. 6825, pp. 268–276, March 2001. [Online]. Available: http://dx.doi.org/10.1038/35065725

[75] R. K. Albert, R. Albert, and A.-L. Barabási, "Statistical Mechanics of Complex Networks," *Reviews Of Modern Physics*, vol. 74, pp. 47–97, 2002.

[76] S. N. Dorogovtsev and J. Mendes, "Evolution of networks," *Advances in Physics*, vol. 51, pp. 1079–1187, Jun. 2002.

[77] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex Networks: Structure and Dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, 2006.

[78] O. Mason and M. Verwoerd, "Graph theory and networks in Biology," *Systems Biology, IET*, vol. 1, no. 2, pp. 89 –119, march 2007.

[79] C. Stam and J. Reijneveld, "Graph Theoretical Analysis Of Complex Networks In The Brain," *Nonlin Biomed Phys*, vol. 1, p. 3, 2007.

[80] E. Bullmore and O. Sporns, "Complex Brain Networks: Graph Theoretical Analysis Of Structural And Functional Systems," *Nature Reviews Neuroscience*, vol. 10, pp. 186–98, 2009, doi: 10.1038/nrn2574.

[81] V. Latora and M. Marchiori, "Economic Small-World Behavior in Weighted Networks," *Eur. Phys. J.*, vol. 32, pp. 249–263, 2003.

[82] *Advanced Biosignal Processing.* Springer, 2009. [Online]. Available: http://www.scribd.com/doc/23654335/Advanced-Biosignal-Processing

[83] K. Ansari-Asl, J. J. Bellanger, F. Bartolomei, F. Wendling, and L. Senhadji, "Time-Frequency Characterization of Interdependencies in Nonstationary Signals: Application to Epileptic EEG," *TBME*, vol. 52, pp. 1218–26, 2005.

[84] L. Astolfi, F. Cincotti, D. Mattia, M. Lai, L. Baccala, F. de Vico Fallani, S. Salinari, M. Ursino, M. Zavaglia, and F. Babiloni, "Comparison of Different Multivariate Methods for the Estimation of Cortical Connectivity: Simulations and Applications to EEG Data," in *Proc. of 27th Annual Conference of the IEEE, EMBS05, Shanghai, China*, September 1-4, 2005.

[85] G. E. P. Box and G. Jenkins, *Time Series Analysis; Forecasting And Control*. Holden-Day San Francisco, 1970.

[86] M. A. B. Brazier and J. U. Casby, "Crosscorrelation And Autocorrelation Studies Of Electroencephalographic Potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 4, no. 2, pp. 201 – 211, 1952. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-485RKDV-B/2/ab455244373d9d24ac0beb3fa85425d2

[87] J. S. Barlow and M. A. Brazier, "A Note On A Correlator For Electroencephalographic Work," *Electroencephalography and Clinical Neurophysiology*, vol. 6, pp. 321 – 325, 1954. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-482R3GY-BY/2/831379b548c0d395968ad5d46f1dc598

[88] K. Kuppusamy, W. Lin, and E. M. Haacke, "Statistical Assessment Of Crosscorrelation And Variance Methods And The Importance Of Electrocardiogram Gating In Functional Magnetic Resonance Imaging," *Magnetic Resonance Imaging*, vol. 15, no. 2, pp. 169 – 181, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/B6T9D-3RJ9F15-T/2/ae18e9b9b3cffbff17160d6d9dae5f6b

[89] B. Whitcher, P. Guttorp, and D. Percival, "Wavelet Analysis of Covariance with Application to Atmospheric Time Series," *J. Geophys. Res. - Atmospheres*, vol. 105, pp. 14 941–14 962, 2000.

[90] G. Carter, C. Knapp, and A. Nuttall, "Estimation of the magnitude-squared coherence function via overlapped fast fourier transform processing," *Audio and Electroacoustics, IEEE Transactions on*, vol. 21, no. 4, pp. 337 – 344, aug. 1973.

[91] G. Carter, "Coherence And Time Delay Estimation," *Proceedings of the IEEE*, vol. 75, no. 2, pp. 236 – 255, feb. 1987.

[92] B. P. van Milligen, E. Sanchez, T. Estrada, C. Hidalgo, B. Branas, B. Carreras, and L. Garcia, "Wavelet Bicoherence: A New Turbulence Analysis Tool," *Physics of Plasmas*, vol. 2, no. 8, pp. 3017–3032, 1995. [Online]. Available: http://link.aip.org/link/?PHP/2/3017/1

[93] P. L. Nunez, R. Srinivasan, A. F. Westdorp, R. S. Wijesinghe, D. M. Tucker, R. B. Silberstein, and P. J. Cadusch, "EEG Coherency: I: Statistics, Reference Electrode, Volume Conduction, Laplacians, Cortical Imaging, And Interpretation At Multiple Scales," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 5, pp. 499 – 515, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-449TYG5-1/2/2dfaa0c544993113c285829e45ef6bdc

[94] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring Phase Synchrony In Brain Signals," *Hum. Brain Mapp*, vol. 8, pp. 194–208, 1999.

[95] W. A. Gardner, "A Unifying View Of Coherence In Signal Processing," *Signal Processing*, vol. 29, no. 2, pp. 113 – 140, 1992. [Online]. Available: http://www.sciencedirect.com/science/article/B6V18-48XCY87-CW/2/eea94d8b22f1c205d206b127049a13ae

[96] J.-P. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. Le Van Quyen, J. Martinerie, and F. Varela, "Estimating The Time-Course Of Coherence Between Single-Trial Brain Signals: An Introduction To Wavelet Coherence," *Neurophysiol Clin*, vol. 32, no. 3, pp. 157–174, Jun 2002.

[97] M. G. Kitzbichler, M. L. Smith, S. R. Christensen, and E. Bullmore, "Broadband Criticality Of Human Brain Network Synchronization," *PLoS Comput Biol*, vol. 5, no. 3, p. e1000314, Mar 2009. [Online]. Available: http://dx.doi.org/10.1371/journal.pcbi.1000314

[98] P. Tass, M. G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H. Freund, "Detection Of n:m Phase Locking From Noisy Data: Application To Magnetoencephalography," *Physical Review Letters*, vol. 81, pp. 3291–3294, Oct. 1998.

[99] A. Pikovsky, M. Rosenblum, and J. Kurths, "Phase Synchronization In Regular And Chaotic Systems: A Tutorial," *International Journal of Bifurcation and Chaos*, vol. 10, pp. 2291–2305, 1999.

[100] K. Ansari-Asl, L. Senhadji, J. J. Bellanger, and F. Wendling, "Quantitative Evaluation Of Linear And Nonlinear Methods Characterizing Interdependencies Between Brain Signals," *Phys. Rev. E*, vol. 74, no. 3, p. 031916, Sep 2006.

[101] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, "Mean Phase Coherence As A Measure For Phase Synchronization And Its Application To The EEG Of Epilepsy Patients," *Physica D: Nonlinear Phenomena*, vol. 144, no. 3-4, pp. 358 – 369, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/B6TVK-411W4TY-7/2/bcd7a6c58942ced05427d69006044544

[102] M. Chavez, M. Besserve, C. Adam, and J. Martinerie, "Towards A Proper Estimation Of Phase Synchronization From Time Series," *Journal of Neuroscience Methods*, vol. 154, no. 1-2, pp. 149 – 160, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/B6T04-4J555SV-2/2/a914a4c469dcbb6061f16d980893004b

[103] D. Gabor, "Theory of communication," *Journal of the Institution of Electrical Engineers*, vol. 93, no. 26, pp. 442–457.

[104] E. Bedrosian, "A Product Theorem For Hilbert Transforms," *IEEE J PROC*, vol. 51, no. 5, pp. 868–869, 1963.

[105] J. J. Brown, "Analytic Signals And Product Theorems For Hilbert Transforms," *IEEE J CAS*, vol. 21, no. 6, pp. 790–792, 1974.

[106] B. Boashash, "Estimating And Interpreting The Instantaneous Frequency Of A Signal. I. Fundamentals," *IEEE J PROC*, vol. 80, no. 4, pp. 520–538, 1992.

[107] P. Panter, *Modulation, Noise, And Spectral Analysis.* McGraw-Hill, New York, 1965.

[108] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, "Phase Synchronization Of Chaotic Oscillators," *Phys. Rev. Lett.*, vol. 76, no. 11, pp. 1804–1807, Mar 1996.

[109] N. Delprat, B. Escudie, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torresani, "Asymptotic Wavelet And Gabor Analysis: Extraction Of Instantaneous Frequencies," *IEEE J IT*, vol. 38, no. 2, pp. 644–664, 1992.

[110] C. E. Shannon, "A Mathematical Theory Of Communication," *The Bell System Technical Journal*, vol. 27, no. 1, pp. 379–423, 1948.

[111] N. Wiener, *The Theory Of Prediction.* In: Beckenbach EF, editors. Modern Mathematics for Engineers. New York: McGraw-Hill, 1956.

[112] C. W. J. Granger, "Investigating Causal Relations By Econometric Models And Cross-Spectral Methods," *Econometrica*, vol. 37, pp. 424–438, 1969. [Online]. Available: http://www.jstor.org/stable/1912791

[113] P. Ktonas and R. Mallart, "Estimation Of Time Delay Between EEG Signals For Epileptic Focus Localization: Statistical Error Considerations," *Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 2, pp. 105 – 110, 1991.

[114] J. Geweke, "Measurement Of Linear Dependence And Feedback Between Multiple Time Series," *Journal of the American Statistical Association*, vol. 77, no. 378, pp. 304–313, 1982. [Online]. Available: http://www.jstor.org/stable/2287238

[115] K. Sameshima and L. Baccalaá, "Using partial directed coherence to describe neuronal ensemble interactions," *J Neurosci Methods*, vol. 94, pp. 93–103, 1999.

[116] E. Baek and W. Brock, "A General Test For Nonlinear Granger Causality," Working Paper University of Iowa, Tech. Rep., 1992.

[117] B. Gourévitch, R. L. Bouquin-Jeannès, and G. Faucon, "Linear And Nonlinear Causality Between Signals: Methods, Examples And Neurophysiological Applications," *Biol Cybern*, vol. 95, no. 4, pp. 349–369, Oct 2006. [Online]. Available: http://dx.doi.org/10.1007/s00422-006-0098-0

[118] V. Fernandes de Lima, J. P. Pijn, C. Nunes Filipe, and F. Lopes da Silva, "The Role Of Hippocampal Commissures In The Interhemispheric Transfer Of Epileptiform Afterdischarges In The Rat: A Study Using Linear And Non-Linear Regression Analysis," *Electroencephalography and Clinical Neurophysiology*, vol. 76, no. 6, pp. 520 – 539, 1990. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-4859SCF-1F/2/86544d26cb7d30ae89711368bc826a39

[119] P. Allen, S. Smith, and C. Scott, "Measurement Of Interhemispheric Time Differences In Generalised Spike-And-Wave," *Electroencephalography and Clinical Neurophysiology*, vol. 82, no. 1, pp. 81 – 84, 1992. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-482RBB4-CJ/2/e63b82a41642e02c1a5df066ca59aa52

[120] S. J. Schiff, P. So, T. Chang, R. E. Burke, and T. Sauer, "Detecting Dynamical Interdependence and Generalized Synchrony Through Mutual Prediction in a Neural Ensemble," *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, vol. 54, no. 6, pp. 6708–6724, Dec 1996.

[121] J. Arnhold, P. Grassberger, K. Lehnertz, and C. E. Elger, "A Robust Method For Detecting Interdependences: Application To Intracranially Recorded EEG," *Physica D: Nonlinear Phenomena*, vol. 134, no. 4, pp. 419 – 430, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/B6TVK-3XNK626-4/2/2417f7cf6467dfbca54131bbab005bf7

[122] R. Q. Quian, J. Arnhold, and P. Grassberger, "Learning Driver-Response Relationships From Synchronization Patterns," *Phys Rev. E*, vol. 61, pp. 5142–5148, 2000.

[123] C. J. Stam and B. W. van Dijk, "Synchronization Likelihood: An Unbiased Measure Of Generalized Synchronization In Multivariate Data Sets," *Physica D: Nonlinear Phenomena*, vol. 163, no. 3-4, pp. 236 – 251, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/B6TVK-44YWT15-1/2/24cf5b2e1ff3aabb008c7d1b275d3d91

[124] D. A. Smirnov and R. G. Andrzejak, "Detection Of Weak Directional Coupling: Phase-Dynamics Approach Versus State-Space Approach," *Phys. Rev. E*, vol. 71, no. 3, p. 036207, Mar 2005.

[125] M. Palus, V. Komárek, T. Procházka, Z. Hrncír, and K. Sterbová, "Synchronization And Information Flow In EEG Of Epileptic Patients," *IEEE Eng Med Biol Mag*, vol. 20, no. 5, pp. 65–71, 2001.

[126] K. K. Jerger, T. I. Netoff, J. T. Francis, T. Sauer, L. Pecora, S. L. Weinstein, and S. J. Schiff, "Early Seizure Detection," *J Clin Neurophysiol*, vol. 18, no. 3, pp. 259–268, May 2001.

[127] F. Mormann, T. Kreuz, C. Rieke, R. G. Andrzejak, A. Kraskov, P. David, C. Elger, and K. Lehnertz, "On the predictability of epileptic seizures," *Clinical Neurophysiology*, vol. 116, no. 3, pp. 569 – 587, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/B6VNP-4F662KT-2/2/4a9a19afab12eedc4d69bf901017a3ba

[128] E. Asano, C. Juhász, A. Shah, S. Sood, and H. T. Chugani, "Role Of Subdural Electrocorticography In Prediction Of Long-Term Seizure Outcome In Epilepsy Surgery," *Brain*, vol. 132, no. Pt 4, pp. 1038–1047, Apr 2009. [Online]. Available: http://dx.doi.org/10.1093/brain/awp025

[129] H. M. Hamer, H. H. Morris, E. J. Mascha, M. T. Karafa, W. E. Bingaman, M. D. Bej, R. C. Burgess, D. S. Dinner, N. R. Foldvary, J. F. Hahn, P. Kotagal, I. Najm, E. Wyllie, and H. O. Lüders, "Complications of Invasive Video-EEG Monitoring With Subdural Grid Electrodes." *Neurology*, vol. 58, no. 1, pp. 97–103, Jan 2002.

[130] C. Onal, H. Otsubo, T. Araki, S. Chitoku, A. Ochi, S. Weiss, I. Elliott, O. C. Snead, J. T. Rutka, and W. Logan, "Complications of invasive subdural grid monitoring in children with epilepsy." *J Neurosurg*, vol. 98, no. 5, pp. 1017–1026, May 2003. [Online]. Available: http://dx.doi.org/10.3171/jns.2003.98.5.1017

[131] E. B. Bromfield, "Epileptiform Discharges," *emedicine*, vol. http://emedicine.medscape.com/article/1138880-overview, 2009.

[132] G. Alarcon, "Electrophysiological Aspects Of Interictal And Ictal Activity In Human Partial Epilepsy," *Seizure*, vol. 5, no. 1, pp. 7–33, Mar 1996.

[133] Y. Lai, W. van Drongelen, K. Hecox, D. Frim, M. Kohrman, and B. He, "Cortical Activation Mapping Of Epileptiform Activity Derived From Interictal ECoG Spikes," *Epilepsia*, vol. 48, no. 2, pp. 305–314, Feb 2007.

[134] J. Bourien, F. Bartolomei, J. Bellanger, M. Gavaret, P. Chauvel, and F. Wendling, "A Method To Identify Reproducible Subsets Of Co-Activated Structures During Interictal Spikes. Application To Intracerebral EEG In Temporal Lobe Epilepsy," *Clinical Neurophysiology*, vol. 116, no. 2, pp. 443 – 455, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/B6VNP-4DFT77T-1/2/1ef24178e2d69ee13768e35702604ab2

[135] W. Penfield and H. Jasper, "Epilepsy and the Functional Anatomy of the Human Brain." *London: J. & A. Churchill.*

[136] X. Zhang, W. Van Drongelen, K. E. Hecox, V. L. Towle, D. M. Frim, A. B. McGee, and B. He, "High-Resolution EEG: Cortical Potential Imaging Of Interictal Spikes," *Clin Neurophysiol*, vol. 114, no. 10, pp. 1963–1973, Oct 2003.

[137] E. Asano, O. Muzik, A. Shah, C. Juhász, D. C. Chugani, K. Kagawa, K. Benedek, S. Sood, J. Gotman, and H. Chugani, "quantitative Visualization Of Ictal Subdural Eeg Changes In Children With Neocortical Focal Seizures," *Clin Neurophysiol*, vol. 115, no. 12, pp. 2718–2727, Dec 2004. [Online]. Available: http://dx.doi.org/10.1016/j.clinph.2004.06.020

[138] S. Micheloyannis, E. Pachou, C. J. Stam, M. Vourkas, S. Erimaki, and V. Tsirka, "Using Graph Theoretical Analysis Of Multi Channel EEG To Evaluate The Neural Efficiency Hypothesis," *Neuroscience Letters*, vol. 402, no. 3, pp. 273 – 277, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/B6T0G-4JWFMPV-5/2/7aa139467cd032988df41e3c72c8a74c

[139] S. Ponten, F. Bartolomei, and C. Stam, "Small-World Networks and Epilepsy: Graph Theoretical Analysis of Intracerebrally Recorded Mesial Temporal Lobe Seizures," *Clinical Neurophysiology*, vol. 118, no. 4, pp. 918 – 927, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/B6VNP-4N3H16J-BC/2/77ba1beb41045945cd00d3e3416debab

[140] R. Srinivasan, W. R. Winter, J. Ding, and P. L. Nunez, "EEG and MEG Coherence: Measures of Functional Connectivity at Distinct Spatial Scales of Neocortical Dynamics," *Journal of Neuroscience Methods*, vol. 166, pp. 41–52, October 2007.

[141] K. Supekar, V. M. D. Rubin, and M. M. M. D. Greicius, "Network Analysis Of Intrinsic Functional Brain Connectivity In Alzheimer'S Disease," *PLoS Comput Biol*, vol. 4, no. 6, p. e1000100, 06 2008. [Online]. Available: http://dx.plos.org/10.1371%2Fjournal.pcbi.1000100

[142] S. Pajevic and D. Plenz, "Efficient network reconstruction from dynamical cascades identifies small-world topology of neuronal avalanches," *PLoS Comput Biol*, vol. 5, no. 1, p. e1000271, 01 2009. [Online]. Available: http://dx.doi.org/10.1371%2Fjournal.pcbi.1000271

[143] S. Grun, "Data-Driven Significance Estimation For Precise Spike Correlation," *J Neurophysiol*, vol. 101, no. 3, pp. 1126–1140, 2009. [Online]. Available: http://jn.physiology.org/cgi/content/abstract/101/3/1126

[144] P. H. Westfall and S. S. Young, *Resampling-Based Multiple Testing: Examples And Methods For P-Value Adjustment.* Wiley, 1993.

[145] K. S. Pollard and M. J. van der Laan, "Resampling-Based Multiple Testing: Asymptotic Control Of Type I Error And Applications To Gene Expression Data," *U.C. Berkeley Division of Biostatistics Working Paper Series*, p. Working Paper 121, 2003.

[146] M. Kaminski and K. Blinowska, "A New Method Of The Description Of The Information Flow In The Brain Structures," *Biol Cybern*, vol. 65, pp. 203–210, 1991.

[147] M. Kaminski, M. Ding, W. Truccolo, and S. Bressler, "Evaluating Causal Relations In Neural Systems: Granger Causality, Directed Transfer Function And Statistical Assessment Of Significance," *Biol. Cybern.*, vol. 85, pp. 145–157, 2001.

[148] L. Amini, S. Achard, C. Jutten, H. Soltanian-Zadeh, G. A. Hossein-Zadeh, O. David, and L. Vercueil, "Sparse Differential Connectivity Graph of Scalp EEG for Epileptic Patients," in *ESANN2009, Bruges, Belgium*, Apr. 22-24 2009.

[149] K. Deb, "Multi-Objective Evolutionary Algorithms: Introducing Bias Among Pareto-Optimal Solutions," *Kanpur Genetic Algorithms Lab (KanGal), Technical report 99002*, 1999.

[150] J. Branke, K. Deb, K. Miettinen, and R. Slowinski, *Multiobjective Optimization, Interactive And Evolutionary Approaches*. Springer, 2008.

[151] J. Triesch and C. von der Malsburg, "Democratic integration: Self-organized integration of adaptive cues," *Neural Computation*, vol. 13, no. 9, pp. 2049–2074, 2001. [Online]. Available: http://www.mitpressjournals.org/doi/abs/10.1162/089976601750399308

[152] A. R. Mcintosh, "Mapping Cognition To The Brain Through Neural Interactions," *Memory*, vol. 7, pp. 523–548, 1999.

[153] E. R. John, "A Field Theory of Consciousness," *Consciousness and Cognition*, vol. 10, no. 2, pp. 184 – 213, 2001. [Online]. Available: http://www.sciencedirect.com/science/article/B6WD0-45BBW60-11/2/6bba2f0a5031e9ce41e7f52b1c4749d2

[154] J. Fell, A. Kaplan, B. Darkhovsky, and J. Röschke, "EEG Analysis With Nonlinear Deterministic And Stochastic Methods: A Combined Strategy," *Acta Neurobiol Exp (Wars)*, vol. 60, no. 1, pp. 87–108, 2000.

[155] D. Van De Ville, J. Britz, and C. M. Michel, "EEG Microstate Sequences In Healthy Humans At Rest Reveal Scale-Free Dynamics," *Proceedings of the National Academy of Sciences*, pp. 1–7.

[156] X. Li, J. Polygiannakis, P. Kapiris, K. Eftaxias, and X. Yao, "Fractal Spectral Analysis Of Pre-Epileptic Seizures Phase: In Terms Of Criticality," *J Neural Eng*, vol. 2, pp. 11–16, 2005.

[157] K. Indiradevi, E. Elias, P. Sathidevi, S. D. Nayak, and K. Radhakrishnan, "A Multi-Level Wavelet Approach For Automatic Detection Of Epileptic Spikes In The Electroencephalogram," *Computers in Biology and Medicine*, vol. 38, no. 7, pp. 805 – 816, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/B6T5N-4SR6FXJ-3/2/caf5cde2d68c3bb655e8e890a3eb87d2

[158] T. Conlon, H. Ruskin, and M. Crane, "Seizure Characterisation Using Frequency-Dependent Multivariate Dynamics," *Computers in Biology and Medicine*, vol. 39, no. 9, pp. 760 – 767, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/B6T5N-4WP4BJJ-1/2/eef11294fb8e1983ff6f8c4e17b97f3e

[159] I. Clark, R. Biscay, M. Echeverría, and T. Virués, "Multiresolution Decomposition Of Non-Stationary EEG Signals: A Preliminary Study," *Computers in Biology and Medicine*, vol. 25, no. 4, pp. 373 – 382, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/B6T5N-3YF49BD-D/2/fcb6563c6cf54171eab1f13b4531789c

[160] B. Whitcher, "Assessing Nonstationary Time Series Using Wavelets," Ph.D. dissertation, Ph.D. thesis, University of Washington, 1998.

[161] L. Senhadji and F. Wendling, "Epileptic Transient Detection: Wavelets And Time-Frequency Approaches," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 32, no. 3, pp. 175 – 192, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/B6VMP-464KJ15-4/2/0cc356a454eae99d15af82b449a4faf5

[162] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis Of EEG Records In An Epileptic Patient Using Wavelet Transform," *Journal of Neuroscience Methods*, vol. 123, no. 1, pp. 69 – 87, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/B6T04-47VPXWC-4/2/86fc6619f794e34f51685d64597d50f1

[163] C. Yamaguchi, "Fourier And Wavelet Analyses Of Normal And Epileptic Electroencephalogram (EEG)," in *Neural Engineering, First International IEEE EMBS Conference*, 2003, pp. 406 – 409.

[164] G. Ouyang, X. Li, Y. Li, and X. Guan, "Application Of Wavelet-Based Similarity Analysis To Epileptic Seizures Prediction," *Computers in Biology and Medicine*, vol. 37, no. 4, pp. 430 – 437, 2007, wavelet-based Algorithms for Medical Problems. [Online]. Available: http://www.sciencedirect.com/science/article/B6T5N-4KWK16C-1/2/01ac8a610524abc2c1d739ddb8bf3edf

[165] D. I., *Ten Lectures on Wavelets*. CBMSNSF Reg. Conf. Series in Applied Math. SIAM, 1992.

[166] I. Daubechies, "Orthonormal Bases Of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909–996, 1988. [Online]. Available: http://dx.doi.org/10.1002/cpa.3160410705

[167] D. Percival, "On Estimation Of The Wavelet Variance," *Biometrika*, vol. 82, pp. 619–631, 1995.

[168] D. B. Percival and H. O. Mofjeld, "Analysis Of Subtidal Coastal Sea Level Fluctuations Using Wavelets," *Journal of the American Statistical Association*, vol. 92, pp. 868–880, 1997.

[169] A. Naït-Ali, *Advanced Biosignal Processing*. Springer, 2009.

[170] V. Menon, W. J. Freeman, B. A. Cutillo, J. E. Desmond, M. F. Ward, S. L. Bressler, K. D. Laxer, N. Barbaro, and A. S. Gevins, "Spatio-Temporal Correlations In Human Gamma Band Electrocorticograms," *Electroencephalography and Clinical Neurophysiology*, vol. 98, no. 2, pp. 89 – 102, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-3VXH769-1/2/345d96708704c1d3b97f693a93c8cffc

[171] C. P. Pawela, B. B. Biswal, Y. R. Cho, D. S. Kao, R. Li, S. R. Jones, M. L. Schulte, H. S. Matloub, A. G. Hudetz, and J. S. Hyde, "Resting-State Functional Connectivity of the Rat Brain," *Magnetic Resonance in Medicine*, vol. 59, pp. 1021–1029, 2008.

[172] J. Gotman, "Interhemispheric Interactions In Seizures Of Focal Onset: Data From Human Intracranial Recordings," *Electroencephalogr Clin Neurophysiol*, vol. 67, no. 2, pp. 120–133, Aug 1987.

[173] J. P. Lieb, K. Hoque, C. E. Skomer, and X. Song, "Inter-Hemispheric Propagation Of Human Mesial Temporal Lobe Seizures: A Coherence/Phase Analysis," *Electroencephalography and Clinical Neurophysiology*, vol. 67, no. 2, pp. 101 – 119, 1987. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-482XKWT-25C/2/0042c06773d3ccbe933de512462919a3

[174] G. Fein, J. Raz, F. F. Brown, and E. L. Merrin, "Common Reference Coherence Data Are Confounded By Power And Phase Effects," *Electroencephalography and Clinical Neurophysiology*, vol. 69, no. 6, pp. 581 – 584, 1988. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-482XNCR-2S7/2/cc3f80a61d0a7f11662b7ea6df8d7b0a

[175] R. B. Duckrow and S. S. Spencer, "Regional Coherence And The Transfer Of Ictal Activity During Seizure Onset In The Medial Temporal Lobe," *Electroencephalography and Clinical Neurophysiology*, vol. 82, no. 6, pp. 415 – 422, 1992. [Online]. Available: http://www.sciencedirect.com/science/article/B6SYX-482XG41-49/2/820f4c9f17ebb0dba23bc678bbec2f3c

[176] V. L. Towle, R. K. Carder, L. Khorasani, and D. Lindberg, "Electrocorticographic Coherence Patterns." *J Clin Neurophysiol*, vol. 16, no. 6, pp. 528–547, Nov 1999.

[177] B. Guntekin, E. Saatci, and G. Yener, "Decrease of Evoked Delta, Theta and Alpha Coherences in Alzheimer Patients During a Visual Oddball Paradigm," *Brain Research*, vol. 1235, pp. 109–116, 2008.

[178] J. P. Pijn, P. C. Vijn, F. H. L. da Silva, W. V. E. Boas, and W. Blanes, "Localization Of Epileptogenic Foci Using A New Signal Analytical Approach," *Neurophysiol Clin*, vol. 20, no. 1, pp. 1–11, Apr 1990.

[179] G. Bettus, F. Wendling, M. Guye, L. Valton, J. R´egis, P. Chauvel, and F. Bartolomei, "Enhanced EEG Functional Connectivity in Mesial Temporal Lobe Epilepsy," *Epilepsy research*, vol. 81, pp. 58–68, 2008.

[180] M. Shen, L. Sun, K. Ting, and F. Chan, "The Investigation Of Time-Varying Synchrony Of EEG During Sentence Learning Using Wavelet Analysis," in *Neural Information Processing, 2002. ICONIP '02. Proceedings of the 9th International Conference on*, vol. 1, Nov. 2002, pp. 92–95 vol.1.

[181] I. Gelfand and A. Yaglom, "Calculation Of Amount Of Information About A Random Function Contained In Another Such Function," *Amer. Math. Soc. Trans*, vol. 12, pp. 199–246, 1959.

[182] N. Mars and G. van Arragon, "Time Delay Estimation In Non-Linear Systems Using Average Amount Of Mutual Information Analysis," *Signal Processing*, vol. 4, no. 2-3, pp. 139 – 153, 1982. [Online]. Available: http://www.sciencedirect.com/science/article/B6V18-4903TGB-1F/2/a4ee40a5606aa17f48d206c81bb313ee

[183] R. Moddemeijer, "A Statistic To Estimate The Variance Of The Histogram-Based Mutual Information Estimator Based On Dependent Pairs Of Observations," *Signal Process.*, vol. 75, no. 1, pp. 51 – 63, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/B6V18-3WHCHYS-5/2/262e2fe30514430ce2657d0ff4a500bc

[184] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 69, no. 6 Pt 2, p. 066138, Jun 2004.

[185] J. F. Alonso, M. A. Mananas, S. Romero, J. Riba, M. Barbanoj, and D. Hoyer, "Connectivity Analysis of EEG Under Drug Therapy," in *Proc. of 29th Annual International Conference of the IEEE, EMBS07*, 22-26 Aug, 2007.

[186] L. Astolfi, F. Cincotti, D. Mattia, F. De Vico Fallani, A. Tocci, A. Colosimo, S. Salinari, M. G. Marciani, W. Hesse, H. Witte, M. Ursino, M. Zavaglia, , and F. Babiloni, "Tracking the Time-Varying Cortical Connectivity Patterns by Adaptive Multivariate Estimators," *IEEE Trans Biomed Eng.*, vol. 55, pp. 902–13, 2008.

[187] F. Babiloni, L. Astolfi, F. Cincotti, D. Mattia, A. Tocci, A. Tarantino, M. G. Marciani, S. Salinari, S. Gao, A. Colosimo, and F. De Vico Fallani, "Cortical Activity and Connectivity of Human Brain during the Prisoner's Dilemma: an EEG Hyperscanning Study," in *Proc. of 29th Annual International Conference of the IEEE, EMBS07*, 22-26 Aug, 2007.

[188] L. A. Baccalá and K. Sameshima, "Partial Directed Coherence: A New Concept In Neural Structure Determination," *Biol Cybern*, vol. 84, no. 6, pp. 463–474, Jun 2001.

[189] B. Pompe, "Measuring statistical dependences in a time series," *Journal of Statistical Physics*, vol. 73, pp. 587–610, 1993.

[190] K. S. Pollard and M. J. van der Laan, "Choice of a null distribution in resampling-based multiple testing," *Journal of Statistical Planning and Inference*, vol. 125, no. 1-2, pp. 85 – 100, 2004, the Third International Conference on Multiple Comparisons. [Online]. Available: http://www.sciencedirect.com/science/article/B6V0M-4BVPB6R-2/2/ecf33ef8cf4844b237027f9ce2972748

[191] P. H. Westfall and J. F. Troendle, "Multiple Testing With Minimal Assumptions," *Biometrical Journal*, vol. 50, pp. 745–755, 2008.

[192] D. Draper, J. Hodges, C. L. Mallows, and D. Pregibon, "Exchangeability And Data Analysis," *Journal of the Royal Statistical Society*, vol. 156, pp. 9–28, 1993.

[193] S. Holm, " A Simple Sequentially Rejective Multiple Test Procedure," *Scand J. Stat.*, vol. 6, pp. 65–70, 1979.

[194] K. Deb, M. Mohan, and S. Mishra, "Evaluating The Epsilon-Domination Based Multi-Objective Evolutionary Algorithm For A Quick Computation Of Pareto-Optimal Solutions," *Evol Comput*, vol. 13, no. 4, pp. 501–525, 2005. [Online]. Available: http://dx.doi.org/10.1162/106365605774666895

[195] K. Deb and H. Gupta, "Introducing Robustness In Multi-Objective Optimization," *Evol Comput*, vol. 14, no. 4, pp. 463–494, 2006. [Online]. Available: http://dx.doi.org/10.1162/evco.2006.14.4.463

[196] C. A. C. Coello, *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, 2007.

[197] M. Zeleny, *Compromise Programming*, J. Cochrane and M. Zeleny, Eds. Columbia: University of South Carolina Press, 1973.

[198] P. Kahane, P. Ryvlin, D. Hoffmann, L. Minotti, and A. L. Benabid, "From Hypothalamic Hamartoma To Cortex: What Can Be Learnt From Depth Recordings And Stimulation?" *Epileptic Disord*, vol. 5, no. 4, pp. 205–217, Dec 2003.

[199] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, Inc., 2006.

[200] H. P. Zaveri, R. B. Duckrow, and S. S. Spencer, "The Effect Of A Scalp Reference Signal On Coherence Measurements Of Intracranial Electroencephalograms," *Clinical Neurophysiology*, vol. 111, no. 7, pp. 1293 – 1299, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/B6VNP-40MT2S1-N/2/3c27c034f28800b75b6aaf75a54fa870

[201] J. P. Lachaux, D. Rudrauf, and P. Kahane, "Intracranial EEG And Human Brain Mapping," *Journal of Physiology-Paris*, vol. 97, no. 4-6, pp. 613 – 628, 2003, neuroscience and Computation. [Online]. Available: http://www.sciencedirect.com/science/article/B6VMC-4BNMR71-2/2/f2de289dc434de436d5ccf543d364313

[202] C. Tallon-Baudry, "Oscillatory Synchrony And Human Visual Cognition," *Journal of Physiology-Paris*, vol. 97, no. 2-3, pp. 355 – 363, 2003, neurogeometry and visual perception. [Online]. Available: http://www.sciencedirect.com/science/article/B6VMC-4B9K947-1/2/a95ffece4b3d75670e0b6815ebd1bd28

[203] J. D. Olden and B. D. Neffd, "Cross-Correlation Bias In Lag Analysis Of Aquatic Time Series," *Marine Biology*, vol. 138, pp. 1063–1070, 2001.

[204] O. David, L. Harrison, and K. J. Friston, "Modelling event-related responses in the brain," *NeuroImage*, vol. 25, no. 3, pp. 756–770, 2005.

[205] T. Blauwblomme, P. Kahane, L. Minotti, F. Grouiller, A. Krainik, L. Vercueil, S. Chabardès, D. Hoffmann, and O. David, "Multimodal imaging reveals the role of activity in eating-reflex seizures," *Journal of Neurology, Neurosurgery & Psychiatry*, in press.

[206] E. Asano, O. Muzik, A. Shah, C. Juhász, D. C. Chugani, S. Sood, J. Janisse, E. L. E., J. Ahn-Ewing, C. Shen, J. Gotman, and H. T. Chugani, "Quantitative Interictal Subdural EEG Analyses In Children With Neocortical Epilepsy," *Epilepsia*, vol. 44, no. 3, pp. 425–434, Mar 2003.

[207] P. Chauvel, E. Landré, S. Trottier, J. P. Vignel, A. Biraben, B. Devaux, and J. Bancaud, "Electrical Stimulation With Intracerebral Electrodes To Evoke Seizures," *Adv Neurol*, vol. 63, pp. 115–121, 1993.

[208] P. Kahane, L. Tassi, S. Francione, D. Hoffmann, G. L. Russo, and C. Munari, "Electroclinical Manifestations Elicited By Intracerebral Electric Stimulation "Shocks" In Temporal Lobe Epilepsy," *Neurophysiol Clin*, vol. 23, no. 4, pp. 305–326, Jul 1993.

[209] M. Paluš, V. Komárek, Z. c. v. Hrnčíř, and K. Štěrbová, "Synchronization As Adjustment of Information Rates: Detection From Bivariate Time Series," *Phys. Rev. E*, vol. 63, no. 4, pp. 0 462 111–6, Mar 2001.

[210] P. J. Franaszczuk, K. J. Blinowska, and M. Kowalczyk, "The Application of Parametric Multichannel Spectral Estimates in the Study of Electrical Brain Activity," *Biological Cybernetics*, vol. 51, pp. 239–247, 1985.

[211] M. Chavez, M. Le Van Quyen, V. Navarro, M. Baulac, and J. Martinerie, "Spatio-Temporal Dynamics Prior to Neocortical Seizures: Amplitude Versus Phase Couplings," *Biomedical Engineering, IEEE Transactions on*, vol. 50, no. 5, pp. 571–583, May 2003.

[212] L. Astolfi, F. De Vico Fallani, F. Cincotti, D. Mattia, M. G. Marciani, S. Bufalari, S. Salinari, A. Colosimo, L. Ding, J. C. Edgar, W. Heller, G. A. Miller, B. He, and F. Babiloni, "Imaging Functional Brain Connectivity Patterns From High-Resolution EEG and fMRI Via Graph Theory," *Psychophysiology*, vol. 44, pp. 880–893, 2007.

[213] D. Mantini, S. Cugini, G. L. Romani, and C. Del Gratta, "Fusion of EEG and fMRI for the Investigation of Functional Connectivity During a Visual Oddball Task," *NFSI-ICFBI07*, pp. 373–376, Oct. 2007.

[214] D. Percival and A. Walden, "Wavelet Methods for Time Series Analysis," *Cambridge, UK: Cambridge UP*, 2000.

[215] L. Hudgings, C. Friehe, and M. E. Mayer, "Wavelet Transforms and Atmospheric Turbulence," *Physics Review Letters*, vol. 71, pp. 3279–3282, 1993.

[216] L. Amini, R. Sameni, C. Jutten, G. A. Hossein-Zadeh, and H. Soltanian-Zadeh, "MR Artifact Reduction in the Simultaneous Acquisition of EEG and fMRI of Epileptic Patients," in *EUSIPCO08, Lausanne, Switzerland*, 2008.

[217] F. Wendling, A. Hernandez, J.-J. Bellanger, P. Chauvel, and F. Bartolomei, "Interictal To Ictal Transition In Human Temporal Lobe Epilepsy: Insights From A Computational Model Of Intracerebral EEG," *J Clin Neurophysiol*, vol. 22, no. 5, pp. 343–356, Oct 2005.

[218] D. Cosandier-Rimélé, J.-M. Badier, P. Chauvel, and F. Wendling, "A Physiologically Plausible Spatio-Temporal Model for EEG Signals Recorded With Intracerebral Electrodes in Human Partial Epilepsy." *Biomedical Engineering, IEEE Transactions on*, vol. 54, no. 3, pp. 380–388, March 2007.

[219] M. Negishi, B. I. Pinus, A. B. Pinus, and R. T. Constable, "Origin of the Radio Frequency Pulse Artifact in Simultaneous EEG-fMRI Recording: Rectification at the Carbon-Metal Interface," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 9, pp. 1725–1727, September 2007.

[220] P. Xu, H. Chen, Z. Liu, and D. Yao, "A New Method Based on Sparse Component Decomposition to Remove MRI Artifacts in the Continuous EEG Recordings," in *Proc. of the 2005 IEEE Trans. Biomed. Eng. 27th Annu. Conf.*, Shanghai (China), September 1–4 2005, pp. 2006–2008.

[221] M. Jing and S. Sanei, "Scanner Artifact Removal in Simultaneous EEG-fMRI for Epileptic Seizure Prediction," in *Proc. of the 18th Int. Conf. on Pattern Recognition (ICPR'06)*, 2006, pp. 722–725.

[222] X. Wan, K. Iwata, J. Riera, M. Kitamura, and R. Kawashima, "Artifact Reduction For Simultaneous EEG/fMRI Recording: Adaptive Fir Reduction Of Imaging Artifacts," *Clinical Neurophysiology*, vol. 117, no. 3, pp. 681 – 692, 2006.

[223] R. Sameni, C. Jutten, and M. Shamsollahi, "Multichannel Electrocardiogram Decomposition Using Periodic Component Analysis," *Biomedical Engineering, IEEE Transactions on*, vol. 55, no. 8, pp. 1935 –1940, 2008.

[224] L. Parra and P. Sajda, "Blind Source Separation via Generalized Eigenvalue Decomposition," *Journal of Machine Learning Research*, pp. 1261–1269, 2003.

[225] L. Tong, R.-W. Liu, V. Soon, and Y.-F. Huang, "Indeterminacy and Identifiability of Blind Identification," *IEEE Trans. Circuits Syst.*, no. 5, pp. 499–509, May 1991.

[226] G. Strang, *Linear Algebra and Its Applications*, 3rd ed. Brooks/Cole, 1988.

[227] G. R. Arce, *Nonlinear Signal Processing: A Statistical Approach.* New York: John Wiley & Sons Inc., 2004.

[228] R. Sameni, "Removing ECG Contaminants from Multichannel Recordings by Deflation," GIPSA-LAB, INP-Grenoble, Tech. Rep., September 2007.

[229] A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999. [Online]. Available: http://www.sciencemag.org/cgi/content/abstract/286/5439/509

[230] M. J. Alava and S. N. Dorogovtsev, "Complex Networks Created By Aggregation," *Phys. Rev. E*, vol. 71, no. 3, p. 036107, Mar 2005.

[231] T.-A. Hochberg, Y., *Multiple Comparison Procedures.* Wiley, New York, 1987.

[232] H. Abdi, "Bonferroni And Sidak Corrections For Multiple Comparisons," *N.J. Salkind (ed.), Encyclopedia of Measurement and Statistics,Thousand Oaks, CA: Sage*, 2007.

[233] Z. Sidak, "Rectangular Confidence Region For The Means Of Multivariate Normal Distributions," *Journal of the American Statistical Association*, vol. 62, pp. 626–633, 1967.

[234] J. W. Tukey, "Some Thoughts On Clinical Trials, Especially Problems Of Multiplicity," *Science*, vol. 198, no. 4318, pp. 679–684, Nov 1977.

[235] M. R. Leadbetter, G. Lindgren, and H. Rootzen, *Extremes And Related Properties Of Random Sequences And Processes: Springer Series In Statistics.* Springer, 1983.