Multi-level Characterization and Information Extraction in Directed and Node-labeled Functional Brain Networks

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

vorgelegt dem Rat der Medizinischen Fakultät der

Friedrich-Schiller-Universität Jena

von Dipl.-Bioinf. Christoph Schmidt

geboren am 30.07.1982 in Saalfeld

Gutachter:

1. Prof. Dr. Herbert Witte (Universitätsklinikum Jena)

2. Prof. Dr. Stefan Schuster (Friedrich-Schiller-Universität Jena)

3. Prof. Dr. Jens Haueisen (Technische Universität Ilmenau)

Tag der Einreichung:

12. August 2016

Tag der öffentlichen Verteidigung:

02. Mai 2017

Zusammenfassung

Ein Hauptforschungsschwerpunkt in den Neurowissenschaften liegt in der Berechnung und Analyse der funktionellen Konnektivität des Gehirns, d. h. der temporalen statistischen Abhängigkeiten (gerichtete Interaktionen) zwischen der an verschiedenen Messpunkts, z. B. an EEG-Elektroden oder fMRI-Voxeln, elektrophysiologisch oder hämodynamisch aufgenommenen Hirnaktivität. Die methodischen Entwicklungen dieser Arbeit beziehen sich auf die umfassende Analyse von Netzwerken, welche die zuvor berechneten funktionellen Konnektivitätsmuster repräsentieren. Im Allgemeinen sind diese aus der Konnektivitätsanalyse resultierenden Netzwerke aufgrund ihrer komplizierten Struktur keiner unmittelbaren Interpretation und dem damit verbundenen Erkenntnisgewinn zugänglich, so dass weiterführende Untersuchungen mittels Analysekonzepten aus der Netzwerktheorie unumgänglich sind. Einen besonderen Schwerpunkt der Dissertation mit Hinblick auf die klinischen Daten, deren Auswertung die Methodenentwicklungen maßgeblich geprägt haben, bildet die Analyse und Informationsextraktion von Mengen funktioneller Gehirnnetzwerke, die spezifische Konnektivitätsmuster unterschiedlicher Probandengruppen widerspiegeln. Mittels gruppenspezifischer Netzwerkanalysen können relevante Eigenschaften der Konnektivitätsmuster erfasst und quantifiziert werden und somit schlussendlich interpretatorisch in Relation zu der zugrundeliegenden neuronalen Informationsverarbeitung zwischen Hirnarealen gesetzt werden. Wie in der Arbeit beschrieben sind solche Analysen und die Extraktion physiologisch relevanter Informationen mit zahlreichen methodischen Herausforderungen verbunden. In der vorliegenden Arbeit wurden in vier Anwendungsstudien funktionelle Gehirnnetzwerke mit unterschiedlichen klinischem Hintergrund auf verschiedene Weisen, mit Kombinationen von etablierten Techniken und eigenen methodischen Entwicklungen, untersucht.

- Für die erste dieser Anwendungsstudien, welche in Kapitel 4 vorgestellt wird, werden die funktionellen Gehirnnetzwerke (EEG, schmerzhafte Stimuli, binäre Netzwerke) von depressiven Personen und gesunden Probanden mittels eigener Methodenentwicklung in kleine funktionell relevante Teilnetzwerke (motifs) zerlegt, welche gruppen-spezifische, innerhalb des EEG-Elektrodenschemas genau lokalisierbare Interaktionsmuster darstellen. Mittels dieser speziellen motifs konnten Gruppenunterschiede bei der neuronalen Prozessierung demonstriert werden.

- Kapitel 5 beschäftigt sich mit der Frage, ob Lithium-Behandlungseffekte in den funktionellen Gehirnnetzwerken (fMRI, Gedächtnisaufgabe, gewichtete kantenvollständige

Netzwerke) von HIV-positiven Probanden mit kognitiver Beeinträchtigung reflektiert sind. Hierzu wurden die Netzwerke hinsichtlich mikroskopischer und makroskopischer Eigenschaften untersucht und Behandlungseffekte nachgewiesen.

- In Kapitel 6 wurden räumlich sehr hochaufgelöste, aus tausenden von Knoten bestehende funktionelle Gehirnnetzwerke HIV-positiver Probanden (resting state fMRI, binäre Netzwerke) hinsichtlich einer funktonalen Segmentierung mittels Identifizierung von module-Struktur (communities) untersucht. In der gleichen Arbeit wurden weiterhin mittels eigener Methodik für die Generierung von Ground Truth Netzwerken mit bekannter module-Struktur umfangreiche Simulationsstudien mit dem Ziel betrieben, Güte und Erhaltung der module-Struktur zu quantifizieren, um Effekte einer neuen Methode zur Konnektivitätsbestimmung (lsGCI) zu evaluieren. Es konnte gezeigt werden, dass die lsGCI-Methode geeignet ist, um räumlich hochaufgelöste Netzwerke zu berechnen, deren funktionelle Segmentierung mit anatomischen Strukturen übereinstimmt.

- Die letzte in dieser Dissertation vorgestellte Anwendungsstudie (Kapitel 7) verfolgt die zeitliche Veränderung der module-Struktur und ihre Stimulus-induzierten Veränderungen (EEG, Balance-Perturbation, gewichtete Netzwerke). Dabei kommt für die Filterung von Kanten parallel zu einem neuartigen Ansatz zur Bestimmung multipler Schwellenwerte ein eigenes leistungsfähiges, auf multikriterielle Optimierung beruhendes Verfahren für die Bestimmung von Schwellenwerten zum Einsatz. Das für die Studie entwickelte aufwändige Analyseverfahren beinhaltet neben einer Selektion von geeigneten Zeitpunkten eine Aggregation von Interaktionsstrukturen zu verschiedenen Zeitpunkten zu Konsensusdaten auf Kantenebene und auf module-Struktur-Ebene, sowie eine state-of-the-art Visualisierung der identifizierten zeitlichen Veränderungen. Die identifizierte zeitvariante module-Struktur entspricht weitestgehend der erwarteten neuronalen Verarbeitung bei der Kompensation der Balance-Perturbation.

Die methodischen Herausforderungen dieser unterschiedlichen topologischen Netzwerkanalysen, aber auch die vielfältigen Möglichkeiten des Gewinns eines tieferen Verständnis der neuronalen Informationsverarbeitung und somit der Funktionsweise des Gehirns wurden anhand der erzielten Resultate aufgezeigt.

Abstract

Current research in computational neuroscience puts great emphasis on the computation and analysis of the functional connectivity of the brain given by temporal statistical dependencies (directed interactions) between the neural activities recorded at different sites, e. g. at EEG electrodes or at fMRI voxels, using either electrophysiological or hemodynamic measuring techniques. The methodological developments presented in this work are concerned with a comprehensive analysis of networks that represent functional interaction patterns, so-called functional brain networks. In general, functional brain networks exhibit intricate interaction patterns that cannot be directly comprehended or interpreted. To gain a deeper understanding of these networks and the underlying neurophysiological processes they reflect, analyses with methods from network science are indispensable. Thereby, the presented methodological developments were substantially influenced by the data and the investigated clinical research problems. A resultant distinctive feature of this work on functional network analysis is the exploration of network samples, which represent the functional connectivity of different groups of subjects. Using group-specific differential network analysis meaningful characteristics of connectivity patterns can be extracted, quantified and ultimately be interpreted with respect to the underlying neural information processing among brain areas. As described in the thesis, such analyses and the extraction of physiologically relevant information is connected with various methodological challenges.

In this work four application studies are presented in which functional brain network samples of different clinical background were analyzed in different ways using combinations of established approaches and own methodological developments.

- The first of these four application studies is concerned with a sample-specific decomposition of the functional brain networks (EEG, painful stimuli, binary networks) of depressed subjects and healthy control subjects into small functionally important and recurring subnetworks (motifs) using own developments (chapter 4). These motifs represent group-specific characteristic interaction patterns that can be located exactly within the EEG electrode layout. By means of these special motifs differences between the considered groups with respect to neural information processing could be shown.

- In chapter 5 it is investigated whether lithium treatment effects are reflected in the functional brain networks (fMRI, memory task, weighted edge-complete networks) of

HIV-positive subjects with diagnosed cognitive impairment. For this, microscopic and macroscopic structural properties were analyzed and treatment effects were shown.

- For the study presented in chapter 6 spatially highly resolved functional brain networks (resting state fMRI, binary networks), which consisted of thousands of nodes and were obtained from brain scans of HIV-positive subjects were explored with regard to a functional segmentation, as given by identified module (community) structure. In the same application study, ground truth networks with known module structure were generated using using own methodological developments. These ground truth networks formed the basis of a comprehensive simulation study that quantified module structure quality and preservation in order to evaluate the effects of a novel approach for the identification of connectivity (lsGCI). Thereby, it could be shown that using the lsGCI approach spatially highly resolved networks can be computed whose functional segmentation conforms with anatomical structures.

- The last application study presented in this thesis (chapter 7) tracks the time-evolution of module structure and its stimulus-induced changes (EEG, balance perturbation, weighted networks). For the filtering of edges in this analysis task, a newly developed and powerful own approach for the determination of edge weight thresholds is applied. It is based on multicriteria optimization and complements another modified approach that was also applied for the determination of multiple thresholds. The analysis concept that was developed for this study is complex. Apart from a selection of suitable time steps it entails an aggregation of interaction patterns at different time steps to generate consensus data at the level of edges and at the level of module structure, as well as state-of-the-art visualization of identified changes. The time-variant module structure identified in this study matches mostly the expected neural processing during the compensation of the balance perturbation.

The methodological challenges that are present in these different topological analyses, but also the various opportunities to gain an improved understanding of both, neural information processing and ultimately the functioning of the brain were highlighted with the achieved results.

Danksagung

Eine Dissertation wäre beinahe nicht vollständig ohne sie. Ich möchte an dieser Stelle meine tiefe Dankbarkeit und meine große empfundene Wertschätzung für die ideelle und tatkräftige Unterstützung denjenigen Personen gegenüber zum Ausdruck bringen, die mich innerhalb meines beruflichen und privaten Netzwerks begleiten. Dankbar bin ich Prof. Herbert Witte für seine vielen unermüdlichen Anregungen hinsichtlich der Arbeit am Institut im Allgemeinen und dem Verfassen meiner Dissertation im Besonderen. Ohne ihn hätte es diese Dissertation sehr wahrscheinlich nicht gegeben. Ich möchte mich bei Lutz Leistritz für seine vielen wertvollen Ratschläge bei der jahrelangen gemeinsamen Forschungsarbeit bedanken. Außerdem danke ich Prof. Stefan Schuster, meinen Bioinformatik-Professor während des Studiums, dass er sich bereiterklärt hat, meine Dissertation zu betreuen und zu begutachten. Bei Prof. Jens Haueisen bedanke ich mich für die Bereitschaft, ein Gutachten der Dissertation zu erstellen. Selbstverständlich bin ich ganz besonders meinen Eltern, meiner Familie und allen anderen Menschen, die mir nahestehen, dankbar. Dieser Dank ist herzlich und lässt sich hier nicht in Worte fassen.

Contents

List of Abbreviations

BD   binary directed network

BU   binary undirected network

CCN   correctly classified nodes

DCM   dynamic causal modeling

dDTF   direct directed transfer function

EEG   electroencephalography

fMRI   functional magnetic resonance imaging

GCI   Granger Causality Index

gPDC   generalized partial directed coherence

HIV   human immunodeficiency virus

LFR   module structure benchmark graphs of Lancichinetti, Fortunato and Radicchi

lsGCI   large scale Granger Causality Index

MD   major depression

MVAR   linear multivariate autoregressive model

WD   weighted directed network

WU   weighted undirected network

Part I

Introduction and prior art

Chapter 1

Introduction

O<small>NE</small> important focus of current research in computational neuroscience (CN) lies on the analysis of connectivity structures of the brain. This reflects the fact that neural information processing is based on two complementary operating principles: functional segregation and integration [1, 2, 3]. Functional segregation means that distinct areas of the brain are involved in or associated with the execution of specific neural functions and tasks. The precise communication between a large number of remote brain areas, i.e. functional integration, is indispensable to guarantee higher order sensory and coherent cognitive functions. During neural processing both operating principles of the brain, segregation and integration, have to be balanced and reconciled, which already suggest that brain functions must be related to complex, networked and dynamic brain structures, which in their entirety are known as the 'connectome'[1]. Accordingly, "the human connectome is a comprehensive map of the brain's circuitry, which consists of brain areas, their *structural connections* and their *functional interactions*" [4] or in other words "the brain is a large-scale network, operating at multiple levels of information processing ranging from neurons, to local circuits, to systems of brain areas" [5]. The notion of the brain as a network [6] implies that, informally, a network can be seen as an abstract object that represents the structure of the relationships (edges) between a set of entities (nodes). The methodological requirements of computational neuroscience are broad and extremely diverse, and are connected with the analysis of large datasets, data sharing and multimodal data integration. Thereby, the availability of advanced measurement technology is the fundamental prerequisite to obtain appropriate data.

---

[1]Various international projects were initiated in order to address the associated experimental and methodological challenges, e.g. the 'Human Connectome Project' (HCP) and the 'Human Brain Project' (CONNECT).

The term 'structural connections' (structural connectivity [2, 3]) is primarily associated with the identification of long range fiber tracts which can be performed on large-scale level by using diffusion-weighted magnetic resonance imaging (DW-MRI), where diffusion tensor imaging (DTI) is the most frequently used version of DW-MRI. Functional interactions (functional connectivity [7, 8, 1]) between brain areas cannot be directly measured, i.e. their study requires sophisticated analysis tools and modeling that uses measured brain activity as input data to estimate the strength of correlations[2]. Large-scale brain activity can be measured by electrophysiological (e.g. electroencephalography (EEG) and magnetoencephalography (MEG)) and hemodynamic/metabolic techniques (e.g. functional magnetic resonance imaging (fMRI), near-infrared spectroscopy (NIRS), and positron emission tomography (PET)). These two groups of functional neuroimaging techniques are characterized by different spatial and time resolution properties. For example, EEG has a high temporal but a low spatial resolution and fMRI has a low temporal but a high spatial resolution. Both properties can be combined by using simultaneous fMRI/EEG recordings which requires considerable measurement and computational effort, e.g. for artifact rejection.

Computational neuroscience might one day be able to precisely identify the basic principles that correlate the structural connectivity architecture of the brain with its activity and the consequent functional connectivity pattern of brain areas. Most likely, new ways of integrating functional neuroimaging results, including analysis and modeling results, with neuroanatomical data will have the potential to guide this development. The precondition for it is an integrated process of computational analysis and modeling, systems analysis, technology development and appropriate experiments.

Using network science methods, the description of a system and its complexity can be largely reduced to an account of the interactions between system elements and the emergent and inherent properties of the connectivity structure [12]. Thereby, network approaches recognize the important role of connectivity patterns for causing functional differences between network elements, i.e. nodes or subnetworks [3]. This role of connectivity patterns in conjunction with the notion of functional segregation introduces one of the main topics of the thesis, namely the identification of mesoscopic scale network structures, e.g. network modules, which are related to separated cohesive groups of strongly interacting nodes with a unique functional meaning. Modules are just one example of interesting information that can be extracted with

---

[2]This implies the possibility of temporal correlations of the activity of anatomically unconnected regions [2], though these regions might be connected by indirect structural paths [3]. Increasing evidence shows that functional connectivity between all corresponding pairs of recorded time series [9] reflects structural connectivity [10, 11, 3].

network science approaches. Consequently, over the last decade or so, network analysis and the application of network science methods has spread across many fields of scientific research, in which an increased focus is directed towards gaining an understanding of the organization, structure, behavior and evolution of complex systems. This is particularly true for research on functional brain connectivity [13]. Networks as models of complex systems and network analysis methods can be applied naturally to characterize functional connectivity data [14] and to describe the dynamic structure of functional interactions between brain areas [15] on both local and large spatial scales.

For this thesis only functional connectivity estimations obtained from either EEG or fMRI recordings of brain activity are considered. Thereby, the information about recorded complex brain activity is contained, or rather hidden, in the functional brain networks (section 1.2) of this thesis, which have dense and intricate patterns of directed interactions that cannot be comprehended by visual inspection. The problem of revealing and understanding this information by an appropriate characterization of the connectivity structure of these networks can and should be adressed by multi-level network analysis approaches.

For the presentation of the work of this thesis there is a lot of ground to cover. I will begin with a brief introduction of EEG and fMRI as methods for recording brain activity, followed by an outline of the functional connectivity estimation methods that were utilized to compute the networks whose algorithmic analysis is the foundation of this thesis.

1.1   Recording brain activity

Multichannel EEG [16, 17] signals are point-wise measurements of the massively synchronous dendritic electrical activity of large neuron populations (pyramidal cells) in the brain using high conductance electrodes typically placed at different standardized sites on the scalp using i. e. the 10-20 system [18]. The EEG has a high time resolution, but the choice of the measurement reference, volume conduction effects and artifacts from non-brain sources (e. g. eye movements, electric heart activity and muscle activity) influence the sensor space signals derived from the electrodes, which were exclusively used for the functional connectivity estimations of this thesis. However, signals can also be algorithmically transformed back into the source space by solving the inverse problem of inferring the unknown distribution and behavior of electrical current sources from EEG signals that best explain the observed external electrical field behavior. More in-depth information on this line of research can be found in e. g. [9, 19, 20].

The signals derived from fMRI [21, 22, 11] registrations are indirect measurements of mass activity of neurons. "The ability to detect changes in brain activity has a biophysical basis in the magnetic properties of deoxyhemoglobin, and a physiological basis in the way blood flow increases more than oxygen metabolism when local neural activity increases. These effects translate to a subtle increase in the local magnetic resonance signal, the blood oxygenation level dependent (BOLD) effect, when neural activity increases." [23]. The advantage of the registration of BOLD signals is that they can be derived from the location of neuronal activity with high spatial resolution.

1.2   Directed functional connectivity

Functional brain networks represent functional interactions, i. e. statistical dependencies, between time series of brain activity recorded at different sites. Nodes of functional brain networks are given by the recording sites (specific EEG electrodes or voxels) and their weighted edges are given by the connectivity estimate (interaction) for corresponding pairs of recorded time series of brain activity. Analysis of functional brain network has clinical relevance as descriptive studies of functional brain networks have offered new insights into basic principles of brain organization and function [3] and have enabled the assessment of specific abnormalities even in structurally normal regions [24].

There are currently two frequently used methodological classes for the computation of directed functional[3] neural interactions [25, 9, 26]: (1) Granger causality (GC) concepts (2) dynamic causal modeling (DCM). The methodological concepts and network analysis techniques presented in this thesis were applied to analyze functional network data obtained exclusively by the first class of connectivity measures. Within this class various approaches are subsumed by the notion of Granger causality. In dependence on the posed research question and signal properties of time series data at hand, different directed connectivity measures were used for the computation of functional brain networks: Granger Causality Index (GCI), large scale Granger Causality Index (lsGCI), direct directed transfer function (dDTF) and generalized partial directed coherence (gPDC). Such linear connectivity measures are normally based on multivariate autoregressive (MVAR) modeling of time series. Functional connectivity estimation is not within the scope of this thesis, as the functional network data was provided by other members of my research group. For the sake of completeness of this thesis, the used connectivity measures are concisely described in appendix B.1.

1.3   Aims of the thesis

The aims of my doctoral thesis cover two main areas: (1) the development of network analysis approaches and (2) their application to functional connectivity networks (left and center column of the table in figure 1.1). Even though the methodological solutions have been tailored for specific applications in functional network analysis they can be generalized for or readily transfered to other areas of application, i.e. analysis of networks from different domains. Their universality with respect to the considered complex system is one of the strengths of network modeling approaches. There are four methodological strands:

1. Testing and adapting of state-of-the-art network analysis methods (right column of the table in figure 1.1) and combining them to comprehensive analysis concepts for the envisaged applications.

2. Advancement and development of structural decomposition approaches in from of a novel approach for the detection of network sample-specific motifs with node labels.

3. Development of an algorithm for the parameterized generation of ground truth networks with known module structure for the simulation of MVAR time series.

---

[3]Friston distinguishes between functional and effective connectivity [1, 7].

4. Development of analysis concepts for the assessment of network module structure quality and preservation and for the extraction of dynamic module structure in time-evolving networks, including novel approaches for network thresholding and visualization of the module structure of network samples.

The developed concepts and methods contributed to the successful execution of the following clinically oriented methodological studies (center column of the table in figure 1.1), i. e. the corresponding analysis tasks were performed by me personally:

- One study was related to the investigation of samples of functional connectivity networks with respect to differences in topological characteristics. Thereby, the functional networks were computed from fMRI data. The study involved the pre-processing of network data and the quantification and interpretation of structural changes in HIV-infected patients in response to medical treatment (methodological strand 1).

- Analysis of functional connectivity during pain processing in patients with Major Depression and healthy subjects (methodological strand 2).

- Two additional studies aimed at evaluating the detrimental effects of a new approach for large scale functional connectivity estimation (lsGCI) on network edge patterns and the preservation of underlying module structure. This evaluation involved ground truth network simulation, the identification of network module structure and the analysis of the module structure quality in lsGCI functional brain networks. Again, HIV-infected subjects were enrolled in both studies, which were based on fMRI recordings of brain activity (methodological strands 1, 3 and 4).

- A not yet completed study aims at the analysis and tracking of dynamic module structure in time-evolving networks in response to external stimuli. EEG data forms the basis for this study and its preliminary results (methodological strands 1 and 3).

1.4 Structure of the thesis and integration of own contributions to network analysis

The following structure for the representation of the methodological approaches and application studies was chosen. In part I, chapter 2 ('Fundamental concepts in complex network analysis') the network theory and the methodological concepts required for the work presented in this thesis are briefly outlined. All reviewed methods for manipulating network

data and algorithms for assessing fundamental network characteristics are implemented in a network analysis toolbox. Only those approaches and network characteristics are described in more detail that are relevant for the application studies, whereas all other implemented approaches and topological characteristics are only briefly explained with references to the most relevant literature. Fundamental methods for network analysis were utilized, combined and complemented by new approaches that are described in part II, chapter 3 ('Contributions to network science methods and the multi-level analysis of functional brain networks'). The developed advanced analysis approaches were tested on simulated networks and applied in different application studies to real-world network data, which is described in the chapters of part III ('Applications and results'). Finally, some concluding remarks and perspectives are given in part IV ('Concluding remarks and perspective') of the thesis.

Depending on the data at hand and the formulated research question, each application study follows a specific methodological (and experimental) design, thereby utilizing different compilations of established network analysis approaches and my own methodological developments. To give a clear overview about how my own methodological developments contribute to the analysis of functional connectivity data from a network perspective and to highlight the central themes of my doctoral thesis an illustrative mapping between my own contributions (left column) and the application studies (center column) is depicted in figure 1.1. Also, the mapping between fundamental concepts (on the right column) and the application studies, for which they were utilized, is shown, enabling a clear distinction between my own work and the established approaches (left vs. right column).
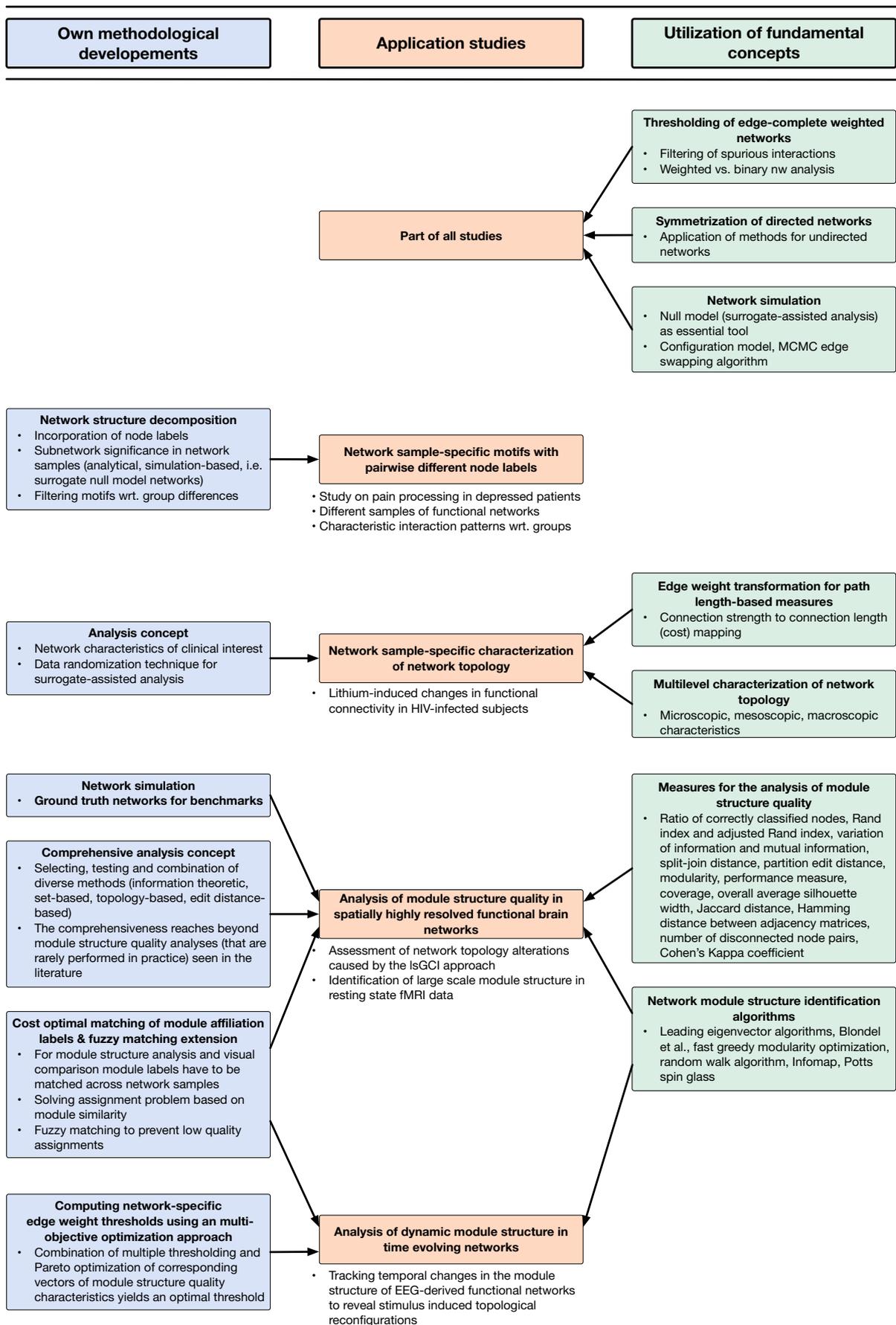
Own methodological developements	Application studies	Utilization of fundamental concepts

Thresholding of edge-complete weighted networks
- Filtering of spurious interactions
- Weighted vs. binary nw analysis

Part of all studies

Symmetrization of directed networks
- Application of methods for undirected networks

Network simulation
- Null model (surrogate-assisted analysis) as essential tool
- Configuration model, MCMC edge swapping algorithm

Network structure decomposition
- Incorporation of node labels
- Subnetwork significance in network samples (analytical, simulation-based, i.e. surrogate null model networks)
- Filtering motifs wrt. group differences

Network sample-specific motifs with pairwise different node labels

- Study on pain processing in depressed patients
- Different samples of functional networks
- Characteristic interaction patterns wrt. groups

Analysis concept
- Network characteristics of clinical interest
- Data randomization technique for surrogate-assisted analysis

Network sample-specific characterization of network topology

- Lithium-induced changes in functional connectivity in HIV-infected subjects

Edge weight transformation for path length-based measures
- Connection strength to connection length (cost) mapping

Multilevel characterization of network topology
- Microscopic, mesoscopic, macroscopic characteristics

Network simulation
- Ground truth networks for benchmarks

Comprehensive analysis concept
- Selecting, testing and combination of diverse methods (information theoretic, set-based, topology-based, edit distance-based)
- The comprehensiveness reaches beyond module structure quality analyses (that are rarely performed in practice) seen in the literature

Analysis of module structure quality in spatially highly resolved functional brain networks
- Assessment of network topology alterations caused by the lsGCI approach
- Identification of large scale module structure in resting state fMRI data

Measures for the analysis of module structure quality
- Ratio of correctly classified nodes, Rand index and adjusted Rand index, variation of information and mutual information, split-join distance, partition edit distance, modularity, performance measure, coverage, overall average silhouette width, Jaccard distance, Hamming distance between adjacency matrices, number of disconnected node pairs, Cohen's Kappa coefficient

Cost optimal matching of module affiliation labels & fuzzy matching extension
- For module structure analysis and visual comparison module labels have to be matched across network samples
- Solving assignment problem based on module similarity
- Fuzzy matching to prevent low quality assignments

Network module structure identification algorithms
- Leading eigenvector algorithms, Blondel et al., fast greedy modularity optimization, random walk algorithm, Infomap, Potts spin glass

Computing network-specific edge weight thresholds using an multi-objective optimization approach
- Combination of multiple thresholding and Pareto optimization of corresponding vectors of module structure quality characteristics yields an optimal threshold

Analysis of dynamic module structure in time evolving networks
- Tracking temporal changes in the module structure of EEG-derived functional networks to reveal stimulus induced topological reconfigurations

Figure 1.1 – Overview of relations between my own methodological developments, the applications studies and fundamental network analysis concepts.

Chapter 2

Fundamental concepts in complex network analysis

NETWORK science and network analysis received a lot of attention since the seminal publications of Watts and Strogatz [27], and Barabási and Albert [28]. The resulting wealth of new developments in this field makes it difficult to give a complete overview of the topic and the large amount of techniques for dealing with different aspects of network analysis also makes choosing a proper analysis strategy for given network data a difficult endeavor. This chapter gives a concise summary of some network theory and those state-of-the-art preprocessing and analysis techniques that are relevant for my work on functional brain network analysis. The methods that I used and that are described in this chapter cover a wide spectrum of network types and their topological features and can be used to characterize functional brain networks across individuals, groups of individuals, developmental stages and disease states. Together with my own methodological developments (see also chapter 3) the state-of-the-art methods and general concepts of this chapter can be beneficially combined to yield a powerful network analysis pipeline to extract essential information with clinical relevance from functional brain network data starting with the weighted edge-complete networks that are the result of the functional connectivity analysis (see also appendix B.1). I start this theory chapter with an outline of the relevant notation and definitions.

2.1 Notation and definitions

In the following, I present the formal notation and precise meaning of relevant network specific terms and language used throughout this thesis. These preliminaries will be useful for the

concise explanation of network science concepts underlying my work, which will be given in this chapter. Because the functional brain networks analyzed in this thesis are edge-directed, the focus of the following definitions lies on directed graphs, but I remind the reader that many of them apply naturally to undirected graphs, too. The chosen notation leans onto the notations used and presented in the textbooks [29, 30, 31, 32, 33].

The directed, unipartite functional connectivity networks I investigated within the scope of this thesis exhibit unilateral edges that encode asymmetric relations as they were based on functional connectivtiy measures that take the direction of interactions into account, i. e. PDC, gPDC, GCI and lsGCI. These measures are described in section B.1. As represented by the mathematical concept of a **directed graph** or **digraph** $\mathcal{D}$, a functional connectivity network consist of a non-empty finite set $\mathcal{V}$ of $N = |\mathcal{V}|$ **vertices** v_k, $k = 1, ..., N$ and a finite set $\mathcal{E}$ of ordered pairs of distinct vertices called **arcs** or oriented **edges** e. In the language of functional brain network analysis, edges can also be called **interactions**. In social network concepts, which often yield intuitive and useful analogies for the function of networks from different domains, edges are called **ties**. In the context of the analysis of networks, specific diagrammatic graph representations of systems, it is custom to use the term **node** instead of vertex. Functional connectivity networks are **unipartite**, which means they have only one type of node. A **bipartite** or **two-mode** network by contrast has nodes that are divided into two nonoverlapping sets and there are no interactions between nodes of the same set. The nodes of functional brain networks are usually uniquely **labeled** with the identifier of the corresponding recording site of brain activity, e. g. the EEG electrode or the voxel location. An ordered pair (v_i, v_j) is called an oriented edge that is originating from (outgoing from) node v_i and pointing towards (terminating at) node v_j. An **oriented edge** is denoted by $v_i \rightarrow v_j$ where v_i is called the **tail** (initial node) and v_j is called the **head** (terminal node) of the edge. An edge is **incident** to its tail and head nodes and conversely, a node is incident to the edges connected to it. Node v_i and v_j are called **adjacent** if they are directly connected by an edge. Contrary to a digraph, in an **undirected graph** $\mathcal{G}$ edges are unordered pairs v_i, v_j of nodes that have no particular relation structure and direction of interaction, which results in a lack of orientation for the edges in the graph. Let $|\Gamma_i|$ denote the **neighborhood of a node** v_i, which consists of all nodes that are adjacent to v_i. In directed networks the **out-** and **in-neighborhood** can be discerned with respect to the direction of the edges that connect adjacent nodes to v_i. A **binary** or **unweighted** (functional connectivity) network is

represented by its **adjacency matrix** $\mathcal{A}$ of size $N \times N$, which is a special case of a similarity matrix and where $\mathcal{A}_{ij} = 1$ if and only if the functional connectivity network contains the directed edge $v_i \rightarrow v_j$. Accordingly, a reciprocated (mutual) edge is indicated by two entries in the adjacency matrix $\mathcal{A}_{ij} = 1$ and $\mathcal{A}_{ji} = 1$, and is denoted by $v_i \leftrightarrow v_j$. The absence of an edge is indicated by $\mathcal{A}_{ij} = 0$. The adjacency matrix of a functional connectivity network and other digraphs is an **asymmetric matrix**. An undirected network has a **symmetric** adjacency matrix since $\mathcal{A}_{ij} = \mathcal{A}_{ji}$. In situations where both types of adjacency matrices should be discerned, $\mathcal{A}_{\mathcal{D}}$ is the notation used for an asymmetric adjacency matrix, whereas $\mathcal{A}_{\mathcal{G}}$ denotes a symmetric one. In case of a **weighted** (functional connectivity) network the entries in $\mathcal{A}$ are real-valued, usually positive numbers indicating the **strength of interaction** between a pair of nodes. Measures for weighted networks assume that edge weights are based on a **ratio scale**. To better discern the representation of a weighted network, also called a **valued graph**, from an unweighted network, the adjacency matrix $\mathcal{A}$ is often formally and conveniently replaced by the **edge weight matrix** $\mathcal{W}$ with real-valued entries $\mathcal{W}_{ij}$. $\mathcal{W}_{\mathcal{D}}$ and $\mathcal{W}_{\mathcal{G}}$ are the notations used for an asymmetric edge weight matrix and a symmetric edge weight matrix, respectively. Functional brain networks do not contain **hyperedges**, a generalized kind of edge that connects more than two nodes (that have a common group membership or classification). Binary functional connectivity networks do also not contain multiple **parallel edges** $(v_i, v_j)_1, (v_i, v_j)_2, ..., (v_i, v_j)_m, m \geq 2$ in the same direction. Directed networks can be transformed to undirected networks by means of performing a **symmetrization of the adjacency matrix** or edge weight matrix (see also section 2.2). Whereas in this thesis the **size of a network** denotes the number of its nodes, often the term **order**[1] is used for it instead, with the term **size** denoting the number of edges.

Functional brain networks are built by abstracting recording sites, single EEG-electrodes or single fMRI voxels, as **labeled nodes** and modeling associated directed interactions between these recording sites by oriented edges. It is crucial that, due to the pairwise different node labeling, all nodes are different – each node has a unique location in the network and plays a unique functional role as it represents (the activity of) a unique area of the brain. A **subdigraph** or **directed subnetwork**, respectively, of a directed network $\mathcal{D}(\mathcal{V}, \mathcal{E})$ is a directed network $\mathcal{D}_s(\mathcal{V}_s, \mathcal{E}_s)$ consisting of subsets $\mathcal{V}_s \subset \mathcal{V}$ and $\mathcal{E}_s \subset \mathcal{E}$ of the sets of nodes and edges of $\mathcal{D}$. This means that a subnetwork is a part of a network. If a subdigraph $\mathcal{D}_s$ contains all edges of its **superdigraph** $\mathcal{D}$ between its nodes, $\mathcal{D}_s$ is called an **induced subdigraph** of

---

[1]In the wider context of my work the term 'order' is already reserved for the time series analysis of brain signals, which is the foundation for obtaining functional brain networks.

$\mathcal{D}$. More precisely, if every edge $e \in \mathcal{E}$ with both end-nodes in $\mathcal{V}_s$ is element of $\mathcal{E}_s$, we say that $\mathcal{D}_s$ is induced by $\mathcal{V}$. For undirected networks the analogous terms **subgraph** or **subnetwork** are used. An **edge-complete network**, or simply, a **complete network**, is a network that possesses all possible directed or undirected edges between its nodes, i.e. for every pair v_i, v_j of distinct nodes, both oriented edges (v_i, v_j) and (v_j, v_i) exist. Two digraphs $\mathcal{D}_1$ and $\mathcal{D}_2$ are **isomorphic** if and only if for any ordering of the nodes of $\mathcal{D}_1$ there is an permutation ϵ of the nodes of $\mathcal{D}_2$, such that their adjacency matrices are equal, i.e. $\{v_k, v_m\}$ is an edge of $\mathcal{D}_1$ if and only if $\{\epsilon(v_k), \epsilon(v_m)\}$ is an edge of $\mathcal{D}_2$.

A **walk** connecting the initial node v_1 to the terminal node v_k is an alternating sequence of nodes and edges, such that each edge connects its preceding (tail) with its succeeding (head) node: $v_1, e_1, v_2, e_2, ..., e_{k-1}, v_k$. If $v_1 = v_k$, then the walk is called a closed walk. In a directed network, a walk in which all edges point in the same direction is also called a **directed-edge sequence** that can be closed or open, too. A **trail**[2] is a walk in which all edges are distinct (but not necessarily all nodes), whereas a **path** is a walk in which all nodes are distinct (and therefore all edges, too). If a path has identical start and end nodes, $v_1 = v_k$, then it is called a **cycle**. A **self-loop** is a cycle of length one, i.e. an edge whose tail and head coincide. Functional connectivity networks do not contain self-loops, i.e. $\mathcal{A}_{ii} = 0$, $\forall i \in \mathcal{V}$. The **path length** is the sum of the weights of all edges that lie on the path, where in the special case of binary networks all edge have a weight of one. The **distance**, also called **graph distance** or **geodesic distance**, $\delta_{sp}(v_i, v_j)$ between two nodes $v_i \in \mathcal{V}$ and $v_j \in \mathcal{V}$ of a finite graph is the minimum length of any path connecting them, i.e. the length of the **shortest path** connecting them. The shortest path[3] between two nodes is called their **graph geodesic** and it can be computed [34, 35] using either a breadth-first traversal algorithm (unweighted network), Dijkstra's algorithm (weighted network, non-negative edge weights), or using the Bellman-Ford algorithm or Johnson's algorithm for networks with arbitrary edge weights (but not containing negative cycles). The **number of shortest path** between the nodes $v_i \in \mathcal{V}$ and $v_j \in \mathcal{V}$ is denoted by σ_{ij}. By convention, $\sigma_{ii} = 1$. The number of shortest paths between $v_i \in \mathcal{V}$ and $v_j \in \mathcal{V}$ on which some node $v_k \in \mathcal{V}$ lies is denoted by σ_{ij}^k. Pairs of nodes that are not connected by any paths, i.e. nodes that are members of different isolated parts of the network, have infinite distance. Distances between all pairs of nodes are stored in the **all-pairs shortest path matrix** $\Delta_{sp}(i, j) = \delta_{sp}(v_i, v_j)$, which is symmetric for undirected networks and asymmetric for directed networks. In a directed or undirected

---

[2]also called a directed-edge train [29]
[3]The path itself, not its length.

network a node v_j is **reachable** from another node v_i, if a directed path or, respectively, an undirected path from v_i to v_j exists that is connecting them. Pairs of nodes that are linked by such paths are said to be **connected**. A directed network is **strongly connected** if every ordered pair of nodes is strongly connected, i.e. mutually reachable, which means there exists a directed (v_i, v_j)- and (v_j, v_i)-path between all pairs of nodes v_i and v_j. If the orientation of edges in paths in a directed network is ignored by taking into account **strict semipaths** between all pairs of nodes or due to considering a symmetrization variant of the directed network where all unordered pairs of nodes are accordingly regarded as being connected if they are reachable by undirected paths, the directed network is said to be **weakly connected**. Similarly, an undirected network is **connected** if every node is reachable from any other node. In a **disconnected** network $\mathcal{D}^{\Vdash}$ or $\mathcal{G}^{\Vdash}$ the set of nodes $\mathcal{V}$ is partitioned naturally into subsets of nodes, which, together with their respective edges, form different connected subgraphs that are separated from each other as they are not connected by any (reciprocated) edge and that are called (strongly- or weakly-connected) **components**. Each node and edge belongs to exactly one component and nodes of different connected components are mutually unreachable. If a network is connected, it necessarily consists of only one component. After determining network components using breadth-first search algorithms, the rows and columns of an adjacency matrix of a disconnected network can be reordered so that the matrix takes **block diagonal form**, in which the components become visible as square blocks of non-zero elements along the main diagonal (similar to the adjacency matrix plots of networks with module structure in section 6.1). The associated network components consisting exclusively of strongly connected node pairs are called **strongly connected components**. The set of nodes reachable from a given parent node v_i via edge-directed paths, including v_i, is called **out-component**. All nodes external to the out-component interact with member nodes of the out-component solely via edges that point towards the out-component. The **in-component** is defined in an analog fashion as the set of nodes from which there is a directed path leading towards a given node v_i, which is included in this set. It follows that the strongly connected component of v_i equals the intersection of its in- and out-components.

The connectedness property is important and is controlled for in the functional brain network analyses presented in this thesis: To yield a more realistic representation of brain activity, the dichotomized functional networks under study are required to be at least weakly connected, which means that they are not allowed to fragment into **isolated nodes**, which are not incident with any edge, or separate **weakly connected components**, which are maximal

subgraphs in which all pairs of nodes are weakly connected. A **cut node** or **articulation point** is a node, whose removal, together with all its incident edges, increases the number of components of a network. The edge equivalent of a cut node is called a **bridge**.

2.2   Symmetrization of directed networks

Despite functional connectivity networks being directed it might be useful to symmetrize their adjacency matrix or their edge weight matrix to assess network characteristics of the underlying undirected network. This way of proceeding allows to incorporate additional measures of network topology that are only defined for undirected networks into the analysis. For example, in the context of module structure identification (see also chapter 6) I found that edge orientation had limited impact on the quality of the results. Thus, in favor of being able to make use of different module structure identification algorithms I also worked on symmetrized versions of the original networks, even though this might entail certain biases and information loss (see also section 2.8). Ultimately, it all comes down to whether the direction of relationships is important for the investigated research question. An interesting application for link reciprocity (see also section 2.7.3) is the assessment of the amount of interaction asymmetry for determining how close the structure of a binary directed network is to being essentially undirected. This information can be used to objectively justify the use of network symmetrization procedures for network analysis.

The asymmetric adjacency matrix of a binary directed network may be symmetrized by addition with its transpose $\mathcal{A}_{\mathcal{G}} = \mathcal{A}_{\mathcal{D}} + \mathcal{A}_{\mathcal{D}}^{T}$ and subsequently setting any resulting entries $\mathcal{A}_{\mathcal{G}_{ij}} = 2$ to $\mathcal{A}_{\mathcal{G}_{ij}} = 1$.

Several basic symmetrization techniques exist for weighted networks. These functions can be applied to each matrix element of $\mathcal{W}$ to generate the undirected network element-wise as

follows [36]

$$sym^{mean}(\mathcal{W})_{ij} = \frac{\mathcal{W}_{ij} + \mathcal{W}_{ji}}{2} \tag{2.1}$$

$$sym^{min}(\mathcal{W})_{ij} = \min(\mathcal{W}_{ij}, \mathcal{W}_{ji}) \tag{2.2}$$

$$sym^{max}(\mathcal{W})_{ij} = \max(\mathcal{W}_{ij}, \mathcal{W}_{ji}) \tag{2.3}$$

The entries of the symmetrized edge weight matrix $\mathcal{W}$ might subsequently be scaled to the interval $[0, 1]$ with the methods of section 2.4.

2.3 Thresholding edge-complete weighted networks

Generally, the quantification of functional connectivity in recordings of brain activity results in edge-complete, fully connected weighted networks (see also appendix B.1). While it is possible to analyze such edge-complete weighted networks directly (see also chapter 5 on page 96), it is common practice to further preprocess potentially noisy, edge-complete networks by a thresholding procedure to remove low-weight interactions that are potentially spurious, i. e. false positive. Working with the resulting non-complete networks circumvents certain problems that can occur in subsequent analyses, primarily the problem of ill-defined network measures and restricted combinatorial options for network randomization, which is essential for surrogate-assisted network analysis (2.12 on page 50). These problems are particularly pronounced in binary networks where all equally sized subsets of nodes exhibit the same topological interaction pattern. In edge-complete weighted functional connectivity networks problems of discerning substructures and topological characteristics arise since weights might be very homogeneously distributed across the edges. Filtering[4] and rejecting edges with sub-threshold weights is not a trivial procedure as it has the potential to introduce serious biases and confounders into the resulting network topology [37, 12, 15, 38]. Corrections for these effects do not exist, because the 'real' underlying network topology needs to be known [37], which is clearly not the case for empirical data of brain activity, as determining the topology and characteristics of functional brain networks is the very objective of the studies. As a consequence, it is difficult to discriminate between a real effect seen in the data and one

---

[4]In principle , network filtering can also be performed on the node level by removing (peripheral) nodes with certain properties.

that is caused by the effects of the thresholding procedure[5]. The severity of such biases is more pronounced for small networks as those obtained by EEG recordings or region-of-interest approaches [37]. Thresholding always discards information on interaction patterns which can obscure underlying structure. An example for this was given in [39], where it was shown for a weighted toy network that thresholding will never succeed in revealing its true underlying module structure in its entirety.

Until now there has been no generally accepted criterion for defining thresholds on edge weights. There are two principle approaches: statistical significance tests [40, 41, 37] and heuristic thresholding procedures [37, 24, 42]. Tests for statistical differences of edge weights are computationally expensive and are feasible in practice only for small networks and small network samples. They are burdened by the arbitrary definition of the type I error and the way alpha-adjustments for multiple comparisons are performed, if any. The reliance on p-values, a key element of such statistical tests, is disputed in general [43, 44, 45, 46, 47, 48, 49, 50]. In contrast, heuristic thresholds are chosen somewhat arbitrarily and usually yield a common cutoff level for all edges or are selected in an attempt to fix arbitrary values of basic network characteristics, such as node degrees or edge density [37]. Fixing arbitrary network characteristics, as e.g. in [51] has recently started to lose acceptance. Instead, often several instances of a network that are obtained by using ranges of thresholds are analyzed. Both principle thresholding approaches assume that only strong interactions contribute meaningfully to the organization of network structure. Contrary to this, the role of certain kinds of weak interactions was recognized as 'the strength of weak ties' [52, 53, 54] and the concept of bridges [55]. Under the assumption that disconnected functional brain networks yield a poor model of neural processing (cf. [56]), the thresholding in this thesis was constrained by the ensured connectedness of the resulting networks, i.e. they were not allowed to fragment into separate connected components or single isolated nodes. This is also circumvents certain problems that arise for disconnected networks down the network analysis pipeline, e.g. the assessment of some network characteristics that rely on path lengths and the malfunction of some module detection algorithms.

---

[5]The analysis can be improved and detrimental effects can be reduced by surrogate-assisted analysis and normalization of network characteristics using a baseline network model (random null model networks or down-sampled networks with randomly removed edges) as described in section 2.12.

2.3.1 Testing for statistical differences in edge weights

Edge weights $\mathcal{W}_{ij}$ in the functional brain networks of this thesis are quantified by a functional connectivity measure, i.e. GCI, lsGCI, dDTF or gPDC (see also sections B.1.1, B.1.2 and B.1.3). To remove subliminal and statistically non-significant interactions from the edge-complete weighted functional networks each edge weight is related to the distribution of the connectivity measure when no influence from time series Y_i to Y_j is present, which corresponds to the null hypothesis H_0 of the associated statistical test. Since the distribution of edge weights under H_0 is usually analytically unknown, a bootstrap approach [57] has to be used to construct it. Following, I give a brief summary of the procedure outlined in [58], which was used for the functional networks of this thesis whenever thresholding and dichotomizaton based on statistical tests was performed.

For the computation of the functional brain networks and their edge weights the parameters $\boldsymbol{A}^r$ of a MVAR model (equation B.1) are estimated based on the neural time series data at hand. To test for statistical differences of the resulting individual edge weights, the model residuals are calculated with respect to the original time series data and resampled by random sampling with replacement. To test the influence of the time series Y_i on the time series Y_j, the associated entries[6] in the autoregressive parameter matrix are set to $\boldsymbol{A}^r_{ji} = 0$ for all $r = 1, \cdots, p$, i.e. all coefficients related to time series Y_i causing Y_j are set to zero. All other autoregressive coefficients remain as they were originally estimated. The autoregressive model under H_0 and the resampled model residuals are used to generate bootstrap time series according to equation B.1. For these time series under H_0 a new MVAR model is fitted to compute a bootstrap replication $\mathcal{W}_{ij}^{H_0}$ of $\mathcal{W}_{ij}$ under H_0 that quantifies the relationship between Y_i and Y_j under H_0. The entire bootstrap process is repeated a large number of times for every pairwise interaction $\mathcal{W}_{ij}$ separately, which results in high computational costs. Therefore, the procedure is infeasible for large networks, e.g. the large functional networks of chapter 6. Finally, for each pairwise interaction the obtained sample of bootstrap replications $\mathcal{W}_{ij}^{H_0}$ is used to either calculate a critical value, given by the $(1 - \alpha)$ quantile, or a p-value for the hypothesis test. Because it can be argued that the functional interaction structure of a network has to be considered as a whole and single interactions should not be tested independently [41], an alpha-adjustment for multiple hypothesis tests is necessary, which can be performed by e.g. the Holm correction procedure [59] or by controlling the false discovery

---

[6]Note the different notation with respect to the order of the indices i and j.

rate [60]. The p-value for each pairwise interaction $\mathcal{W}_{ij}$ is then compared with its adjusted α value. Edges for which no statistical difference is found are removed from the resulting network.

To statistically threshold functional networks that were generated based on simulated ground truth networks (see also section 3.4), a similar strategy to estimate the distribution of connectivity measures under H_0 can be used [61]. This approach is based on a Monte Carlo simulation and works directly on the predefined autoregressive parameter matrix $\boldsymbol{A}^r$ (see also appendix B.2). As before, the coefficients for each tested interaction are set to $\boldsymbol{A}^r_{ji} = 0$ prior to the realization of a new MVAR process, which generates a set of time series under H_0 from which the connectivity measure between time series Y_i and Y_j is recalculated. This simulation is repeated to generate a large sample of values of the connectivity measure $\mathcal{W}^{H_0}_{ij}$ under H_0 for the statistical hypothesis test.

2.3.2 Multiple threshold strategy

To gain robustness against specific edge weight threshold choices, multiple thresholds should be considered for the analysis of networks and the identification of persistent topological features [42, 24]. To actually threshold an edge weight matrix, the edge weight corresponding to a given percentile of all edge weights can be used as threshold parameter τ [62]

$$\mathcal{A}_{ij} = \begin{cases} 1 & \text{if } w_{ij} \geq \tau \\ 0 & \text{if } w_{ij} < \tau \end{cases} \tag{2.4}$$

In practice, whenever statistical tests were not feasible from a time resources point of view the strategy that was pursued in this thesis consisted of using multiple pre-selected percentiles of the edge weights as thresholds. Using edge weight percentiles circumvents specifying certain edge weights directly, but results in a fixed number of retained edges independent of the underlying edge weight distribution. In practice, a fixed number of edges for all networks in a network sample might be undesirable, as it potentially masks differences in structure of some networks. If samples of related networks are analyzed, it might be worthwhile to pool the edge weights of all individual networks to select a weight as common threshold using a pre-defined percentile. This results in individual networks with different numbers of retained edges. Thresholding effects have to be analyzed in an exploratory fashion to justify the choice of selected edge weight thresholds.

In section 3.5 on page 74 I present a different method that obtains an 'optimal' edge threshold value for module structure analyses. In a sense, this method accounts for an on average minimum information loss with respect to the module structure of the resulting thresholded network and yields non-arbitrary, objective and network-specific thresholds. I applied this new approach for network data analysis in the study presented in chapter 7 on page 129.

2.4   Comparison of weighted vs. unweighted network analysis

Depending on research objectives, after removing low-weight edges one can either proceed with the analysis of the non-complete weighted network, which is a form of soft thresholding, or one can dichotomize all remaining weighted edges by assigning them a weight of 1, which yields non-complete binary networks and is called hard thresholding [62]. Analysis of weighted networks has the advantage that the information about the continuous nature of edge weights is preserved and can consequently be exploited to obtain a more detailed, fine-grained understanding of the network and the modeled relationship between its entities [63]. On the other hand the underlying binary network, with limited richness of interaction information, is often easier to interpret, to analyze and to visualize, using a wealth of available techniques and measures. In particular, normalization of network characteristics is usually less intricate in binary networks as compared to weighted ones. Dual approaches combine information from both network types to discover different forms of latent structure [39], e.g. to account for the situation that the same binary characteristic corresponds to different weighted counterparts [64, 65].

2.5   Rescaling of edge weights

Narrow ranges of similar edge weights are often found in weighted functional connectivity networks. For analyses it might be beneficial to resolve a narrow range of edge weights by mapping them into the entire $[0, 1]$ interval. Such a mapping can be obtained by the following function [36]

$$f(w_{ij}, \beta, \mathcal{W}) = \left(\frac{w_{ij} - \min(\mathcal{W})}{\max(\mathcal{W}) - \min(\mathcal{W})} \right)^{\beta} \tag{2.5}$$

where $\beta > 0$, e.g. $\beta = 1.5$ yields a good edge weight resolution for the EEG-derived networks analyzed in chapter 7. Other functions for the purpose of separating similar edge weights are given in [62]

$$f(w_{ij}, a, b) = \frac{1}{1 + e^{-a(w_{ij}-b)}} \tag{2.6}$$

$$f(w_{ij}, c) = |w_{ij}|^c \tag{2.7}$$

with positive parameters a, b and c. The choice of the parameters a, b and c determines the sensitivity of the mapping and therefore how good similar edge weights are resolved and how much the mapping increases the differences between their newly assigned values. Alternatively, edge weights might be scaled by dividing with the maximum value in the network.

2.6 Transforming edge weights for path length-based indices

In the context of network characteristics that are based on (shortest) path lengths between nodes the weight of an edge is interpreted as its connection length. Accordingly, a high edge weight is regarded as a large distance between the nodes connected by the edge. This implies that the edge has high traversal costs and corresponds to a weak connection between its nodes, which are too distant from each other to strongly interact. In structural networks, which describe anatomical connections, network methodologies based on path lengths can be applied directly and have a clear meaning in terms of neuronal signaling or communication [3]. However, in functional connectivity networks the situation is opposite, as high edge weights indicate a strong interaction between the connected nodes. Therefore, edge weights have to be transformed so that high strength interactions become short distances associated with low edge traversal costs [66]. For it, several transformations were proposed, e.g. subtracting from the maximum weight plus one or another upper bound, taking the inverse $1/w_{ij}$ or taking a negative exponential $e^{-w_{ij}}$ [63]. Equations 2.6 and 2.7 can also be used to obtain an connection strength to connection length (cost) mapping. For it, parameter a and c have to take on negative values, whereas b can be either positive or negative. A potentially good range for parameter a might be given by $-15 < a < -5$ and for parameter b it might be $0 < b < 0.6$. Setting parameter c to $-0.5 < c < -0.1$ might give reasonable results. The transformation most often used for the work presented in this thesis is given by

$$d(v_i, v_j) = \sqrt{2(1 - w_{ij})} \tag{2.8}$$

which defines a metric distance between any two nodes v_i and v_j [67].

2.7 Comprehensive quantification of network topology

In this section I give a concise but broad overview of network characteristics that taken together quantify many different structural aspects of a network and yield a good overall picture of its topology. Thereby, I discern three topological scales: the microscopic, mesoscopic and macroscopic scale. Although the described measures provide different perspectives on major features of network architecture, several of them are inevitably interrelated. Due to the methodological independence of these measures, which can be assessed individually and separately, the given overview rather takes on the form of a listing. These measures and many more are part of my network analysis toolbox. Currently, a manuscript is under preparation, which reports the results for a study in which large scale functional brain networks of HIV infected patients and healthy controls were analyzed using a multitude of macroscopic scale characteristics. Subsequently, the network samples were compared and the correlations among network characteristics and several clinical scores were investigated. Since this study is not described in this thesis and reviewing every single measure of network topology is beyond the scope of this thesis, I only state measures that I used for my work on analyzing functional brain networks. For an extensive overview of network characteristics I refer to review articles like e. g. [68, 31, 69, 70, 71, 72, 2] and textbooks like e. g. [73, 31, 36, 14].

Contrary to the descriptions typically found in the scientific literature on network science I emphasize the network types for which measures of network structure, network models or algorithms are defined, which I belief will add clarification and value to the explanations and the overview given in this thesis. Thereby, the notation is the following: "BD" and "BU" denote, respectively, *binary directed* and *binary undirected* networks. Consequently, the abbreviations "WD" and "WU" denote, respectively, *weighted directed* and *weighted undirected* networks. Due to the nature of functional brain networks, which have weighted or binary directed edges, I focus primarily on presenting WD and BD network characteristics and related algorithms, if available, despite stating all network types for which definitions and concepts apply. Otherwise I state the undirected versions, which can be applied to functional brain networks, too (see also section 2.2). As already described, weighted networks can be dichotomized to apply binary network measures (see also section 2.3 on page 16).

2.7.1 Microscopic scale network characteristics

Characteristics on the level of single nodes and edges are located on the microscopic scale of network structure. A special meaning have node centrality indices that rank nodes by assigning them a structural centrality value according to different notions of their role, position and prominence in the network. This concept originated in the work on social networks and communication [74] and in operations research [75]. Accordingly, nodes that have high values for many centrality measures are important functional elements of network structure [66]. They can be thought of as being strategically located at the center of star-like network configurations, where due to their high degree and short distances to all other nodes they constitute a major gateway for flow and communication along most of all the shortest paths in the network. Central nodes can be regarded as able to avoid the control potential of other nodes that act only as intermediaries of network flow [76]. On the contrary, nodes with low centrality scores are likely to be on the outer layers of their network where they cannot access many other nodes. Thus, centrality indices make for an important network analysis tool for the extraction of information from networks.

To counteract the dependence of a node centrality measure on network size, a linear rescaling of individual node centrality values to the $[0, 1]$ interval can be performed. Another usual way of normalizing a node centrality value is by dividing by the maximum possible score in a network of the same size N or by dividing by the sum of all scores if the distribution of values is important [63]. To control for the influence of purely mechanistic effects of network topology, centrality measures can be compared to the ones obtained from randomized network counterparts (see also section 2.12 on page 50).

Following, I list relevant microscopic scale network characteristics together with their main references.

The **node degree**$^{\text{BD,BU}}$ κ_i of a focal node v_i in a binary network equals the number of adjacent nodes [69, 76]. The **node strength**$^{\text{WD,WU}}$ ζ_i in a weighted network equals the sum of incident edge weights [69]. Therefore, node degree and strength measure a node's importance for the network structure in its proximity as given by its interaction activity (many interactions vs. strong ones). In a binary directed network, due to the added complexity of considering edge orientation there are four kinds of node degrees, the out-degree κ_i^{out}, the

in-degree κ_i^{in} [66], the total degree κ_i^{tot} and the reciprocal degree $\kappa_i^{\leftrightarrow}$ [77].

Closeness centrality[BD,BU,WD,WU] captures the centrality of a node with respect to its independence from intermediaries and with respect to how easily and efficiently it can access other nodes and their ressources in the network or route flow to them [66, 78, 76, 79, 80].

Betweenness centrality[BD,BU,WD,WU] measures the number of shortest paths that pass through a node. Thus, betweenness centrality captures the potential of a node to coordinate, withhold or distort flow along geodesics between other nodes and other parts of the network [74, 78, 81, 76, 63]. Different variants of betweenness centrality [63, 82, 66] were developed and algorithmic aspects were improved to compute betweenness in large and dynamic networks, e.g. [83, 84, 85, 86, 87]. Betweeness is also analogously defined for edges, which can be exploited for network community detection.

Nodes centrality can also be assessed by the concept of **key player** nodes [88] or by classification into **hub nodes** and non-hub nodes [89].

For any node v_i the local node-level **clustering coefficient**[BD,BU,WD,WU] is given by the ratio between the number of triangles with v_i as center node (and two of its adjacent nodes that are connected themselves) and the number of triangles that node v_i could have formed, given its number of neighboring nodes [90, 27, 91]. It is a measure of segregation. Nodes with low clustering coefficient have relatively many missing edges in their neighborhood, which implies the existence of so-called 'structural holes' (cf. [55]). This, in turn, gives the respective nodes some additional control over the information flow between its mutually unconnected neighbor nodes [31]. Influential nodes, e. g. hub nodes, have a low node-level clustering coefficient [91, 92]. Different definitions of the clustering coefficient in weighted networks can be found in [64, 93, 90, 94]. Clustering in network samples was analyzed in the study presented in chapter 5

The **eccentricity**[BD,BU,WD,WU] $\epsilon(v_i)$ of a node v_i in a connected network is the largest geodesic distance between v_i and any other node [75]. It can be thought of as how far a node is from the one most distant from it.

A node with an elevated **Page Rank**[BD,BU,WD,WU] value is structurally and functionally

important, acting as e.g. a switch and broker of network flow. High Page Rank values result from either many other nodes being connected to the node, or at least some nodes that themselves have a high PageRank [95, 96].

2.7.2 Mesoscopic scale network characteristics

Elements of network structure on this topological scale depend both on local and global aspects of network organization. Network motifs and in particular network modules are indicative of functional segregation in the brain [2]. Mesoscopic scale topological features might be used for network-based data reduction techniques that summarize network structure by coarse graining or focusing only on important substructures [39]. As a major part of my work focussed on network analysis at the mesoscopic level, I describe these topological feature and their detection in more detail.

Network motifs[BD,WD] are statistically overrepresented connection patterns given by small local connected and induced subnetworks [97]. Motifs are assumed to act as functional meaningful building blocks or as elementary information processing circuits [98] of a network. It was discovered that individual real-world networks (or classes of networks) possess characteristic combinations of network motifs that might reflect topological constraints related to the functionality of the represented system and its history of development [99]. Standard network motif detection in a single directed binary network with unlabeled nodes, which are indistinguishable from each other, is performed in three fundamental and computationally expensive subtasks [97]:

1. Solving the subgraph census problem [100] by exhaustively enumerating [101] or sampling [102, 103] the number of occurrences of each subnetwork induced by a set of k nodes in the input network. This quantity is affected by the kind of allowed node and edge overlap, e.g. non-identical counting (arbitrary overlaps) [104, 105]. Even in comparably small networks, the number of subnetwork occurrences is potentially large due to its exponential increase with the size of the input network. Moreover, the number of k-node subnetworks in a network grows very fast with k, which is commonly chosen to be 3 or 4.

2. The second subtask in network motif detection encompasses determining graph isomorphism (see also section 2.1) for found subnetworks to group them into topological equivalence classes and obtaining their counts. It is believed that graph isomorphism

cannot be solved in polynomial time. Several algorithms for solving graph isomorphism with miscellaneous performance in practice have been presented [106].

3. The last subtask is assessing statistical significance of subnetwork occurrences. Subnetworks with a statistical difference in their number of occurrence in the input network as compared to their number of occurrence in a large set of simulated null model random networks are deemed motifs (see also sections 2.11.2 and 2.12). Usually, a subnetwork occurrence is defined to be significant if it occurs a certain multiple of standard deviations more often in the input network than would be expected in the set of random networks, which is expressed by the Z-score [104, 102, 107]. Making use of Z-scores for assigning statistical significance to subnetwork occurrences is flawed by the unsafe assumption being made that subnetwork occurrences follow a normal distribution [104] and it was shown that this is not always the case [108]. Statistical significance of subnetwork occurrences might in general be assessed by computing p-values based on the distribution of subnetwork occurrences under the null model [101, 109], which corresponds to a non-parametric evaluation that avoids the problem of making assumptions about a particular type of distribution.

The centerpiece of the network decompostion into motifs is the third subtask, which is concerned with uncovering statistical differences in subnetwork occurrences. It is the statistical overrepresentation of subnetworks that indicates their putative role as functional meaningful building blocks or elementary local computational circuits of their network [101, 109, 110, 99, 111, 112]. Indeed, it was shown that motifs can act as elementary computational circuits for information processing in gene transcription networks [109, 110, 111]. A complicated interplay between motifs and global network structure seems to exist. On one side, the statistical overrepresentation of subnetworks might be a consequence of global topological constraints (architectural type and function of the network) [98, 113]. Conversely, global network structure might be shaped by the presence of motifs, which stabilize dynamics [98] and thus support robust stability of the network with regard to small-scale perturbations. A slightly modified variant of motif detection was used in the context of anatomic brain networks to investigate *structural motifs* and the instances of functional motifs contained within them [107]. A somewhat more specialized variant of motif detection considers *topological motif generalizations*, larger subnetworks with a common architectural subpattern as defined based on replications of nodes in basic subnetworks [114]. Another network fingerprinting method that analyzes smaller substructures of networks and that can be used for network similarity comparisons is based

on undirected and directed *graphlets* [115, 116, 100, 73], which are obtained by a network decomposition into subnetworks without assigning statistical significance to their counts, and the corresponding graphlet frequencies and graphlet degree distributions. Motif detection was generalized to weighted directed networks in [93].

Network modules$^{\text{BD,BU,WD,WU}}$, also known as communities, are disjoint cohesive groups of densely connected nodes, which probably share common properties and play similar roles within a network. The similarity of nodes is reflected in a higher probability of edges between comember nodes of the same module than between nodes in different modules [70, 117, 118]. In a network module in the strong sense, each node has more intra-module edges than inter-module edges [119]. Many algorithms for identifying module structure exist (see also section 2.8). Detecting modules in a network with an unknown number of present modules is inherently subjective, depends on vaguely defined constraints and requires domain knowledge for proper interpretation [120, 70]. Module structure was detected in the studies presented in chapters 6 and 7.

2.7.3 Macroscopic scale network characteristics

The global network characteristics on the macroscopic scale are usually highly non-trivial functions of the adjacency matrix. On the macroscopic scale we find network properties that summarize the network structure in its entirety. Such global network signatures provide a parsimonious representation of a network's functional interaction patterns. Macroscopic network characteristics based on paths between nodes are measures of functional integration as they indicate costs of routing of information flow within the network [2]. Small-world network topology, which supports wiring cost efficient complex information processing, is reflected by a combination of the transitivity measure and the characteristic path length [121]. Following, I list relevant macroscopic scale network characteristics.

The **degree distribution**$^{\text{BD,BU}}$ or, respectively, the node **strength distribution**$^{\text{WD,WU}}$ constitutes an elemental characteristic of a network [31, 69] that can be used e.g. to discern network types (see also section 2.11) or to reveal the existence of network hubs [122]. The degree distribution is given by the fraction of nodes that have degree k, for all k. Since in directed networks edge directions have to be taken into account, the in-degree and out-degree distributions are discerned. Alternatively, a two-dimensional joint distribution of in-degrees

and out-degrees might be considered [31]. The strength distributions in weighted networks are defined analogously [69].

Network density[BD,BU], also known as edge density or sparsity, is quantified as the ratio of actual edges to the possible number of edges.

The **average node strength**[WD,WU] measures the strength of interactions nodes receive or transmit on average, i. e. the average edge weight in the network. For weighted directed networks, the average in-strength equals the average out-strength, which in turn equals the average edge weight in the network. Network samples were analyzed with respect to the average node strength in the study presented in chapter 5.

The **network-wide clustering coefficient**[BD,BU,WD,WU] is the average of the microscopic scale clustering coefficients [27] (see also section 2.7.1). Therefore, it measures the density of triangles in the entire network, i. e. the tendency of a network to form tightly connected groups of nodes [90]. The average node-level clustering coefficient may be disproportionately influenced by nodes with low degree [14] and highly skewed degree distributions [31], a drawback that is avoided with the similar transitivity measure.

Transitivity[BD,BU,WD,WU] denotes the fraction of transitively closed node triplets [72, 65, 31]. It can be interpreted as the global clustering coefficient. In binary directed networks only so-called non-vacuous triplets (one edge terminating at and another edge originating from the center node v_i: $a_{hi} = a_{ij} = 1$ or $a_{ji} = a_{ih} = 1$) are considered and counted (denominator of the fraction). Only non-vacuous node triplets that are additionally closed in a transitive way with a directed edge from the first to the third node of the chain are also contributing to the numerator of the fraction of transitive node triplets. To incorporate edge weights, a triplet value ω is computed for each non-vacuous triplet centered on node v_i [65].

The **characteristic path length**[BD,BU,WD,WU] denotes the average length of the shortest paths between all pairs of nodes in the network. It is a measure of the typical separation between two nodes and therefore of the efficiency and cost of information transmission or spread of entities on a network, assuming that such processes tend to utilize direct, i. e. shortest, paths [68, 31, 27]. The characteristic path length is the most commonly used measure of functional

integration [2]. Weighted network samples were analyzed with respect to the characteristic path length in the study presented in chapter 5.[]

The efficiency e_{ij} in the information transfer between two nodes i and j is defined to be inversely proportional their shortest path distance $\Delta_{sp}(i,j)$ and is given by $e_{ij} = 1/\Delta_{sp}(i,j)$. The **global efficiency**$^{\text{BD,BU,WD,WU}}$ is the average efficiency of a network, i.e. the average inverse shortest path distance between all pairs of nodes [123] (also called the average conductance (inverse resistance) [124]). It is argued that global efficiency is generally better at capturing so-called small-world properties and the efficiency of parallel information transfer than the characteristic path length [125, 123].

The **local efficiency**$^{\text{BD,BU,WD,WU}}$ is the average efficiency of subnetworks consisting of all nodes adjacent to a node v_i, over all nodes [125, 123]. The role of the local efficiency is similar to the clustering coefficient as it measures for every node v_i how efficient the communication between its neighbors is [123].

The **radius**$^{\text{BD,BU,WD,WU}}$ is the smallest eccentricity (see also 2.7.1) in the network and is given by $rad = \min_{v_k \in \mathcal{V}} \epsilon(v_k)$ [75]. For directed networks, eccentricites take shortest paths to and from the respective node into account.

The **diameter**$^{\text{BD,BU,WD,WU}}$ is the maximum eccentricity in the network (i.e. the longest graph geodesic between any two nodes; see also 2.7.1) and is given by $dia = \max_{v_k \in \mathcal{V}} \epsilon(v_k)$ [75]. In other words, the diameter is the greatest distance using only shortest paths between any pair of nodes and describes the interconnectedness of a network.

In addition to the analysis loops of length three given by node triangles, in a directed network loops of length two and their frequency of occurrence might also be considered [31]. **Link reciprocity**$^{\text{BD,WD}}$ measures the propensity of node pairs to form reciprocated (mutual) edges and is given by the proportion of reciprocated edges between all pairs of nodes. A different definition of reciprocity as a correlation coefficient between entries of the adjacency matrix was proposed in [77]. Networks of the same domain seem to display similar ranges of reciprocity values [77] and heightened values of reciprocity were observed in functional networks [126].

The **assortativity coefficient**[BD,BU,WU] [127, 128, 31, 14], essentially a linear correlation coefficient, measures the level of assortative mixing (homophily) between connected nodes according to a scalar characteristic in a network, in particular an intrinsic property of the network nodes as given by node centrality indices like the degree (see also section 2.7.1). Networks that are assortatively mixed according to some property are sometimes also said to be stratified by that property. Assortative and disassortative behavior of nodes was found to have implications for network robustness (resilience) with respect to random failure of nodes and targeted deletion of most central nodes [129], with degree assortative networks being robust to targeted removal of nodes, which in some contexts might be a bad thing, e. g. due to the facilitation of epidemic outbreaks [128]. Networks that are assortative by degree typically have densely connected cores consisting of high degree nodes, which are surrounded by a periphery of chains of nodes with low degrees [31, 130, 122], whereas degree disassortative networks adopt star-like structures [31]. In [129] directed assortativity measures based on Pearson correlation for all four possible degree-degree correlations were proposed. An extension of the node degree assortativity coefficient to weighted networks was presented in [64].

Centralization[BD,BU,WD,WU], also known as network centrality, denotes a technique for defining a macroscopic, network-level scalar centrality quantity based on a comparison of the difference between the value $C_{nc}(v_i)$ of a particular node centrality measure (see also section 2.7.1) of a structural central node v_i and the respective centrality values $C_{nc}(v_k)$, $k = 1, ..., N$ of all other nodes [76, 131, 72]. Thereby, the centralization may be computed as $Cen(C_{nc}) = \sum_k \left(\max_i C_{nc}(v_i) - C_{nc}(v_k) \right)$.

Functional (and structural) brain networks are assumed to possess small-world network characteristics, which associate network structure with efficient information processing and transmission. **Small-world-ness**[BD,BU,WD,WU] is a continuous quantitative measure (as opposed to a categorical distinction) of the manifestation of small-world characteristics [132]. It relates the transitivity and the characteristic path length of the investigated network with the respective quantities obtained from an ensemble of null model networks (see also section 2.11.1). However, in [133] it was argued against the widespread belief that structural brain networks represent small-worlds (the investigated functional brain networks of this thesis also usually show only weak small-world characteristics).

2.8 Network module detection algorithms

Structural analyses of networks often require a network decomposition into functionally self-contained subnetworks [80] like the ones induced by network modules or other mesoscopic structures like motifs [134] or connected components defined by the removal of high-degree nodes [135], high-betweenness nodes [136] or nodes with other particular properties, e.g. [88]. The obtained subunits are not only functionally interesting on their own, using them as the basis for subsequent anlyses might help to substantially reduce the complexity of the analysis task. This section is concerned with the algorithmic detection of network modules (see also section 2.7.2) by means of assigning densely interconnected, cohesive groups of nodes the same module affiliation. For it, a large number of module detection methods are available, which utilize different algorithmic strategies to exploit and interpret structural information inherent to the network data [70]. Therefore, an identified network partition into modules is not necessarily unique, and consequently different partitions of similar quality and equal legitimacy might exist. To further fine-tune and optimize the result of a module detection algorithm, the Kernighan-Lin algorithm [137] might be applied to the obtained network partition [118]. This algorithm attemtps to further increases the quality of the network partition by iteratively swapping subsets of nodes or single nodes between modules. A commonly used module structure quality function is called modularity, which quantifies the extent to which edge densities within modules exceed the densities expected on the basis of chance (see also section 2.10). The lack of comparability of these algorithms makes it hard to decide on the most suitable one for identifying modules in given network data [70]. Most module detection algorithms are not deterministic and their results typically depend on specific random seeds, initial conditions generated at random, random selections made by the algorithm at run time and the tie-break rules adopted for their execution [138]. Consequently, different runs of an algorithm might yield different network partitions. If for a given network multiple different network partitions are returned by an algorithm, my strategy was to always select the highest quality network partition as measured by the modularity statistic. Depending on available time and computational resources a consensus clustering of different results can be generated, which enhances the accuracy of the finally obtained partition (see also section 2.9). Very recent approaches incorporate metadata on nodes to improve the accuracy of module detection [139].
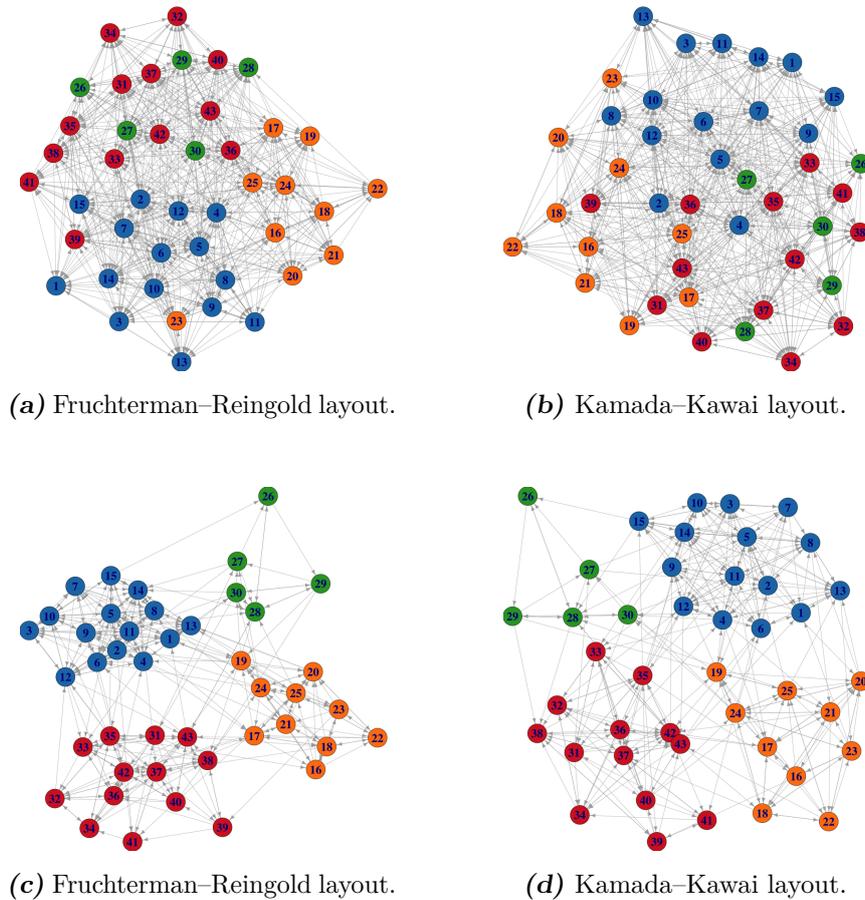
(a) Fruchterman–Reingold layout.

(b) Kamada–Kawai layout.

(c) Fruchterman–Reingold layout.

(d) Kamada–Kawai layout.

Figure 2.1 – **Under the right conditions, module structure can be already evident from network drawings.** If the adjacency matrix of a network has a clear intrinsical module structure, it becomes apparent in network plots where node positions are determined by force-directed layout algorithms, e.g. Fruchterman–Reingold [140] or Kamada–Kawai [141]. Such layout algorithms place densely interconnected subsets of nodes together using attractive forces while using repulsive forces to separate weakly interacting nodes (e.g. node pairs that are connected by long shortest paths). Panels (a) and (b) show two different plots of the same network without module structure. This network was obtained from the simulated network with module structure shown in (c) and (d) by adding random interactions ('noise'). The color-coded module affiliations of nodes were given by construction.

As a remark, it is often observed empirically, e.g. in [142], that network drawing algorithms such as force-based layout algorithms, e.g. [140], and force-spring layout based algorithms, e.g. [141], potentially reveal the existence of module structure to some extent in the visualized network structure. By way of illustration, figure 2.1 shows how under ideal circumstances network modules can be identified by visual inspection of different network layouts, which tend to automatically separate such node groups without depending on a dedicated network module detection algorithm. Despite the criticism and caution outlined above, identification of module

structure in brain networks represents an opportunity to gain insights into the functional segmentation of brain areas during certain recorded states of neural processing. To analyse functional connectivity network data with respect to their module structure I selected several algorithms for application that have a good reputation in the network science research community and that are based on very different algorithm design strategies. An additional constraint for the selection of module detection algorithms is yielding interpretable results with respect to underlying neural mechanisms. Thus, in my work I was focused solely on finding clear-cut and unambiguous modules, which means I did not take into account hierarchical [143, 92] or overlapping module structure [144, 145, 146], even though, by their very nature, some of the selected algorithms are uncovering hierarchies of module structure. While the orientation of edges might encode potentially useful functional information that should not be discarded, the effect of edge orientation on module detection accuracy is not a priori obvious and might not be generalizable from one network data set to another, as an edge in any direction indicates a potential commonality of nodes by virtue of their interaction. Following this reasoning, I also applied algorithms designed for undirected networks on symmetrized versions of directed adjacency matrices or edge weight matrices, in which any directed edge is replaced by an undirected one (see also section 2.2). Among other lines of research on network topology, from research on stochastic block modeling (e. g. [147, 148, 149, 150, 151]) it is known that module structure inferred from weighted networks might reveal latent structures that are qualitatively distinct from the ones obtained in their dichotomized unweighted counterpart networks [39]. Module structure analyses could be gainfully augmented by analyzing the network at hand with respect to a (multi)core-periphery structure, as module detection algorithms cannot discern peripheral nodes from the dense modules to which they are connected [122, 130, 80]. The following concise descriptions only cover the algorithms that were used for this thesis.

The **leading eigenvector algorithm of Newman**[BU,WU] [152, 153] is based on spectral decomposition of the modularity matrix to express modularity (see also section 2.10) in terms of eigenvalues and eigenvectors (cf. [147]). It forms the basis for the Leicht-Newman algorithm described next.

The **Leicht-Newman leading eigenvector algorithm**[BD,WD] [154] is a generalization of the spectral modularity optimization approach for undirected networks [152, 153] to directed networks. It relies on the so-called modularity matrix $B = \mathcal{A} - P$, where $\mathcal{A}$ is the adjacency

matrix of the network and P contains the probability of each edge in a random network drawn from the configuration model [155, 156, 157]. The method computes the eigenvector for the largest positive eigenvalue of the symmetric matrix $B + B^T$. The signs of the elements of this eigenvector are used to divide the network into two modules and assign module membership to nodes. In a subsequent heuristical "fine-tuning" step similar to the Kernighan-Lin algorithm [137] nodes are assigned and re-assigned to any of the two modules in an attempt to further increase the modularity of the network partition. Modules found this way are further subdivided by repeated bisection as long as modularity can be increased.

The algorithm of **Blondel *et al.*** [BD,BU,WD,WU], the so-called **'Louvain algorithm'** [158] (cf. [159]), is based on iterated greedy local optimization of the network's modularity characteristic (see also section 2.10). The first iteration of the algorithm assigns one module to each node and then repeatedly and sequentially identifies for each node the gain in modularity when removing it from its module and placing it into the module of any of its neighbor nodes. Subsequently, the displacement that yields the maximum positive gain is then performed and the possible displacements of the next node are considered. The first iteration comes to an end when no further modularity improvement is possible by individual node displacements. The nodes of every identified module are contracted into a single node and their edge weights are aggregated to build a new network on which further iterations takes place. In these further iterations the greedy step is repeated and modules are merged until no further improvement of modularity can be achieved.

The **fast greedy modularity optimization**[BU] algorithm [160] is an optimized version of the greedy algorithm for optimization of the modularity criterion (see also section 2.10) that was proposed in [161]. The key idea is, starting with single nodes, to repeatedly choose and merge a pair of modules that are connected, so that in each merge step the biggest gain or, if no increase is possible, the smallest decrease in modularity is attained. Consequently, in each step this algorithm can only exploit local information about the module structure to base its merges upon. The optimizations are achieved by using efficient data structures for sparse matrices and by eliminating needless operations on the network's adjacency matrix during runtime. This algorithm has the reputation of being fast enough for applying it to very large networks, however its estimations of the modularity maximum is not as accurate as the results of other methods [138].

The **random walk algorithm of Pons and Latapy**[BU] [162, 163] detects network modules using a random walk process along edges to measure distances between nodes. The basic intuition here is that short random walks are likely to stay for most of the time within a network module, where the edge patterns are more dense as compared to inter-module regions. The probabilities for the transition from a node v_i to an adjacent node v_j are only determined by the degree of v_i and define a transition matrix. Two nodes v_i and v_j of the same module have very similar probabilities for a transition to a third node v_k. This idea is incorporated into the distance measure which uses powers of the transition matrix and accounts for module membership of nodes and thus captures the module structure of the network. The problem of finding network modules is reduced to a clustering problem by generalizing the node-node distance to a distance between modules and using it in an agglomerative clustering approach which computes network partitions. Finally, the partition maximizing a certain quality criterion is selected as the result.

The **Potts spin glass algorithm**[BD,BU,WD,WU] reformulates the problem of module detection as the problem of finding the global ground state configuration of an infinite range spin glass [164, 165, 166] with the states of the N spins being equivalent to the module labels [167]. The principle of the spin glass model is that edges should preferably connect nodes of the same spin state, whereas nodes with different spin states should be disconnected. Respective edges are energetically rewarded or penalized by the Hamiltonian (cost function) of the spin model, which is the sum over all interaction energies, which depend on the spin states, i.e. module affiliations, of the nodes of each considered node pair and the strength of their interaction. The interaction energy of two nodes is only added if the nodes are in the same spin state. It is positive (ferromagnetic, aligned interaction) only if the interaction strength of two nodes is larger than expected under a given null model [118]. If their interactions strength is less than expected under the null model, the energy is negative (antiferromagnetic, differently oriented interaction). The equation of Hamiltonian bears a strong resemblance with the expression of modularity (equ. 2.22) and contains the resolution parameter γ that expresses the relative contribution from existing and missing edges to the energy (adjusts importance of the null model term) and allows to tune the number of discovered modules in the minimum energy partition, thus, enabling the exploration of the hierarchical organization of module structure. It was shown that finding the spin configuration for which the Hamiltonian is minimal is

equivalent to maximizing modularity [167]. The ground state of the Hamiltonian is found using a simulated annealing algorithm [167].

The **Infomap**[BD,BU,WD,WU] algorithm [168] finds modular structure with respect to flow and is based on the idea that the problem of module structure identification can be turned into an equivalent coding problem[7], where a network partition is sought that yields a minimum description length of a random walk across the network structure, which is given by the 'map equation' objective function [170]. The idea of the 'map equation' is using Huffman codes [171] for an optimal two-level description of the random walk, which is obtained by differently labeling regular important structures (modules) and insignificant details (nodes in the modules) that are encountered by it. Whenever the random walk transitions into a different module, where it likely has a long persistence time, a unique binary codeword will be used for the encountered module in the description of the random walk. The lengths of Huffman codewords is derived from the visit rates of the structures, with frequently encountered structures having short codewords. The nodes that are encountered by the random walk process within a module also get unique codewords in the description of the random walk, but these are much shorter than the module codewords and they are reused among different modules to compress the description. This 'network map' resembles somewhat geographic maps where different cities have mostly unique names, but share a large number of street names. Thus, finding an optimal code that describes the random walk solves the dual problem of finding the regularities and important structures of the network. To minimize the 'map equation' a certain number of modules have to be used (the description length associated with transitions among modules increases with the number of modules, whereas the description length associated with the movements within modules decreases with the number of modules) and it has to be determined which nodes belong to which modules. Given a network partition, the 'map equation' yields the lower bound for the description length, without actually generating the description code. Whereas modularity accounts for the relationships between any two nodes, the 'map equation' accounts for the way local interactions induce a system-wide network flow. As a result, optimization of these quality functions uncovers different module structures. The optimization of the map equation may be carried out by combining a greedy search with a simulated annealing step or by using an adapted variant of the fast greedy modularity

---

[7]Minimum description length statistics is concerned with the duality of data compression and extracting patterns within this data. The more the data can be compressed, the more can be learned about its regularities [169].

optimization technique of Blondel *et al.* (see above) [170].

2.9 Consensus clustering

Consensus clustering is a method for integrating different related network partions into a single one, while enhancing trends and consistent features in the partition data and removing or smoothing out noise in the module assignments. Such partitions might be obtained from different runs of a module detection algorithm or from different network configurations at different time steps. In this work, consensus clustering is primarily used for dynamic module structure analysis in small and dense EEG-derived networks (see also chapter 7) as an approach to simplify the amount of generated network partition data and to extract trends in the network module structure time course [138]. In essence, a consensus network partition of snapshot networks in a given time window accounts for the history of the network system and constitutes a summary of its time-evolving module structure in either the vicinity of pre-selected isolated time frames corresponding to specific network snapshots or in snapshot networks of several overlapping time windows that cover the entire network evolution. The selected time window length, i.e. the number of combined partitions is important. If too few partitions are combined, their consistent features might not be sufficiently extracted. If, however, the selected time window is too broad, time-specific meaningful features of network module structure are also removed by the smoothing associated with the procedure. Therefore, one should strive for a balanced approach and avoid merging network partitions referring to a time range that is much broader than the natural time scale of the network evolution. Identification of a meaningful time scale [172] for the network evolution usually requires prior knowledge and exploratory data analysis. Snapshot time frames need not be equidistant to each other and the number of combined snapshot networks, i.e. the width of time windows, along the network sequence might be chosen to vary.

At the heart of the methodology proposed in [138] is the so-called $N \times N$ symmetric consensus matrix $\mathscr{C}$, whose entries $\mathscr{C}_{ij}$ store the number of partitions in which nodes i and j are comember nodes of the same module, divided by the number of partitions, which contain both nodes. As said before, in the context of this thesis nodes in the functional connectivity networks correspond to brain recording sites and therefore their number remains static over time. The partitions used for the construction of $\mathscr{C}$ are obtained from applying a module detection

algorithm to a number of snapshot networks[8] at different time steps. This module detection algorithm can be designed to handle either binary or weighted networks. Nodes that often share the same module affiliation get relatively large entries in the consensus matrix, while nodes at the boundary of modules that are often misclassified due to interaction noise are linked by low weights in $\mathscr{C}$. In a subsequent filtering step, all entries $\mathscr{C}_{ij}$ with a weight below a threshold τ are set to zero. Nodes that would become disconnected due to this filtering keep their largest entry in $\mathscr{C}$. Then a module detection algorithm for weighted networks is applied to the consensus matrix to obtain a new set of partitions. If the newly obtained set of partitions is still heterogeneous, a new consensus matrix is constructed from this new set of partitions and the procedure is iterated. The iterations are stopped if the partitions obtained from a consensus matrix $\mathscr{C}$ are homogeneous, which means that $\mathscr{C}$ has a block-diagonal structure with entries of weight of (almost) 1 for nodes of the same block and entries of weight of (almost) 0 for nodes of different blocks. The convergence of entries in $\mathscr{C}$ usually occurs after only a few iterations of the procedure.

2.10 Measures for module structure quality assessment

For many applications obtaining a natural partition of the node set into groups of well interconnected nodes provides crucial phenomenological or functional information about the underlying system (cf. [15]). This suggests the importance of making a precise distinctions about 'clear-cut' high quality or 'weak' low quality network partitions. Naturally, the quality and definiteness of identified network partitions depends on the local topological context of clustered nodes and single modules given by the spatial distribution of edges, the heterogeneity of edge density and the technique used to uncover partitions. Module structure quality might also depend on correlated metadata on nodes, if available, which is not considered here [139]. Different network partitions might be compared with each other, using the module preservation and quality measures described in this section. Following aspects of module structure quality might be assessed: Module structure preservation can be quantified with cluster resemblance measures (ratio of correctly classified nodes, van Dongen metric, Rand index, Jaccard index), information theoretic measures (variation of information, mutual information), edit distance-based measures (Hamming distance between adjacency matrices, partition edit distance) and measures from statistics (Cohen's Kappa). Module structure quality and definiteness

---

[8]To improve module detection results in a single network instance, the described procedure might be applied to several of its partitions obtained from different runs of one or several module detection algorithms.

of obtained module structure can be assessed by topological measures (modularity measure, performance measure, coverage, number of disconnected node pairs) and measures assessing the clustering validity, i. e. the fit of node–module assignments based on structural similarity and equivalence of node pairs (overall average silhouette width).

The **ratio of correctly classified nodes** is obtained by comparing the network partition yielded by a particular module detection algorithm with another (benchmark) classification of nodes. It measures the 'goodness' of a partition simply by counting misclassified nodes. To compute this percentage, module labels in the different network partitions have to be matched (see also chapter 3 section 3.2). This approach causes problems when the number of modules in both partitions is different, e. g. when modules in one partition might have been separated into several modules in the other partition. In this case, only one of separated modules can be matched to their counterpart in the other partition. Consequently, the ratio of correctly classified nodes will be greatly reduced by virtue of this split, even though a large fractions of the same node pairs are still clustered together. This situation is handled more robustly by the Rand index that is described next.

The **Rand index** [173], one of the most commonly used partition resemblance measures [174], compares two different partitions $\mathscr{P}_A$ and $\mathscr{P}_B$ of the same node set on the basis of counting and comparing classifications of pairs of nodes in both partitions. Thereby, it does not make use of topological information, i. e. adjacency information of the network. The Rand index is given by the number of identically classified comembership and non-comembership of node pairs normalized by the total number of node pairs

$$R(\mathscr{P}_A, \mathscr{P}_B) = \frac{\rho_{11} + \rho_{00}}{\rho_{11} + \rho_{00} + \rho_{01} + \rho_{10}} \qquad (2.9)$$

where ρ_{11} (true positives) represents the number of node pairs that are comembers in both, $\mathscr{P}_A$ and $\mathscr{P}_B$, and ρ_{00} (true negatives) denotes the number of node pairs that are separated and assigned to different modules in $\mathscr{P}_A$ and $\mathscr{P}_B$. There are two more types of classified node pairs that represent disagreement. More specifically, the number of pairs of nodes that are assigned to different modules in $\mathscr{P}_A$ but are placed in the same module in $\mathscr{P}_B$ (ρ_{01}, false positives) and the number of pairs of nodes where the situation is the opposite (ρ_{10}, false negatives). In practice the Rand index does not necessarily range over the entire $[0, 1]$ interval and instead often concentrates in a small interval close to 1. Therefore, it might be adjusted

for chance assignment of modules [175, 176, 174]. The **adjusted Rand index** is given by

$$AR(\mathscr{P}_A, \mathscr{P}_B) = \frac{R(\mathscr{P}_A, \mathscr{P}_B) - E[R]}{1 - E[R]} \tag{2.10}$$

The Rand index is a measure of similarity, which can be easily converted to a measure of distance between two network partitions by taking the one-complement [173]. This distance is a metric on the set of all clusterings of a given set of nodes

$$d_R(\mathscr{P}_A, \mathscr{P}_B) = 1 - R(\mathscr{P}_A, \mathscr{P}_B) \tag{2.11}$$

The **Jaccard index** is a similarity measure, which compares network partitions and the extent to which their module assignments overlap. It is based on counting pairs of nodes in which both partitions agree or disagree in a way similar to the Rand index (see above) [177, 176]. The Jaccard index is given by

$$J(\mathscr{P}_A, \mathscr{P}_B) = \frac{\rho_{11}}{\rho_{11} + \rho_{01} + \rho_{10}} \tag{2.12}$$

where, as before for the Rand index, ρ_{11} denotes the true positives, ρ_{10} denotes the false negatives and ρ_{01} denotes the false positives. The Jaccard index can also be easily transformed into a distance function (dissimilarity measure) by the one-complement, which measures the dissimilarity between two network partitions

$$d_J(\mathscr{P}_A, \mathscr{P}_B) = 1 - J(\mathscr{P}_A, \mathscr{P}_B) \tag{2.13}$$

Other similarity measures for binary data based on $\rho_{11}, \rho_{00}, \rho_{01}, \rho_{10}$ can be found in [178].

The **variation of information** compares two network partitions of N nodes into modules, $\mathscr{P}_A = \{\mathcal{M}_1^A, ..., \mathcal{M}_K^A\}$ and $\mathscr{P}_B = \{\mathcal{M}_1^B, ..., \mathcal{M}_L^B\}$, by measuring the change in their information content [176]. It is a true metric on the space of network partitions and can consequently be used to calculate the distance (i.e. dissimilarity) between two partitions of the same network data or between the module structure of two different but equally sized

networks (i.e. for comparing members of a functional connectivity network sample). Variation of information is not using topological information of the input network as it relies only on module affiliations of nodes, i.e. the information on the subsets of nodes that are grouped together. Thereby, it does also not depend on the ordering of assigned module affiliation labels. To derive this network partition distance measure, two quantities are established: the amount of information with respect to node assignments contained within each partition and the amount of information one partition has about the other. First, for each partition $\mathscr{P}_A$ and $\mathscr{P}_B$ a discrete random variable is defined that models the module affiliation of a node picked uniformly at random from the network's N nodes. The probability mass functions associated to these random variables are $Pr(k) = N_k/N$ and $Pr(l) = N_l/N$, giving the probability of a randomly picked node being member of module $\mathcal{M}_k^A$ consisting of N_k nodes or module $\mathcal{M}_l^B$ consisting of N_l nodes, respectively. The entropy H associated with a network partition quantifies the uncertainty of node assignments (the uncertainty of the module of a randomly picked node), that is, the information needed on average to describe the node assignments in each partition. The entropy of node assignments in partition $\mathscr{P}_A$ is calculated as

$$H(\mathscr{P}_A) = - \sum_{k=1}^{K} Pr(k) \, log \, Pr(k) \tag{2.14}$$

The calculation for $H(\mathscr{P}_B)$ is analog. $Pr(k, l)$ is the joint probability distribution of the two random variables associated with $\mathscr{P}_A$ and $\mathscr{P}_B$

$$Pr(k, l) = \frac{\mid \mathcal{M}_k^A \cap \mathcal{M}_l^B \mid}{N} \tag{2.15}$$

It specifies the probability that a node is at the same time assigned to module $\mathcal{M}_k^A$ in $\mathscr{P}_A$ and module $\mathcal{M}_l^B$ in $\mathscr{P}_B$. Then, the reduction of uncertainty of node assignments in one partition due to knowledge about the other partition is given by the **mutual information I** [179] between the associated random variables

$$I(\mathscr{P}_A, \mathscr{P}_B) = \sum_{k=1}^{K} \sum_{l=1}^{L} Pr(k, l) \, log \, \frac{Pr(k, l)}{Pr(k)Pr(l)} \tag{2.16}$$

Only if the two network partitions to be compared are equal, the following holds

$$I(\mathscr{P}_A, \mathscr{P}_B) = H(\mathscr{P}_A) = H(\mathscr{P}_B) \tag{2.17}$$

Finally, the normalized variation of information is given by

$$VI(\mathscr{P}_A, \mathscr{P}_B) = \frac{H(\mathscr{P}_A) + H(\mathscr{P}_B) - 2I(\mathscr{P}_A, \mathscr{P}_B)}{log\ N} \tag{2.18}$$

The numerator represents the information about $\mathscr{P}_A$ that is lost and the information about $\mathscr{P}_B$ that is gained when partition $\mathscr{P}_A$ is changed or converted to partition $\mathscr{P}_B$. Normalization by dividing by $log\ N$ ensures that VI is bounded to the intervall $[0, 1]$ [179].

The **van Dongen metric** on the space of partitions of the same set of nodes is based on set matching and is given by [180, 176, 179]

$$d(\mathscr{P}_A, \mathscr{P}_B) = 2N - pn(\mathscr{P}_A \to \mathscr{P}_B) - pn(\mathscr{P}_B \to \mathscr{P}_A) \tag{2.19}$$

Again, $\mathscr{P}_A$ and $\mathscr{P}_B$ denote the two partitions to be compared, which consist of a number of modules k and l, respectively. N denotes the number of nodes in the network and $pn(\mathscr{P}_A \to \mathscr{P}_B)$ is the projection number of $\mathscr{P}_A$ onto $\mathscr{P}_B$, which reflects how close $\mathscr{P}_A$ is to being a subpartition of $\mathscr{P}_B$. The projection number measures the amount of module overlap in both partitions and is defined as

$$pn(\mathscr{P}_A \to \mathscr{P}_B) = \sum_{\mathcal{M}_k^A \in \mathscr{P}_A} \max_{\mathcal{M}_l^B \in \mathscr{P}_B} \mid \mathcal{M}_l^A \cap \mathcal{M}_l^B \mid \tag{2.20}$$

The **partition edit distance** quantifies topological alterations of intra-module edge patterns. For each pair of corresponding, matched modules in $\mathscr{P}_A$ and $\mathscr{P}_B$ the Levenshtein edit distance [181, 182] of intra-module edges is computed, which is the cost for their optimal alignment. For it, the adjacency matrix for each module is vectorized and typecasted to a string in the alphabet $\{0, 1\}$. The Levenshtein distance is the minimum number of insertions, deletions and substitutions to make both strings equal. Single edit distances for each pair of corresponding modules are added up to yield the partition edit distance. Alternatively, Kappa coefficients could be computed to quantify the inter-rater agreement of intra-module edge patterns (see also section 2.10).

The **modularity**$^{\text{BD,BU,WD,WU}}$ measure [154, 148, 70, 183, 184] quantifies the degree to which a given network partition into modules is clearly delineated by means of accounting for

the magnitude of local edge densities. Networks possessing a strong module structure have statistically surprisingly high intra-module edge densities and low inter-module edge densities, thus yielding high values of modularity. It is defined as

$$
\begin{aligned}
Q = &\text{(fraction of intra-module edges)}- \\
&\text{(expected fraction of such edges if placed at random)}
\end{aligned}
\tag{2.21}
$$

For binary undirected networks modularity is given by the following equation

$$
Q^{BU} = \frac{1}{2|\mathcal{E}|} \sum_{ij} (\mathcal{A}_{ij} - \frac{\kappa_i \kappa_j}{2|\mathcal{E}|}) \, \delta(c_i, c_j)
\tag{2.22}
$$

where κ_i denotes the degree of node v_i and c_i denotes the module of node v_i. The Kronecker-δ function is defined as usual, i.e. it is unity if and only if both arguments are equal and it is zero otherwise. Thus, Q^{BU} quantifies the existence of each edge inside a module by comparing it with its expected probability in an equivalent null model network, which could be generated by the configuration model (see also section 2.11.2). Equation 2.22 also states that edges between large degree nodes are not 'surprising' and add only little to the modularity of a given network partition, with the rationale being that in a modular network edges are not homogeneously distributed since the density of edges inside modules is higher than expected under random edge placement. Modularity for binary directed networks quantifies 'surprising' edges with respect to the in-degree and out-degree information of their tail and head nodes, e.g. edges that fall between pairs of nodes where the tail-node has small out-degree and the head node has small in-degree. It is defined analogously to equation 2.22 as follows

$$
Q^{BD} = \frac{1}{|\mathcal{E}|} \sum_{ij} (a_{ij} - \frac{\kappa_i^{out} \kappa_j^{in}}{|\mathcal{E}|}) \, \delta(c_i, c_j)
\tag{2.23}
$$

Modularity for weighted directed networks can be calculated as [185]

$$
Q^{WD} = \frac{1}{w} \sum_{ij} \left(w_{ij} - \frac{\zeta_i^{out} \zeta_j^{in}}{w} \right) \delta(c_i, c_j)
\tag{2.24}
$$

where $w = \sum_{ij} w_{ij}$ is the total weight of all edges in the network under consideration. The definition for modularity in weighted undirected networks can be found e.g. in [185, 184].

Since modularity might occasionally have increased values in random networks [179], for

module structure quality assessment it is a good strategy to not rely on the modularity characteristic alone. It was also found that modularity optimization, while it works well in many situations, suffers from a resolution limit [186], which means that a maximum modularity partition can fail to resolve modules smaller than a certain size. In particular, if the number of internal edges of an identified module is smaller than about $\sqrt{2|\mathcal{E}|}$, then the identified module might be an artificial merger of two or more sufficiently smaller modules. In the worst case, if all modules of a network are 'fuzzy communities' that maintain a large number of inter-module edges, the identified supermodule could in principle have a size of the order of the size of the whole network. Incorporating a resolution parameter into the optimized quality function is a way to address the resolution limit [167, 118]. Optimization of modularity is a nondeterministic polynomial-time complete ($\mathcal{NP}$-complete) problem [187] and is therefore computationally intractable[9], i. e. there is no hope for designing an fast and exact algorithm that computes maximum modularity partitions for all inputs. As a consequence and unless $\mathcal{P} = \mathcal{NP}$, any polynomial-time algorithm has to make use of heuristics. On the other hand, the result of the $\mathcal{NP}$-completeness of modularity maximization can be seen as a justification for using approximation algorithms. To overcome these limitations in weighted networks, an exact procedure for network size reduction that preserves modularity and allows for a more exhaustive search of the network partition space by heuristic modularity optimization algorithms was proposed [185]. In addition to these difficulties, it was found that the modularity function exhibits degeneracies as it admits an exponential number of distinct high-scoring, near-optimal solutions and typically lacks a clear global maximum [188]. Consequently, analysis of real-world network partitions obtained by modularity maximization warrants cautious interpretations and the output of several runs of an algorithm should be compared if feasible in practice. It also implies that results of different module structure identification algorithms are not unlikely to differ from each other. These issues are more pronounced in larger networks and in binary networks [15]. On the other hand, the exponential number of near-optimal solutions provides an explanation for the good performance of modularity optimization in practice, despite its $\mathcal{NP}$-completeness: good partitions are not hard to find, even though the optimal solution is obscured [188]. A similar behavior is expected for other quality functions and identification techniques, too [188, 186]. A local version of the modularity measure was proposed in [189].

The **performance measure**[BU] is given by the fraction of node pairs that are correctly

---

[9]No polynomial-time algorithm has yet been discovered for an $\mathcal{NP}$-complete problem and also it has not been proved that no polynomial-time algorithm can exist for any such problem [34].

'interpreted' by an obtained network partition with regard to the network adjacency information [70]. It takes into account the node pairs that are assigned the same module and that also interact via an edge and those node pairs where both nodes are classified to belong to different modules that are not connected by an edge. For the performance measure only the presence of an interaction is relevant as it indicates a putative common role of nodes and a less unlikely common module affiliation. Since edge directions are ignored, directed networks have to be symmetrized. The performance measure is given by

$$Per(\mathscr{P}) = \frac{|\{(i,j) \in \mathcal{E}, \mathcal{M}^{(v_i)} = \mathcal{M}^{(v_j)}\}| + |\{(i,j) \notin \mathcal{E}, \mathcal{M}^{(v_i)} \neq \mathcal{M}^{(v_j)}\}|}{N(N-1)/2} \qquad (2.25)$$

where $\mathcal{M}^{(v_k)}$ denotes the module of node v_k. In other words, the performance measure penalizes edges that are ignored by a given network partition, i.e. if both nodes of an edge are assigned to different modules and it penalizes edges implied by the network partition that are not present in the network, i.e. if nodes with the same module affiliation are unconnected.

Coverage[BD,BU] denotes the ratio of the number of intra-module edges by the total number of edges [70]. The motivation behind it is the following: in an ideal module structure, e.g. if a network is fully fragmented into isolated connected components with no inter-module edges linking nodes of different connected components, there would be little ambiguity with respect to the (non-hierarchical) module structure. In this case all edges are intra-module edges and the value of coverage is one. Thus, coverage measures the goodness of the obtained network partition into modules in dependence of the quality of the network's inherent module structure.

The **overall average silhouette width**[BD,BU,WD,WU] quantifies the disbalance between the average intra-module dissimilarity of all nodes and their smallest average inter-module dissimilarities. It equals the average over all individual silhouette values $s(v_i)$ [190] of each node v_i, which measure how well each v_i is classified with respect to the network partition at hand. For it, the average intra-module dissimilarity $a(v_i)$ is compared with the minimum of its average inter-module dissimilarities $b(v_i)$ (the average dissimilarity of v_i to its second-best module). In a high quality network partition $a(v_i)$ values are small and $b(v_i)$ values are high, which means that node v_i is similar to its comember nodes, whereas all other modules consist of dissimilar nodes. Using an appropriate node dissimilarity measure to compute $a(v_i)$ and $b(v_i)$

the silhouette of a node is given by

$$s(v_i) = \frac{b(v_i) - a(v_i)}{max\{a(v_i), b(v_i)\}} \tag{2.26}$$

Consequently, silhouette values close to 1 indicate good module assignments, whereas silhouette values close to 0 and close to -1 indicate ambiguous module affiliations and misclassification, respectively. To obtain a characteristic for a network partition in its entirety the overall average silhouette width is computed. An example silhouette plot is presented in figure 6.5. To determine node dissimilarity in the absence of spatial embedding of network data and information on contextual similarity of nodes, a structural equivalence[10] measure similar to ones presented in [191] could be used. Structural equivalence measures consider adjacency relationships between neighboring nodes, i.e. common neighbors that pairs of nodes share. This corresponds to a common theme of data clustering, where similarity between objects can be related to the number of features they share [192]. This notion of similarity between nodes corresponds also to a well-established result in sociology, that people are increasingly likely to be friends, the more commonalities they share [193]. One possible dissimilarity measure and the one I used for this thesis accounts for the direction of edges between shared neighboring nodes (cf. [70]). It takes on the following form

$$D(v_i, v_j) = \sqrt{\sum_{k \neq i,j} [\mathcal{A}_{ik} - \mathcal{A}_{jk}]^2 + [\mathcal{A}_{ki} - \mathcal{A}_{kj}]^2} \tag{2.27}$$

Its suitability depends on the network data and the actual application, due to the way this measure determines dissimilarity of nodes: Any node v_m connected to only either v_i or v_j increases the dissimilarity. Any shared neighboring node v_m linked by the same kind of edge to both, node v_i and v_j doesn't affect the value of D. However, any shared neighboring node v_m linked by different kinds of edges to v_i and v_j contributes to a further increase of dissimilarity. Consequently, D might take on large values for nodes of high degree, even if they share a large number of neighboring nodes, since high-degree nodes are more likely to also have a larger number of neighboring nodes they don't share. In this situation the resulting dissimilarity value D might be even larger than for some nodes of small degree, that have only a small overlap of neighboring nodes and at the same time have only few neighboring nodes they don't share. Depending on the application this emphasis on the

---

[10]*Structurally* equivalent nodes share the same neighboring nodes to a large extent. *Regularly* equivalent nodes do not necessarily have to share neighboring nodes, but their neighboring nodes are themselves similar [31], therefore two nodes are similar to the extent that their neighbors are similar.

direction of edges linking neighboring nodes (and not the fact alone that two nodes share a certain number of neighbors) might be an undesirable trait of the proposed distance measure D. The similarity/dissimilarity of nodes could also be established by other means, e.g. by the topological overlap measure [143, 62], m-th order generalized topological overlap measure [194], Jaccard distance between node neighborhoods Γ_i and Γ_j [70], the measure based on regular equivalence and structural equivalence proposed in [191], correlations between rows or columns of the adjacency matrix [31, 70], correlations of time-variant node properties (cf. [195]), cosine similarity of node adjacency vectors [31, 192, 70], the distance of Adamic–Adar [193], 'most reliable route' similarity [196] or measures based on edit distances or the Hamming distance of node adjacency information [31].

For each module the **number of disconnected nodes**[BD,BU,WD,WU] can be interpreted as a simple measure of fragmentation. Total or median numbers accounting for all modules can be used for comparing entire partitions.

Cohen's Kappa coefficient[BD,BU] [197] can be used as a topological measure of network module structure similarity by means of quantifying the agreement between edge patterns given by corresponding adjacency matrix entries.

2.11   Network models

Network models are used in this thesis primarily as null models to contrast regularities in investigated real-world networks with random network structure in the absence of functional constraints. However, it was found that network topology might be classified into four different topological classes [198, 199], which cannot be completely reproduced by the commonly deployed network models (*Erdös–Rényi random networks* with uniform edge probability distribution [200, 201], *Barabási–Albert preferential attachment* (scale-free, i.e. power-law) networks [28, 201, 143] and *random networks with prescribed degree sequence* (see also section 2.11.2)). Of these network models, the random network model with prescribed the degree sequence is of particular relevance for the analysis of functional brain networks (see also sections 2.11.2 and 2.12). As geometric constraints can give rise to unique subnetwork distributions, *random geometric graphs* (edges connect pairs of nodes that are proximal in space [202, 115] or interactions decay with the distance between nodes [203]) might be used as an analysis tool for discerning whether the abundance of certain subgraphs is caused by such purely geometric

constraints or whether they arise due to additional optimizations and characteristics of network structure design. *Small-world networks* [27, 204] provide an appealing description of brain networks due to their inherent incorporation of integration (high clustering coefficient) and segregation (short characteristic path length) that can be achieved with low wiring costs [124]. These network models are also interesting from the perspective of studying the resistence of network structure to random failures and targeted attack on nodes [31] (cf. [205, 206]).

2.11.1   Network null models

Null models correspond to a null hypothesis and are used to analyze patterns and test hypotheses [207]. A null model is broadly defined as a pattern-generating model that is based on constrained data randomization [208]. For the construction of the null model, certain aspects of the observed data are preserved while others are randomized and vary stochastically to create new patterns in the null model data [209, 210, 211, 212, 213, 214, 215, 128]. Network null models like the network models and random graph models of the previous sections are important for network analysis to validate results using surrogate-assisted approaches (see also section 2.12), since identified network characteristics might be the product of random effects and contingencies in the network structure or are simply caused by mechanistic effects of basic network properties. Consequently, distinctly nonrandom characteristics of network topology are linked to functionally important substructures. Therefore, real-world network measures should be compared with 'default' values identified in appropriate null model networks. Thereby, a network null model is needed to construct a reference system that contrasts such topological regularity with random effects that also influence the network topology. Particularly, the effects of any process that created structures with functional relevance in the networks have to be reversed during the generation of the null model networks. Choosing a suitable null model that fits given network data is an open problem. In particular, it is difficult to decide which low-level topological properties of the network data should be captured by the null model networks while at the same time the connectivity between nodes vary stochastically. Imposing too many constraints on the null model will reduce its statistical power [122]. Thus, type II statistical errors are increased. Imposing too few constraints on the randomization procedure that yields null model networks increases false positive discoveries, i.e. type I errors. The problem of balancing type I and type II errors is inevitable in null model analysis and statistical tests in general [207]. Thereby, it is difficult to assess whether a particular null model sufficiently captures the range of patterns specified or implied by the null hypothesis

[207]. Consequently, using an inappropriate null model in the statistical test might introduce a bias in the assignment of statistical significance [216, 217]. The null model widely employed for network analysis preserves the in-degree and out-degree sequence. To ways for constructing such null model networks are presented in the following section.

2.11.2 Random networks with prescribed degree sequence

Both degree sequences determine the topology of a directed network by imposing constraints on potential locations of edges connecting specific pairs of nodes and therefore it ultimately affects many of the network's properties. Thus, the in-degree and out-degree sequence is a basic and important attribute of a directed network which consequently should be accounted for in the generation of reasonable null model random networks [216, 210, 213]. Incorporation of the vertex degree sequences into the null model yields a statistical test for significant subnetwork counts (see also section 2.7.2) with a 'good' amount of restrictiveness so that not too many false positive results nor too many false negative results are expected. The associated random networks are usually either generated by the configuration model ('stubs-pairing') [155, 157, 218, 219, 220] or by a Markov chain Monte Carlo (MCMC) method ('edge-switching') [213, 101, 157, 214, 221].

The basic idea of the process of drawing networks from the configuration model ensemble is to consider the in-degrees and out-degrees of nodes as 'stubs' that have to be randomly paired by connecting them with directed edges. For it, a node is chosen uniformly and repeatedly from the set of nodes which have not used up all their outgoing edges and another one is randomly chosen from the set of nodes that can still accept ingoing edges, both nodes are linked by an directed edge. This procedure is performed until all nodes have acquired all of their connections as specified by their degree sequence. In practice, the configuration model approach suffers from the possible introduction of multiple edges and loops into the randomly constructed network, thereby creating multigraphs (multiple parallel edges in the same direction) or pseudographs (multigraph with loops), which provide an undesired model for functional brain networks, which do not exhibit these properties. Rejections of selected node pairs that would be required to avoid the addition of degenerating edges are unfortunately problematic on their own. A sampling bias would be introduced if following the rejection of a multi-edge an alternative node pair is chosen at random from the set of available nodes with free 'stubs' [219, 214]. This modified strategy would be equivalent to an extended exploration

of the search space in the neighborhood of non-simple partial directed networks, causing the final simple network to be drawn non-uniformly from all possible 'stub'-pairings. The algorithm might also be modified to reject partial networks upon introducing a degenerate edge. It then uniformly samples simple networks with prescribed degrees but its acceptance rate would be too small to apply it to real-world problems [219]. Nevertheless, numerical experiments have revealed that the modified configuration model algorithm that discards degenerating edges and instead selects a new node pair can be acceptable in practice despite of its sampling bias [157]. The non-Markov chain Monte Carlo method 'go with the winners' that also constructs networks with prescribed degree sequences generates statistically correct samples but is too inefficient for generating large null model network ensembles [157, 222].

The Markov chain Monte Carlo method (MCMC) 'edge-switching' algorithm rewires an input network by means of a series of random reconnections of edges. Repeatedly, two directed edges (v_k, v_m) and (v'_k, v'_m) are selected uniformly at random and their head nodes are exchanged to yield the edges (v_k, v'_m) and (v'_k, v_m). If this would generate multiple edges or loops, the edge switch is not performed and a new pair of edges is randomly selected. These edge switching attempts are repeated $\Omega |\mathcal{E}|$ times, where Ω is a ('mixing') parameter which has to be chosen large enough to allow the underlying Markov chain to converge to its stationary distribution. In the course of this, rejected edge switches, which correspond to the transition from a network to itself, are also counted. For the selection of Ω values and the number of randomized network instances I refer to section 3.1.3. With respect to functional brain networks the resulting randomization disintegrates network structure with functional relevance that represents the recorded neural processing. The MCMC edge-switching algorithm is recommended for network motif detection as its application yields a good trade-off between speed and accuracy (uniform sampling of random graphs) [157].

2.12 Surrogate-assisted network analysis

To assess if observed network characteristics are relevant and reflect true functionally meaningful properties of the underlying network topology or whether they are caused by random processes and simple mechanistic effects that stem mainly from very basic topological properties, like the degree sequence or edge density, surrogate-assisted analysis [223, 224, 132, 209] has to be performed. Central to surrogate-assisted analysis approaches is the utilization of topological information obtained from appropriate null model network ensembles (see also section 2.11.1).

Essentially, surrogate-assisted analysis looks whether null model networks can generate the same patterns and characteristics that were observed in the empirical networks. After generation of a large sample of hundreds or thousands of null model adjacency matrices, the patterns that each of them contains have to be quantified with the network characteristics of interest. The obtained null model network characteristics can be used in several ways. A common approach is to normalize the values of real-world network characteristics relative to the null model ensemble, i. e. by dividing the real-world network characteristics with their mean value in the null model ensemble [224, 225]. Obviously, this ratio is larger than one if a topological characteristic is larger and more pronounced in the real-world networks as compared to its surrogate counterparts. Another approach consists in obtaining the distribution of the network characteristic under the null model to calculate p-values for assessing statistical significance (see also section 3.1.3).

Surrogate-assisted analysis might be influenced by network size effects. In particular, simple randomization schemes for large networks potentially yield null model network ensembles whose topology will almost always be deviating significantly from the associated real-world networks, causing the underlying null hypothesis to be rejected regardless of the data [207].

Part II

Contributions

Chapter 3

Contributions to network science methods and the multi-level analysis of functional brain networks

FOR the purposes of my research on the analysis of functional network data I developed a network toolbox in the form of several R packages and additional Matlab code. In this chapter I present my methodological contributions to the multi-level analysis of functional brain networks. They are roughly illustrated in figure 1.1, left column, and consist of (1) development of analysis concepts by combining state-of-the-art methods, (2) complementation of these concepts by own application-tailored approaches, (3) conduction of the network analysis parts of application studies and (4) visualization of the results. To provide additional context for the topics of this chapter, I begin with an overview that relates my methodological developments to the scientific problems and the processing steps of the application studies, which are then presented in the following chapters.

Section 3.1, "*Motif detection in samples of binary directed networks with pairwise different node labels*": The presented method is an extension of classical network motif analysis. It accounts for the case of network samples and it includes pairwise different node labels into the definition of a motif. This method was applied in a study of functional brain connectivity in patients with major depression vs. a healthy control group during neural processing of

painful electrical stimuli, where characteristic patterns of functional interactions were to be identified and compared among groups. For this data my analytical approach for assigning subnetwork significance enables a fast discrimination between functional networks of both groups, whereas my network randomization-based approach enables a more precise resolution of sample-specific network motifs. The involved processing steps are illustrated in figure 3.1a and the corresponding methodological application study is presented in chapter 4 on page 84.

Section 3.2, "*Cost optimal matching of module affiliation labels and a fuzzy matching extension*": Module structure analyses in network samples leads to the problem of matching module affilition labels of nodes among different network partitions, so that similar modules will have the same assigned labels (identities). The presented algorithms solve this problem by exploiting module similarity to perform cost-optimal assignments. They were used in all studies in which module structure was identified. They are presented in chapters 6 and 7.

A concept that combines state-of-the-art techniques for the network sample-specific analysis of weighted edge-complete networks was devised. It entails quantification of topological characteristics with relevance for the clinical research question (weighted clustering coefficients, average node strength, characteristic path length; see also sections 2.7.1 and 2.7.3) and surrogate-assisted statistical analysis of the results, which is often not reported in the literature on network analysis (see also section 2.12). This network analysis concept was used to investigate whether an expected reconfiguration of functional brain network topology in response to medical treatment (lithium therapy) took place in HIV-infected subjects with diagnosed cognitive impairment. The application study is presented in chapter 5 and the corresponding methodological overview is shown in figure 3.1b.

Section 3.3, "*Comprehensive analysis of module structure quality*": A comprehensive module structure analysis concept was developed that goes beyond the typical analysis schemes reported in the literature. I combined several different popular network module identification algorithms (e. g. algorithm of Blondel *et al.* , Infomap, Potts spin glass-based algorithm, leading eigenvector algorithms) for unweighted networks with many different kinds of measures of module structure quality, similarity and preservation (e. g. modularity, variation of information, intra-module edge pattern edit distance, Rand index, 'split-join' distance). This analysis framework was assembled for the purpose of evaluating the effects of the new large scale
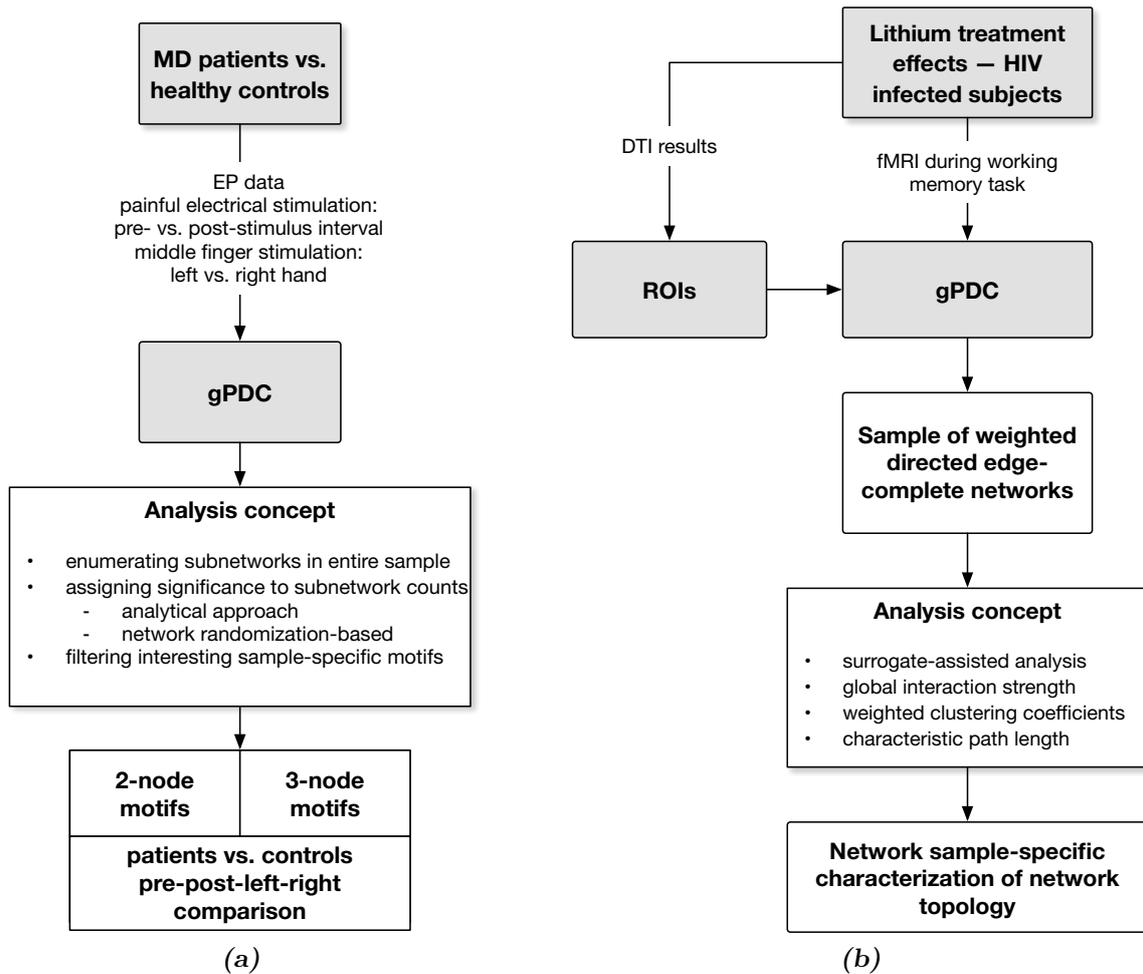
Figure 3.1 – **Flow chart of the data processing steps and results for two application studies.** Data acquisition and connectivity analysis (gray rectangles) were performed by others. (a) Motif detection in network samples with node labels (chapter 4). (b) Characterization of network sample topology (chapter 5).

Healthy athletic subjects

EEG data
compensation of a sudden
balance perturbation

tvMVAR dDTF

Sequence of weighted directed edge-complete networks

time grid definition

Global threshold computation based on pooling network sample information

Individual Pareto-optimal thresholds

Consensus network structure

Soft thresholding

Soft thresholding

Weighted network module detection

Weighted network module detection

Consensus clustering

Module matching

Visualization of tracking results

(b)

Simulation study I and II
lsGCI vs. GCI networks with simulated module structure

Analysis concept

- several different module structure detection algorithms
- quantification of module structure quality and preservation
- assessment of network topology alterations caused by lsGCI

1 healthy subject	**5 HIV-infected subjects**
Resting state fMRI data	

Multiple thresholding and dichotomization

Module structure detection

- spatially high-dimensional lsGCI functional networks
- module structure projection onto fMRI voxel masks

(a)

Figure 3.2 – **Flow chart of the data processing steps and results for two application studies.** Data acquisition and connectivity analysis (gray rectangles) were performed by others. (a) Module structure detection and quality analysis in lsGCI functional brain networks (chapter 6). (b) Dynamic module structure analysis (chapter 7).

Granger (lsGCI) approach (see also appendix B.1.1) on network edge patterns. This approach enables the computation of spatially very highly resolved functional interactions at the cost of an information loss with respect to the underlying interaction patterns. Thereby, the general problem of quantifying the effects of edge pattern alterations was shifted to the quantification of module structure quality and preservation. For this evaluation an extensive simulation study using ground truth networks with known module structure was performed. Within this framework, it was my task to design the needed network simulation tools, which are presented in section 3.4. Furthermore, module structure and the associated functional segmentation was identified in spatially highly resolved functional brain networks that were obtained from resting state[1] fMRI recordings using the lsGCI approach. The flow chart for the respective study is shown in figure 3.2a and the study is presented in chapter 6 on page 103.

Section 3.4, "*Simulating networks with pre-defined module structure*": The presented algorithm simulates networks with ground truth module structure. The resulting networks possess a constrained topology so that they can be used for the simulation of MVAR time series (see also appendix B.2). The parametrization of the simulation allows to tune several quality characteristics of the module structure. Therefore, the algorithm could also be used to generate benchmark networks for testing module structure identification algorithms. The algorithm was used for the study presented in chapter 6 on page 103.

Section 3.5, "*Computing edge weight thresholds using a multi-objective optimization approach*": Defining optimal edge weight thresholds to filter potentially spurious interactions and to binarize networks is still an open research problem. The presented approach allows to compute thresholds so that the resulting non-complete binary or weighted networks exhibit their most strongly marked module structure. The algorithm was applied in the application study presented in chapter 7 on page 129.

Section 3.6, "*Extraction and visualization of time evolving module stucture*": For successful dynamic module structure extraction several available methods were combined with my own approaches. The clinical background of the corresponding application study were EEG recordings of subjects during a single-leg balancing task on a initially stable surface that abruptly becomes unstable. From these EEG recordings time-variant connectivity networks

---

[1]Studies have shown that, unexpectedly, spontaneous brain activity at rest is not idle or unstructured, e. g. [226, 227]. Instead, the activity of functionally related brain regions at rest is organized and correlated.

were computed using time-variant dDTF (see also appendix B.1.3). The execution of cognitive tasks is associated with time-varying changes of the functional connectivity structure. Thus, the working hypothesis of this study is that dynamic module structure in functional brain networks reflects the neural processing during optimization of balance control. The data analysis concept entails my method for network thresholding (section 3.5), module structure detection in weighted networks, my cost optimal module label matching approach (section 3.2) and visualization of the results using Sankey diagrams that illustrate the 'flow' of nodes through the modules. The network analysis concept was extensively evaluated and continually modified to cope with the noisy nature of the data and to reveal their underlying patterns. The concept-related flow chart is represented in figure 3.2b. The application study is presented in chapter 7.

In the following sections I describe my network analysis methods in more detail.

3.1 Motif detection in samples of binary directed networks with pairwise different node labels

Analysis of complex systems is most commonly performed on single network instances of the investigated system, which can be problematic as isolated instances are usually not representative of the system itself. To take account of this situation I present an extension of the standard approach for network motif detection in single, directed networks without node labeling to the case of a sample of directed networks with pairwise different node labels. A network sample is obtained by pooling functional brain networks of a group of related subjects, e.g. a group of patients with a similar manifestation of a neural disease and similar medical treatment regime, or a control group (cf. illustration of processing steps in figure 3.1a). All sample element networks consist of the same set of nodes. The involved statistical tests are adjusted to assign significance to subnetwork counts derived from the entire network sample. In addition, my methodological extension does not discard but rather preserve the functionally important topological information associated with the node labels. This network decomposition procedure can be seen as a way to filter intricate network topology by omitting less important substructures with the resulting motifs being a distinguishing characteristic to reveal sample-specific differences in network topology.

Dealing with samples of directed networks with identical pairwise different node labeling, given by e. g. EEG recording sites, instead of single networks without such labeling imposes certain constraints on network motif detection. Given a pairwise different node labeling, two subnetworks are identical if and only if they consist of the same subset of labeled nodes and if they share the exact same set of directed edges (with labeled tail and head nodes) as opposed to only sharing their patterns of interconnections. As a consequence, the problem of determining graph isomorphism for subnetworks does not have to be addressed and topological equivalence classes of isomorphic subnetworks do not exist. The situation is completely different for networks without node labeling, where equivalence classes for isomorphic subnetworks exist (see also section 2.7.2). The existence of node labeling also means each subnetwork can occur at most only once in a single sample element network, which affects the statistical analysis of subnetwork occurrences. It is not possible to assign statistical significance to subnetwork counts in a single network or in very small samples of networks. Consequently, motif detection in a sufficiently large sample of networks constitutes not only a novel approach to reveal common topological characteristics of all sample element networks but is also a necessity.

In order to deal with network samples and to incorporate the information of the node labeling, i. e. the recording sites, into the network decompostion approach one has to extend the original notion of network motifs [101, 109] to define the special case of network motifs of a sample of directed networks with pairwise different node labeling. In this context, network motifs are defined as small connected subnetworks which differ in their set of edges, as opposed to differ in their patterns of interconnections only, which appear in their sample of networks significantly more often than in random networks according to a suitable random graph model. In this way the node labeling is taken into account that yields not only an advantage with respect to the computational complexity of motif detection but also has the important advantage of conserving the positional information of motifs in the network. This positional information is associated with underlying neural processes and is therefore important for a subsequent functional interpretation of the results. If node labeling is discarded or not available, as in standard motif detection, then the only information obtained from motifs in the absence of localization is about the existence and the type of significant patterns of directed influences between nodes.

Detecting sample-specific and node labeled network motifs encompasses the following steps.

1. Exhaustive enumeration of subnetworks of a certain size in all sample element networks.

2. Identifying subnetworks that have a statistically significant occurrences in the network
   sample using either a fast analytical approach or an approach based on extensive and
   computationally expensive null model network simulations. The analytical approach
   save considerable computation time since no random network ensembles have to be
   generated and no subnetwork counts have to be obtained from them. The alternative
   approach on the other hand yields a more realistic null model than the analytic one.

The following descriptions of the motif detection steps are taken from my publications on
sample-specific network motifs [97, 156].

3.1.1   Exhaustive enumeration of subnetworks

Let $\mathcal{D} = (\mathcal{D}_1, \ldots, \mathcal{D}_n)$ be a sample of n node-labeled directed networks $\mathcal{D}_k = (\mathcal{V}, \mathcal{E}_k)$ all
having the same set of nodes $\mathcal{V}$ with pairwise different labels and a specific set $\mathcal{E}_k$ of directed
edges. $\mathcal{A}^k$ denotes the adjacency matrix that represents network $\mathcal{D}_k$. The first step in our
approach is to explicitly enumerate all subnetworks $\mathcal{S}$ of a certain size $N_{\mathcal{S}} \geq 2$ in every network
$\mathcal{D}_k$. Thereby, for each sample network $\mathcal{D}_k$ every combination of $N_{\mathcal{S}}$ nodes is investigated to
determine whether a connected subnetwork is induced by it in $\mathcal{D}_k$. This exhaustive enumeration
technique is computationally tractable for small to medium sized networks, e.g. the ones
encountered in the computational neuroscience setting of EEG-derived network analysis, where
128-channel caps or 256-channel caps for recording the EEG represent typical upper bounds.
EEG-derived functional brain networks are usually much smaller than these numbers of nodes.
Alternatively, for larger networks and lower edge densities, subnetwork sampling algorithms
can be used instead of exhaustive enumeration (see also section 2.7.2). Subsequently, the
number of occurrences of each induced subnetwork over the entire sample is counted. Based
on these subnetwork counts statistical significance is assigned to subnetworks.

3.1.2   Analytical approach to determine subnetwork significance

To identify subnetworks with statistically significant counts in an investigated network sample,
the counts have to be contrasted with the counts obtained in corresponding random networks,
i.e. in absence of any functional constraints on the network topology. For it, a null model has
to be specified, which entails what 'randomness' with respect to edge patterns exactly means.
For a sufficiently large sample of node-labeled binary directed networks a suitable null model
can be derived as follows.

Let $0 \leq |\mathcal{E}_k| \leq N(N-1)$ be the number of edges of the directed network $\mathcal{D}_k$ and let

$$q = \frac{1}{n} \cdot \frac{1}{N(N-1)} \sum_{k=1}^{n} |\mathcal{E}_k| \tag{3.1}$$

be the normalized mean number of edges of the sample $\mathcal{D}$. Then, the i.i.d. variables $\mathcal{A}_{ij}^0$, $1 \leq i \neq j \leq N$, with

$$P\left(\mathcal{A}_{ij}^0 = 1\right) = q, \tag{3.2}$$

$$P\left(\mathcal{A}_{ij}^0 = 0\right) = 1 - q, \tag{3.3}$$

describe a random network $\mathcal{D}^0 = (\mathcal{V}, \mathcal{E}^0)$ with a mean number of edges $qN(N-1)$. It provides the basis of the null model. Let $\mathcal{S}$ be an arbitrary subnetwork with at least $N_{\mathcal{S}} \geq 2$ nodes of the set $\mathcal{V}$ and η_S edges. Obviously, the subnetwork $\mathcal{S}$ can exhibit at most $\eta_{S_{max}} = N_{\mathcal{S}}(N_{\mathcal{S}} - 1)$ edges. The count of $\mathcal{S}$ in the sample $\mathcal{D}$ is considered, where it can occur at most once in a sample network $\mathcal{D}_k$. Since all sample networks $\mathcal{D}_k$ are associated with the same null model, the count of subnetwork $\mathcal{S}$ in the sample $\mathcal{D}$ is binomially distributed under the null model with $B\left(n, q^{\eta_S} \cdot (1-q)^{\eta_{S_{max}} - \eta_S}\right)$. Finally, all subnetworks of a certain size $N_{\mathcal{S}}$ are tested with respect to a significant overrepresentation in the sample. Thus, an alpha-adjustment has to be applied. In the present study, generally the Bonferroni–Holm correction [59] with a multiple significance level of $\alpha = 0.05$ was adopted for all multiple test procedures to conservatively control the familywise error rate for all hypotheses at α in the strong sense instead of controlling the expected proportion of incorrectly rejected null hypotheses (false discovery rate [60]).

The time complexity of the associated computations is comparably low, since no simulations of large ensembles of random networks are involved and consequently no subnetwork counts have to be obtained from them. As a downside, the null model that is used for contrasting the network data is somewhat simple. This is the starting point of another approach that, in return for greatly increased computational costs, offers a much more realistic null model for the investigated network data. It is presented in the following.

3.1.3 Network randomization-based approach to determine subnetwork significance

Similar to classic motif detection in single networks the Markov chain Monte Carlo (MCMC) network randomization approach that preserves the in-degree and out-degree sequences of an input network is applied to each network of the investigated network sample (see also sections 2.11.2 and 2.7.2). Determination of suitable randomization parameter values often seems to be overlooked in the literature on motif detection, where the choice of parameters is usually not specified or not justified [156]. Contrary to this, the presented randomization-based approach includes a parameter determination step [156]. In particular, no a-priori bound exists for the mixing time of the underlying Markov chain [212]. The 'mixing' parameter Ω, which determines the number of edge swap attempts as $\Omega \cdot |\mathcal{E}|$, has to be chosen to ensure sufficiently uniform generation (sampling) of every directed network with prescribed in-degree and out-degree sequences. Often values of Ω around 100 are described in the literature. Also, the choice of the size of the random network ensemble, given by the number of random realizations b^* for every sample element network, is important. On one hand it should be as small as possible to save computational resources, and on the other hand the random network ensemble has to be large enough to ensure that the distribution of relative subnetwork frequencies in these random networks is likely to differ only within sufficiently small bounds from a distribution obtained by generating a larger number of realizations. Often just 100 or 1000 random network realizations were reported, but sometimes up to 10000 random networks were used, but not much justification was given for any choice [156]. Here, both quantities are determined more accurately with regard to the investigated network sample data.

For large network samples it is unfeasible from a computational cost point of view to determine Ω for each sample element network separately. Thus, one sample element network is identified that is representative of all other analyzed networks of its sample with respect to the property preserved during randomization, namely the in-degree and out-degree sequence. One way to identify the representative network is to calculate the mean in-degree and out-degree sequence of the network sample and to select the network whose degree sequences have minimal distance to it, according to the maximum norm. With the representative network a large range of Ω values is analyzed with respect to their effect on uniform network sampling by the MCMC edge-switching algorithm. The idea is to use the exact chi-squared goodness-of-fit test to investigate the evidence against the null hypothesis of a uniform generation of networks for the preselected Ω values. For it, the number of all networks with the degree sequence of the

representative networks needs to be known. This number cannot be calculated analytically, hence, it has to be obtained by simulations. Due to the time complexity of the simulations and the large number of network instances with a given degree sequence, often only an approximation, i. e. a good lower bound, of the number of networks with the degree sequence of the representative network can be obtained. To obtain this lower bound and to perform the chi-squared goodness-of-fit test, a large number of randomized instances of the representative network have to be generated using the edge-switching algorithm with different values of Ω. The results of the independent simulations are then pooled together, since for different Ω the MCMC algorithm is likely to sample networks with prescribed degree sequences from different regions in the network configuration space. After detecting and removing all duplicate networks from the generated union set, a lower bound for the number of networks with the same degree sequences as the representative network is given by the cardinality of the resulting set of unique networks. The sampling distribution of the corresponding test statistic under H_0 is determined by means of Monte Carlo simulations. The smallest Ω for which the test statistic falls below the $(1 - \alpha)$-quantile is selected. Otherwise, if it exceeds the $(1 - \alpha)$-quantile for all Ω the edge-switching algorithm does not uniformly sample networks with the same given prescribed degree sequences for any Ω. In this case, one has to use the Ω for the randomization for which the test statistic is minimal. In addition to identifying the value of the 'mixing' parameter Ω, the number of random realizations b^* for all sample element networks required for a reliable detection of motifs is determined. This procedure starts by generating an upper bound B of bootstrap network samples using the MCMC edge-switching algorithm, where each bootstrap network sample consists of one randomized instance of each input sample element network. Based on these bootstrap network samples, a reference distribution R_B^i of relative subnetwork frequencies is calculated by enumeration of all interesting node-labeled directed subnetworks $\mathcal{S}_i$. The reference distribution is then compared to distributions R_b^i obtained in the same way from lower numbers b of bootstrap network samples. R_b^i is accepted to be sufficiently close to R_B^i if it holds $\delta_b = \max_i |R_b^i - R_B^i| < \varepsilon$ for an arbitrary fixed $\varepsilon > 0$. Finally, b^* is defined by $b^* = \min_b(\forall k \geq b : \delta_k < \varepsilon)$. By implication, if no such b^* exists the upper bound B of bootstrap network samples was chosen too low and more bootstrap networks have to be generated to augment the set of randomized networks and to increase B.

Statistical significance is assigned to subnetwork counts that have been obtained from the input network sample using the b^* random realizations of every sample element network's

degree sequences. Relative subnetwork frequencies in each of the b^* bootstrap network samples were already obtained in the preceding step for the computation of distributions R_b^i. Relative subnetwork frequencies are used to compute p-values for the corresponding subnetwork counts in the input network sample. As for the analytical approach presented in section 3.1.2 various subnetworks are tested with respect to a significant overrepresentation in the sample, which makes an alpha-adjustment necessary. Again, the conservative Bonferroni–Holm correction [59] with a multiple significance level of $\alpha = 0.05$ might be applied. Subnetworks with statistically significant counts in the input network sample are network sample-specific motifs with pairwise different node labels, i. e. locatable characeristic topological patterns of the network sample.

3.1.4 Filtering motifs

Depending on the number of different investigated network samples, their cardinality and the size of the networks they contain, a large number of subnetworks are enumerated, some of which are potential motifs. A simple data reduction technique is needed to reduce the number of subgraphs to be evaluated. For example, one might focus the further evaluation on 3-node node subnetworks that occur in at least one of the investigated network samples at least a given minimum number of times, because these subnetworks would represent the most promising candidates for network motifs. Typically, such a restriction would dramatically decrease the number of 3-node subnetworks from several thousands to just a few hundred that have to be further analyzed with respect to an overrepresentation in the input network samples by comparison with their counts or their probability in null model network ensembles [156].

It is also advised to filter identified motifs so that only 'interesting' motifs are finally interpreted. Filtering motifs might be achieved by considering only those motifs whose occurrence is either sample-specific or which occur often in most samples of functional connectivity networks.

3.2 Cost optimal matching of module affiliation labels and a fuzzy matching extension

A problem often encountered when working with network modules in different networks of a network sample consists in dealing with mismatched module labels. Due to the randomized nature of network module detection algorithms, different module labels might be assigned

to similar or identical modules across distinct networks, which complicates analyses and graphical representation of module structure and its changes. An example of such a situation is illustrated in figure 3.3. Module labels had to be matched in the application studies presented in chapters 6 and 7. The corresponding methodological overview of both studies is shown in figures 3.2a and 3.2b.

Following, I present three module label matching algorithms that are based on the classical two-dimensional (n:n)–assignment problem with quadratic cost matrix[2] [228] and the unbalanced assignment problem with rectangular cost matrix, where two sets of entities with different cardinality have to be matched in a cost optimal way [229, 230]. In both assignment problems an injective, one-to-one assignment of "workers" (i.e. modules in one network partition $\mathscr{P}_A$) to "tasks" (i.e. modules in another network partition $\mathscr{P}_B$) with minimal sum of assigned costs is sought. Thus, for the unbalanced assignment problem, the number of matchings is determined by the smaller set of either "workers" or "tasks". A schematic overview of assignment problems can be found in a publication I had co-authored [231].

Formally, this kind of considered assignment problem is given by a set W containing p elements, a set T containing q elements, with $p \leq q$ and a $p \times q$ cost matrix C where $C(i, j)$ is the cost of assigning element i of W to element j of T [229, 230]. An assignment can be interpreted as a permutation of a p-element subset of T, such that element i of W gets assigned the ith element of that permutation. To solve the assigment problem, the objective is to

---

[2]The classical assignment problem is mathematically identical to the weighted bipartite matching problem from graph theory.

(a)

(b)

(c)

(d)
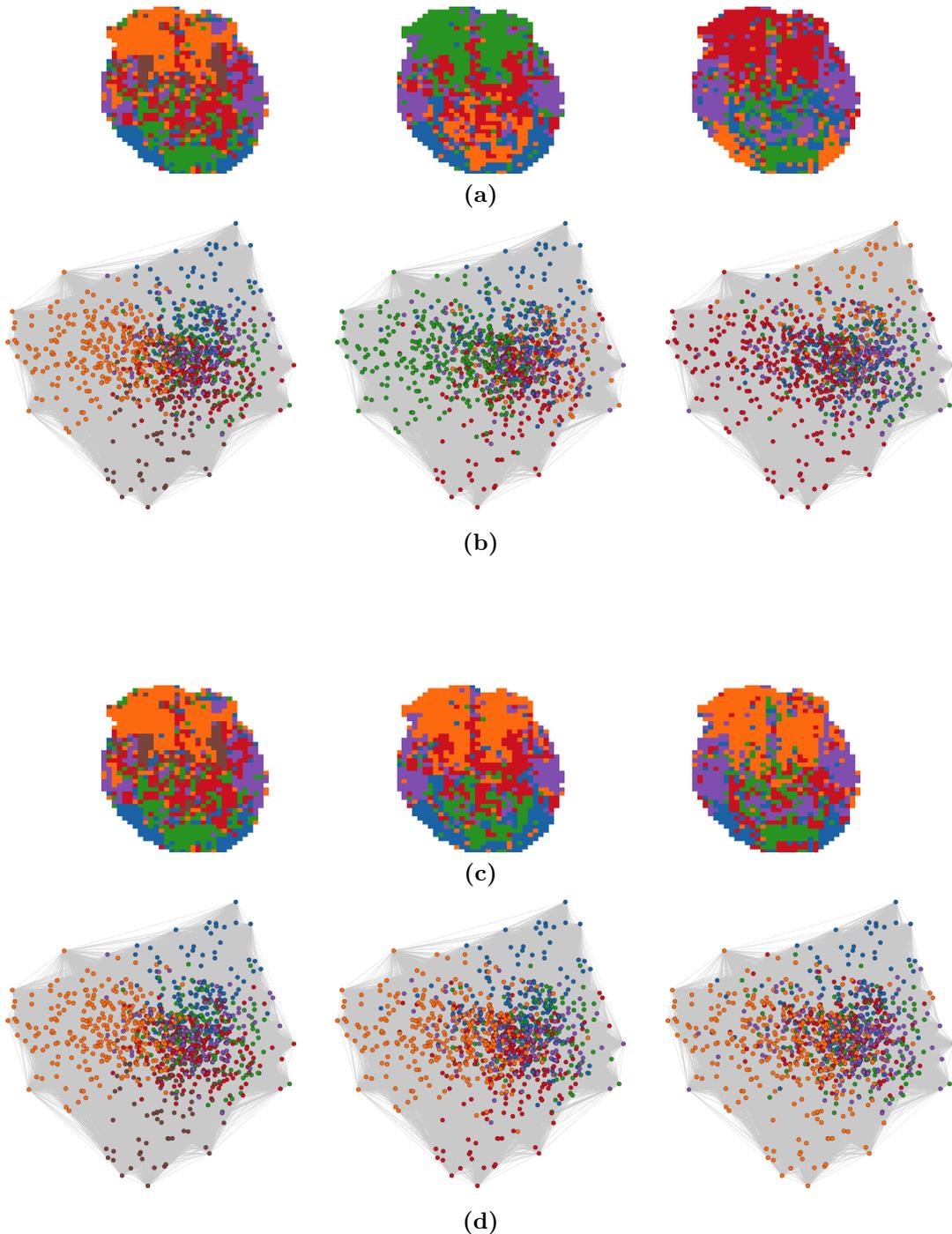
Figure 3.3 – **Example of the module label matching problem.** A typical situation is illustrated using three different fMRI-based networks of a network sample for which module structure was identified (see also chapter 6). (a) Unmatched, originally obtained module affiliations of nodes (voxels) are projected to their respective fMRI slice voxel mask. Modules with a large overlap in voxel locations have different assigned module labels (colors) among the three networks (frontal and occipital regions). For subsequent module structure analysis and to improve the visual representation a matching of module labels across the three networks has to be performed. With the methods described in this section, module label matching can be accomplished automatically with little computational costs even for large network samples. (b) Unmatched module affiliations in the corresponding network layouts. (c) Matched module affiliations shown using the voxel mask. The matching was performed using fuzzy matching with a threshold of 0.15. (d) Matched module affiliations shown in the corresponding network layouts.

$$\text{minimize} \sum_{i=1}^{p} \sum_{j=1}^{q} C(i,j)x_{ij} \tag{3.4}$$

subject to constraints

$$\sum_{i=1}^{p} x_{ij} \leq 1, \qquad (j = 1, \ldots, q) \tag{3.5}$$

$$\sum_{j=1}^{q} x_{ij} = 1, \qquad (i = 1, \ldots, p) \tag{3.6}$$

$$x_{ij} \in \{0, 1\}, \qquad (\forall\, i, j) \tag{3.7}$$

In equation (3.7) x_{ij} is an indicator variable for the assignment of element W_i to element T_j.
Each element of W is assigned to exactly one element of T (equ. 3.6). Each element of T is
assigned to at most one element of W, i.e. each "task" cannot be assigned to more than one
"worker" (equ. 3.5).

A function for solving these two kinds of assignment problems is part of my R packages
that constitute my network analysis toolbox. The resulting linear programs are solved using
an R interface to the popular mixed integer linear programming solver *lpSolve* [232].

I developed the following algorithms to compute cost optimal module label matchings across
networks of a network sample. As primary input parameter they accept a matrix of node
module affiliations for at least two different networks with the same number of nodes.

The first step in the procedures is selecting the network partition with the maximum
number of modules as a reference against which the module affiliations of all other networks
are being matched separately. Then all pairs of modules of the reference partition and each
non-reference partition are compared systematically to compute a cost matrix C for every
single reference partition—non-reference partition comparison. The entries of C contain the
costs of matching all pairwise combinations of reference network modules and the modules of
the respective non-reference partition. In detail, this cost matrix C is defined as the difference
between the number of network nodes and the cardinality of the intersection of member nodes

of each considered module pair[3]. Consequently, the costs for matching a reference network module label to a pair of reference—non-reference network modules are decreasing with an increasingly larger overlap of module member nodes. Using this cost matrix, the assignment problem is solved. According to the solution obtained, module labels of each non-reference network partition are changed to the assigned module labels of the reference network partition.

Despite the cost optimal nature of this module label matching algorithm, in practice situations can and do arise in which the used module detection algorithm returns conflicting module labels for different networks that cannot be plausibly matched. After making good low-cost assignments, the matching algorithm ultimately has to assign the remaining non-reference modules to the remaining modules of the reference partition even though their member node sets are very different. A further complication arises if the number of modules differs between the reference partition and a non-reference partition or between different non-reference partitions. Since in this situation not all modules can be matched, modules of different partitions automatically appear to be merged, splitted or non-existent and the resulting module label matching across such a network sample cannot be unequivocal. Since nodes in functional brain networks are associated with unique identity labels (e.g. EEG electrode identifiers or voxel positions), module structure of several networks should be interpreted with respect to the spatial information of nodes and their location in the network, while the colors associated to the modules are acting merely as a guide. By way of illustration, figure 3.4 (a) shows an example situation where module labels in three different network partitions are difficult to match. In this figure the module label matching results of the cost optimal matching algorithm described above is shown in figure 3.4 (b). Panel (c) shows the results for an improved version of the label matching algorithm, which is described below.

Two further improvements of the cost optimal module label matching algorithm were implemented. The first improvement concerns the handling of low-quality assignments. In a sort of 'fuzzy' matching extension, a non-reference network module is compared to the reference network module to which it is assigned by the standard approach. If their similarity is below a (user-defined) matching threshold θ, the label of the corresponding reference network module is not assigned and instead, under the assumption of the novelty of the module and its composition of nodes, a new label (and thus a new color), which is not already in use in the

---

[3]An equally good choice for the cost function would be the Jaccard distance (section 2.10 on page 40), which can be seen as a normalized version of the cost function described in the text (cf. [172]).
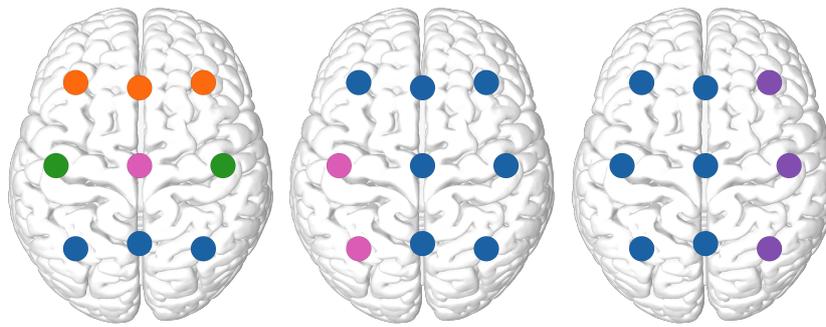
(a) Original node module affiliations.
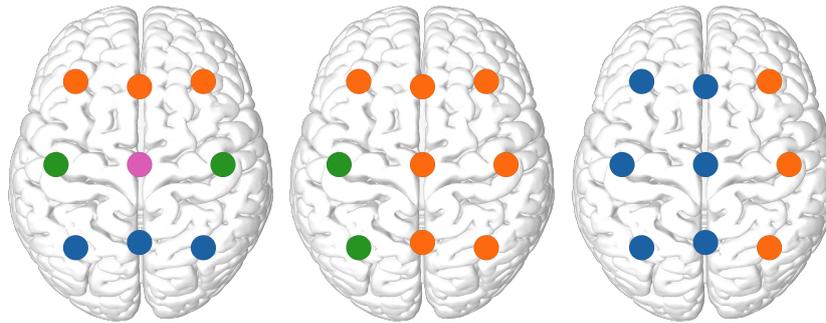
(b) Cost-optimized matching of module labels.
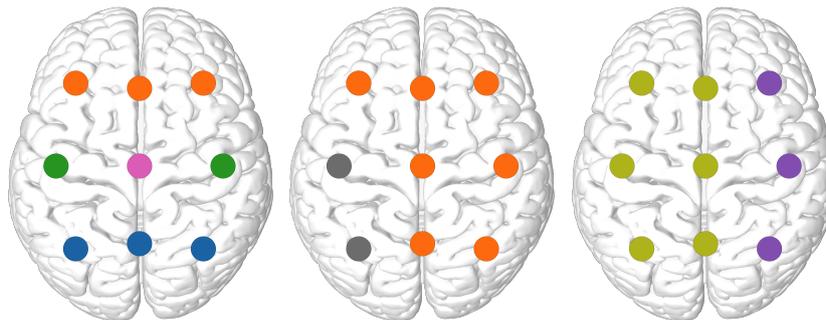
(c) Cost-optimized label matching subject to module similarity constraints.

Figure 3.4 – **Module label matchings obtained by two different algorithms.** Three different network instances of the same size ($|\mathcal{V}| = 9$, presented as EEG electrodes) with very different, constructed module affiliations are shown. The first network, which was partitioned into four modules, was automatically selected as the reference for matching module labels. The module affiliations of the reference partition are not changed. (a) The conflicting original, unmatched module labels, as they could have been returned by a module detection algorithm. (b) Results of the cost optimal label matching algorithm that solves the assignment problem and strictly applies the resulting label matches. For the 2nd network, the ● module is correctly changed to the ● one, since for this choice the overlap of member nodes (3 nodes) is the largest (the overlap with ● is only 2 nodes). The label matching for the 3rd network is also algorithmically correct, but it is ambiguous and implausible. The original ● module could have been equally well matched to the ● reference module and the original ● module to either the ● or ● reference module. (c) Results of the "fuzzy" label matching algorithm that solves the assignment problem, but actually assigns reference module labels to non-reference network modules only if the matched module pairs have at least a minimum user-specified similarity to each other. Implausible, low-similarity matchings of spatially distinct modules are *not* performed and instead new module labels are assigned. Consequently, the original ● module of the 2nd network becomes ● and the original ● module of the 3rd network becomes ●. Although being dissimilar to all reference modules, the original ● module of the 3rd network does not get assigned a new identity, because this module label is already unique among all three networks.

network sample, is allocated. The new color might be selected in a way that similar modules, e. g. two related modules that occured as the result of a splitting event of a larger module, are assigned similar colors, i. e. by adjusting the hue or opacity. If on the other hand their similarity exceeds the matching threshold θ, the non-reference network module gets assigned the label of the matching reference network module, as before. The similarity of a pair of modules and their member nodes might be measured in terms of the relative overlap of module member nodes [233]. This is calculated using the Jaccard coefficient (section 2.10 on page 40) which can be reformulated for dealing with sets as

$$J(\mathcal{M}_i, \mathcal{M}_k) = \frac{|\mathcal{V}_{\mathcal{M}_i} \cap \mathcal{V}_{\mathcal{M}_k}|}{|\mathcal{V}_{\mathcal{M}_i} \cup \mathcal{V}_{\mathcal{M}_k}|} \tag{3.8}$$

Clearly, the parameter θ has a big influence on the results of this matching strategy, as a trade-off between eliminating poorly fitting matches and an excessively colored, diffuse module structure has to be made, in particular if large network samples are considered. Illustrative module label matching results of the improved cost optimal matching algorithm are shown in figure 3.4 (c).

Another improvement of the module label matching strategy is concerned with an additional matching of the low quality assignments that were identified by the 'fuzzy' matching approach. As explained above, these low quality assignments were too dissimilar to any reference network module and as a result they were not actually matched to any reference module, but were assigned a completely new label instead. Since modules of non-reference network partitions are matched separately to the module structure of the reference partition, it is possible that modules with very similar member node composition got assigned different unique labels across the network sample. Thus, the problem that the second improvement solves consists in a re-matching of these unmatched modules (that got new labels) across the non-reference partitions by pairwise comparisons of module similarity, using equ. 3.8 as before, in an attempt to further reduce the number of unique labels in the network sample, which improves subsequent module structure analyses.

The presented module label matching algorithms have internal logic for dealing with certain special cases that might occur, for example module labels are not consecutive integers starting

with 1 (i.e. input module labels are allowed to have gaps), all input module affiliations consist of the same module label, or each network's community affiliation vector has only one module label. The fuzzy matching algorithms also contains code for intelligently managing the reuse of already existing module labels for non-matchable modules in special cases. Further improvements of the module label matchings might be obtained by a more sophisticated strategy for selecting the reference network partition, as the choice of the reference partition plays a role for generating plausible module label assignments across different networks.

3.3   Comprehensive analysis of module structure quality

The development of a data analysis concept that is specific to the context of the data is in many cases not possible without the use and integration of established and time-proven concepts and methods, i.e. a sophisticated application-tailored combination and modification of established methods often creates innovations. Following this line of thought, I extend the frequently used somewhat simple analyses of identified module structure quality and similarity by combining a vast number of measures of different scope: module quality and resemblance measures based on exploiting features of network topology in association with measures comparing the module membership assignments of nodes and information theoretic measures (e.g. modularity, variation of information, intra-module edge pattern edit distance, Rand index, 'split-join' distance). This concept enables the comparative analysis of network module structure quality and preservation. This strategy proved to be useful for evaluating a new Granger Causality approach of my institute (lsGCI, see also appendix B.1.1) that makes it possible to compute directed interactions in spatially very high-resolved fMRI brain recordings. The network-based analysis of the effects of lsGCI on connectivity patterns and the resulting recoverability of module structure is the main topic of chapter 6 on page 103. The selected measures, which were used in combination for pairwise comparisons of network partitions of ground truth networks and corresponding lsGCI networks are listed and described in section 2.10. The flow chart for the corresponding application studies is shown in figure 3.2a.

3.4   Simulating networks with pre-defined module structure

The following algorithm simulates unsymmetrical binary random adjacency matrices with edge patterns that link nodes to form a pre-defined non-overlapping module structure, where the module structure quality and definiteness is adjusted by several parameter settings. The

corresponding edge-directed ground truth networks can be used as benchmark to evaluate the performance of community detection algorithms or to control whether network visualization methods are able to separate single modules in the layout of the network structure. In one application study (chapter 6, cf. overview in figure 3.2a) simulated ground truth networks with pre-defined module structure were used for the analysis of the adverse effects the lsGCI approach (see also appendix B.1.1) has on network topology and how the associated alterations of edge patterns impact the recoverability of module structure. The data basis for this analysis was a set of realized first order MVAR processes that were simulated on the basis of constructed ground truth networks (see also appendix B.2).

Each edge-directed ground truth network consists of a specified number of N nodes, where N is a multiple of 25. The generated networks can be partitioned naturally into non-overlapping modules restricted to consist between 10 to 15 nodes each. The number of modules scales with the network size, so that for every increase of 25 nodes two additional modules are build into the network topology. Thereby, the size of each module is chosen randomly, such that the sum of all module sizes equals N. Nodes are selected and accordingly assigned membership to the modules. Pairs of nodes are then linked by edges placed uniformly at random under constraints that define the module structure, using separate probabilities for intra-module edges ("internal", p_{int}) and inter-module edges ("external", p_{ext}) that have to be specified as parameters of the algorithm. Clearly, the larger the ratio $p_{int}/p_{ext} > 1$, the more pronounced is the network module structure. Optionally, the algorithm allows to specify global restrictions on the column sums (in-degrees) of the generated adjacency matrix. Incorporating this constraint in the network simulation might become necessary due to technical reasons in the case of using the generated adjacency matrices for MVAR time series simulation as described in B.2. For each node the network size determines its number of possible inter-module interaction partners. Thus, if the column sums are constrained, the probability of directed inter-module edges p_{ext} depends on the network size so that the constraint on the column sums holds true: To account for outlier nodes that have an above average number of intra-module and inter-module edges that would violate the column sum constraint, each node is allowed to have, on average, three edges to and from nodes of different modules. Consequently, the algorithm sets the probability of directed inter-module edges to $p_{ext} = 3/(N-15)$. Outlier nodes are likely to occur in synthetic networks with large N, and networks of size $N >= 500$ can already be difficult to generate otherwise. Moreover, additional conditions on *minimum*

intra-module ($\kappa_{int}^{in}, \kappa_{int}^{out}$) and *maximum* inter-module ($\kappa_{ext}^{in}, \kappa_{ext}^{out}$) in- and out-degrees have to hold true so that the resulting module structure is more clear-cut, while the modules remain connected. By means of specifying these four parameters, nodes are bound to maintain a given minimum number of edges to and from member nodes of their own module, while their number of interactions with nodes of other modules must not exceed a maximum number of edges in each of both directions. Clearly, the constraints on internal and external degrees do not entail a determination of specific connection patterns between nodes. To make the computations involved much more tractable, the network simulation problem is being split into simulating intra-module edge patterns that satisfy the constraints on intra-module node degrees separately from the simulation of eligible inter-module edges. In each of the two parts of the simulation, edge patterns are constructed a module at a time. Only if the constraints on the node degrees hold, the algorithm proceeds with generating the next module's intra-module or, respectively, inter-module edge patterns. Otherwise, the algorithm keeps on restarting the simulation of the current module's edge patterns until an upper bound of iterations (e.g. $10,000$) is reached and the algorithm terminates without success. If the simulation of intra-module and inter-module edges eventually succeeds, their corresponding adjacency matrices are combined to yield a ground truth network as the algorithm's output.

The computational costs of the simulation increase strongly with network size N, as the aforementioned constraints, which strongly limit the space of allowed network configurations, become increasingly harder to satisfy. In fact, during the simulation a large number of attempts to generate a network under the given topological constraints have to be discarded and the process has to be started all over again. Finally, the adjacency matrices of both separate simulations are then combined to yield a network with ground truth module structure. If potential additional contraints on column sums are satisfied, the algorithm terminates and returns the network. Unfortunately, in many cases the deployment of long compute times to construct a network having a correct degree sequence is in vain and the process has to be repeated, as the network will have (a few) columns for which the column sum constraint will be unsatisfied. In particular for large values of N, I found that setting all input paramter values of the algorithm is not straightforward, as their values have to be balanced in a delicate way to make the conjunction of all topological constraints and the column sum constraint satisfiable.

It is straightforward to extend the algorithm and to devise a second variant of this algorithm that slightly varies the parameters p_{int}, p_{ext}, κ_{int}^{in}, κ_{int}^{out}, κ_{ext}^{in} and κ_{ext}^{out} to yield a

greater diversification of the resulting networks. In detail, the user-specified probabilities for internal and external edges might be changed by drawing from normal distributions, e.g. $p_{int} = \mathcal{N}(p_{int}, 0.03^2)$ and $p_{ext} = \mathcal{N}(p_{ext}, 0.001^2)$, respectively. Also, similar to LFR benchmark graphs [234] a power law distribution could be used instead. Similarly, the user-specified minimum intra-module and maximum inter-module node degrees are changed by drawing from normal distributions with the mean being the user-specified parameter value and a certain variance, e.g. 0.45^2. Of course, the empirically-derived parameterization of these probability distributions doesn't have to be fixed, as the variances constitute additional input parameters of the simulation algorithm.

Depending on the size of the generated networks, the satisfaction of the maximum column sum constraint for all nodes, as required for the autoregressive parameters of the MVAR time series simulation (see also B.2 on page 150), requires careful fine-tuning of all other network simulation parameters to a narrow range of combined values, which results in ground truth networks with a topology that potentially makes uncovering their module structure challenging. Example ground truth networks are shown in figures 6.2a and 6.7a.

As already explained, the presented algorithm was specifically designed to account for the requirements of MVAR time series simulation while still incorporating many parameters that define the notion of module structure, i.e. it is possible to tune the resulting module structure in many different ways. Therefore, it is similarly powerful as the LFR benchmark and more powerful than the classic and commonly used planted l-partition model [235], which constructs l modules of identical size using a fixed probability for intra-module edges and a fixed probability for inter-module edges.

3.5 Computing edge weight thresholds using a multi-objective optimization approach

Thresholding edges of a network is a common processing step in the network analysis pipeline to remove spurious interactions, to yield edge non-complete networks and to dichotomize weighted networks for the application of binary network analysis techniques. The difficulties and challenges of defining edge weight thresholds were already outlined in section 2.3 on page 16. Here I describe a novel approach that I developed for the analysis of module structure in samples of time-evolving EEG-derived functional networks (cf. overview in figure 3.2b), for

which statistical significance tests (see also section 2.3.1) are not computationally feasible and determining a global edge weight threshold for all networks turned out to be too inaccurate for individual networks so that no clear and stable module structure could be uncovered. While the following descriptions focus on the optimization of binary network module structure, the presented approach can be used in an analog fashion for obtaining optimal thresholds with respect to the inherent module structure of weighted edge non-complete networks.

The novel approach takes a weighted input network and computes a global network-specific edge weight threshold for which the resulting binary (or weighted) edge non-complete network has Pareto optimal module structure. A specified range of edge weight percentiles is considered and the corresponding edge weights are computed [236]. They serve as a threshold to dichotomize the weighted input network. The calculations described in the following are only performed for those resulting binary networks that are at least weakly connected. Several topological characteristics of the resulting networks are evaluated. These networks characteristics mainly give information on the network's module structure quality (cf. 3.3 on page 71). In particular, the modularity measure, performance measure, coverage, maximum Page Rank and small-world-ness are computed. For it, the maximum modularity partition is found by comparing the results of several module detection algorithms (walkTrap BU, infoMap BD, fastGreedy BU, Louvain BD and leadingEigenvector BD, cf. 2.8 on page 31). This analysis is repeated for each specified edge weight percentile yielding vectors containing the values of each of the five considered network characteristics. Thereby, each network characteristic represents an independent objective that has to be optimized. The goal of multi-objective optimization (Pareto optimization) [237] is to find solution vectors that represent a reasonable trade-off among different, conflicting objective functions that are optimized simultaneously. In general, for nontrivial multi-objective optimization problems the existence of a single unique solution that would yield an optimal value for each objective simultaneously cannot be expected. Hence, for real-world problems, the objectives are usually at least partly conflicting. A unique vector of different objectives is called nondominated or Pareto optimal if none of its components (objective function values) can be further improved without deteriorating the value of at least one of its other components. That is, there exists no other possible solution point for which all objectives have clearly better values, i. e. at least one of the objectives has to be improved in value, while all other objectives are at least as good. Such a nondominated point (nondominated feasible solution, i. e. nondominated vector of objectives) is an 'efficient

solution' [238] of the multi-objective optimization problem. Since the vectors of objective function values cannot be naturally ordered in the objective space (they can only be partially ordered), multi-objective optimization problems are in a sense ill-defined, hence all Pareto optimal solution points can be considered as equally acceptable. That means each Pareto optimal solution represents a different instance of the unavoidable trade-offs. Therefore, the set of Pareto optimal vectors consists of 'compromise solutions' from which a final solution has to be sorted out by a human decision maker[4] with problem domain expertise [239]. The set of Pareto optimal solutions is called a Pareto front. Scalarized problems in which the optimization of multiple objectives is transformed to a single objective optimization problem [238] are not considered here.

Thus, a multi-objective optimization problem has the following form [237]

$$
\begin{aligned}
\text{minimize} \quad & \{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \dots, f_k(\boldsymbol{x})\} \\
\text{subject to} \quad & \boldsymbol{x} \in S
\end{aligned}
\tag{3.9}
$$

The k ($k \geq 2$) objective functions (criteria) $f_i : S \to \mathbb{R}$ quantify the values of the network characteristics of interest. The decision variable vector $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^T$ belongs to the non-empty feasible region set S, which is the set of alternatives (feasible solutions) of the decision problem [238]. For the problem of computing good edge weight thresholds the decision variables are the individual edge weight percentile values or, respectively, the thresholded binary networks that correspond to these percentiles. In this context, one has finitely many and explicitly given alternatives. Note that multi-objective optimization problems are usually formulated as simultaneous *minimization* problems, whereas in the context of this thesis the selected network characteristics need to be maximized. Multiplication of the objectives with -1 easily converts the maximization problem to yield an equivalent minimization problem. The dimension k of the criterion space $\mathbb{R}^k$ is given by the number of objective functions $f = (f_1, \dots, f_k)$. Depicting $f(\boldsymbol{x})$, the feasible set in criterion space, can help identify nondominated, Pareto optimal points [238], an example of which is given in figures 3.5a and 3.5b. For computing edge weight thresholds, no ranking of the objectives is employed, even though it would be a reasonable choice, for example, to rank modularity values higher than the values of other characteristics. Such a ranking yields a weak definition of order of vectors in criterion space.

---

[4]The decision maker might be algorithmically supported by methods of multicriteria decision aid (MCDA).

The above given informal definition of Pareto optimality can be stated more precisely. A decision variable vector $\boldsymbol{x}^* \in S$ is (globally) Pareto optimal if there does not exist another decision variable vector $\boldsymbol{x} \in S$ such that $f_i(\boldsymbol{x}) \leq f_i(\boldsymbol{x}^*)$, $\forall\, i = 1, \ldots, k$ and $f_j(\boldsymbol{x}) < f_j(\boldsymbol{x}^*)$ for at least one index j. $\boldsymbol{x}^*$ is called efficient point. An objective function vector (outcome) $\boldsymbol{f}^* = (f_1^*, \ldots, f_k^*)^T$ with $f_i^* = f_i(\boldsymbol{x}^*)$ is Pareto optimal if its corresponding decision variable vector is Pareto optimal. $\boldsymbol{f}^*$ is called nondominated point.

The objective values of feasible solutions in S are not known in advance and have to be enumerated. For this reason all considered network characteristics are computed for all considered edge weight percentiles as described above. The sign of the components of the resulting (objective function) vectors are interchanged to handle the situation that the underlying multicriteria optimization problem is formulated in terms of a maximization problem. The Pareto front is computed based on pairwise comparisons and nondominant sorting [240]. To obtain a consensus, single most representative solution the centroid of the set of nondominated objective vectors (points of the Pareto front) is computed. Finally, the nondominated point with minimum Euclidean distance to the centroid vector is considered as the best compromise and is returned.

Preliminary results obtained from network sequences indicate an improved module detection, which yields more clear-cut module structures due to the network-specific edge weight thresholds computed by the Pareto optimization approach. An example is shown in figures 3.5c - 3.5e. Pareto optimal thresholds have a clear meaning with respect to the obtained binary network structure – these thresholds guarantee that the underlying binary network module structure is most pronounced with respect to the evaluated module structure characteristics (objectives). With regard to other network properties, the possibility that bias is introduced cannot be ruled out. For a discussion of thresholding biases I refer to section 2.3 on page 16. A strength and advantage of the presented method for computing edge weight thresholds is its specificity to the structure of individual networks of a network sample. It might replace the commonly used approach of explicit verification of several distinct thresholds applied to the entire network sample at once.

For my work in network module tracking in EEG-derived functional networks (see also
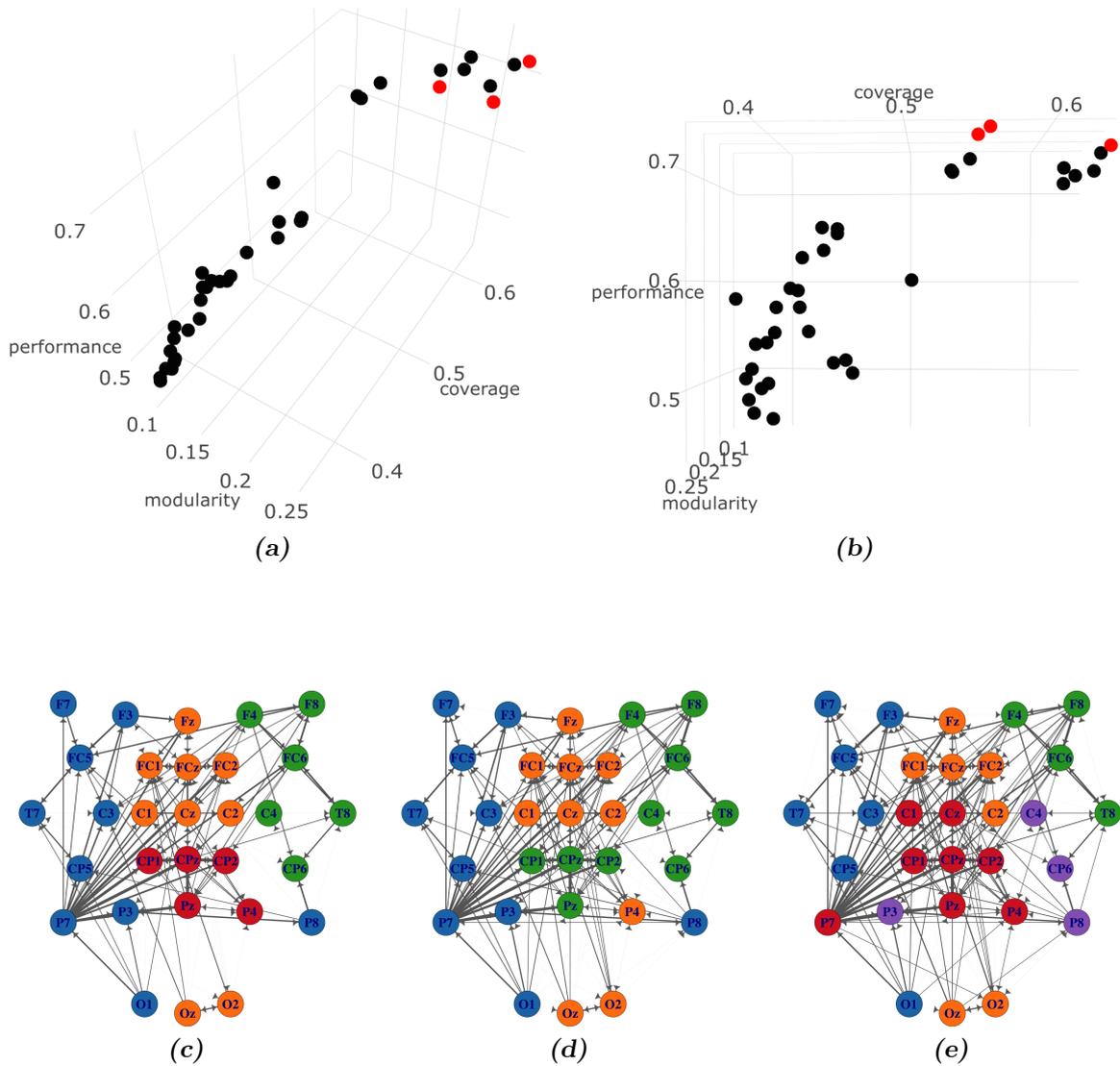
Figure 3.5 – **Pareto front of module structure quality objective vectors for different edge weight percentile decision variables (and module structure in corresponding thresholded networks).** Edge weight threshold analysis was performed for an EEG-derived functional network (dDTF interactions) [241] using the novel multicriteria optimization method. The range of evaluated edge weight percentiles was $0.5, 0.51, \ldots, 0.99$. Large percentile values starting with 0.86 were excluded by the algorithm as the underlying binary network starts to fragment into isolated components. (a) - (b) For each remaining percentile a sphere depicts the corresponding vector of objective values: modularity, performance index and coverage. The red spheres indicate the Pareto front, which consists of the nondominated objective function values that represent the set of optimal trade-offs between network module structure quality characteristics. Finally, the nondominated objective function vector with lowest Euclidean distance to the centroid of the Pareto front was found. It corresponded to an edge weight percentile of 0.85 and an edge weight threshold of 0.069. (c) Network module structure of the weighted network where edges were thinned out using the optimal Pareto front threshold (given by the 85% percentile). (d) Network module structure of the weighted network where edges were thinned out using a Pareto front threshold (0.060, given by the 81% percentile). (e) Network module structure of the weighted network where edges were thinned out using an arbitrary threshold (0.055, given by the 77% percentile). For the actual network data, with the optimal threshold one finds the split of the green module into a red and green submodule (c). This split is not visible with the other Pareto front threshold (d). With the arbitrary threshold the resulting module structure has lower quality and appears more fragmented (e). Module structure was detected with the algorithm of Blondel *et al.* for edge-weighted directed networks.

chapter 7) I also developed and applied another edge filtering approach for network module detection, which is different from standard global thresholding. This approach aims to reverse construct the network and shows some promising first results, too. For it, the algorithm first sorts all edges in decreasing order with respect to their weight. Then, in this order edges are placed back to their original position in the initially empty original node set – the edge with largest weight is placed first, then comes the edge with the second largest weight and so on. This process stops once the constructed network is weakly connected. The resulting network represents the backbone of the strongest interactions of the original network. Note that this procedure does usually not correspond to constructing a maximum spanning tree, as the backbone network likely contains cycles. Depending on the specific spatial distribution of edge weights in the network at hand, it is not unreasonable to expect the strongest connections of the network being primarily situated inside modules. In practice, however, this must not always be the case.

I included an Rcpp [242] function to compute the Pareto front for explicit feasible sets based on nondominated sorting in my utility R package of the network analysis toolbox. The function that implements the presented method for computing edge weight thresholds is part of another one of my R packages for network analysis, as is a function that uses multicriteria optimization in a similar fashion to find network-specific module detection algorithms that yield the highest quality network partitions. To perform module detection in thresholded weighted networks, a different implementation of the above described method exists, which Pareto-optimizes weighted network module structure characteristics instead of binary ones. Also, the weighted edge backbone function is implemented in my network analysis toolbox.

3.6 Extraction and visualization of time evolving module stucture

In many neuroscience settings, longitudinal network data can be extracted naturally from brain recordings. Identification of structural features in time-variant sequences of networks adds another layer of analysis for studying brain processes. This section deals with the problem of extracting information on time-variant module structure in samples of longitudinal networks, i. e. samples of snapshot network sequences. A snapshot network refers to the network corresponding to a particular time step. The analysis concept presented in this section is specifically tailord towards samples of small and high edge density EEG-derived functional

network sequences, in which consistent time-variant module structure identification poses additional challenges.

The work presented in this section does not aim at a direct modeling of module structure dynamics in a sequence of snapshot networks, rather, the aim is to uncover the module structure hidden within the intricate topology of the snapshot network sequence, to visualize this complex information and in this way to direct it to an interpretation with regard to underlying brain activity. The dynamics of module structure changes that have to be identified and visualized is given by the 'life cycle' of individual modules that is briefly described in appendix B.4. Sankey diagrams (alluvial diagrams) [159] represent a suitable means to visualize the time evolution of module structure. In Sankey diagrams the time-varying module memberships of nodes is displayed in the form of different amounts of node flows (ribbons) between modules at different time steps. Example Sankey diagrams can be found in figures 7.2b and 7.2d. This graphical representation gives a good overview over long-term trends in the module affiliations of nodes and it shows time steps with large fluctuations and rapid reconfigurations of module structure. Sankey diagrams help to discern persistent, interesting modules from short-lived noisy ones. In addition to the Sankey diagrams, network plots using the EEG electrode layout delineates the spatial location of modules and their affiliated nodes. To color-code the module representation in the Sankey diagrams, module structure has to be matched between the considered time steps, for which I use the algorithms presented in section 3.2. This matching inherently accounts for the time evolution of module structure, i. e. the similarity of modules at different time steps with respect to shared member nodes. Now that the setting and the desired outcome of the analysis concept is clearly defined I describe the pipelining of all network processing and analysis techniques required to achieve the aims.

The descriptions of the methodology follow the flow chart depicted in figure 3.2b. This concept was successfully applied to EEG data in the application study described in chapter 7. Once a sequence of weighted directed edge-complete networks is obtained and a time grid for investigating snapshot networks is defined, two separate analysis approaches for identification of time-variant module structure can be applied, each offering a different perspective on the time evolution of network modules. The difference between both analysis approaches is given by the way of thresholding the snapshot network sequence.

The first analysis approach is shown on the left branch of the flow chart in figure 3.2b.

It combines empirical multiple thresholding with a pooling of all available snapshot networks
so that the edge weight thresholds corresponding to the predefined percentiles are deter-
mined with respect to the distribution of all available edge weights at all time steps (see
also section 2.3.2). Extraction of time evolving module structure is performed separately
for each obtained threshold. Crucially, the snapshot networks at each considered time step
are not soft thresholded (see also section 2.4) directly. Rather, the predefined time grid is
used to center a time window on each considered time step and to compute a consensus
network for each centered time window. In the process, a consensus network aggregates the
connectivity structure of each considered snapshot network and the surrounding networks in its
temporal neighborhood. Thereby, for each snapshot network on the time grid a given number
of snapshot networks directly preceding it and directly succeeding it in the network sequence
are selected to compute a consensus network. The consensus network is given by the median
interaction strength of each individual edge of all networks within the time window. This
processing step incorporates connectivity information in the vicinity of a given time step into
the module structure identification and in the same time it reduces the influence of random
fluctuations of interaction strengths. The consensus networks at each time step in the time
grid are used for subsequent thresholding and module structure identification. The already
obtained thresholds are then applied to soft threshold each considered consensus snapshot
network, which results in weighted directed networks with different edge densities. This way
of thresholding and aggregating connectivity information potentially makes it possible to
reveal more detailed differences in the connectivity structure of individual time grid snapshot
networks. The rationale behind it is that for many experimental settings where a stimulus
effect is considered a global increase of interaction strengths with time can be expected as a
result of the processing of the stimulus. Therefore, snapshot networks that appear early in
the network sequence are more extensively thresholded and their interaction 'noise' is more
strongly suppressed as compared to time grid snapshot networks that appear later in the
network sequence. The sequence of module structure reconfigurations that is obtained by the
subsequently performed module detection will most likely reflect clearer, more stable changes
and less fluctuations between time steps. Standard multiple thresholding (see also section
2.3.2) on the contrary would result in less clear module structure detection results. Since
this way of global threshold computation potentially yields disconnected networks I advise to
assign all disconnected (isolated) nodes the same separate module affiliation, which is contrary
to the default output of most module detection algorithms, which would assign these nodes

to different separate modules. As disconnected nodes are normally being excluded from the interpretation of the results this step greatly reduces visual clutter from the presentation of the results. In summary, the first approach is focused on revealing module structure in single snapshot networks with respect to the respective connectivity structure and the changing strength of interactions over the entire network sequence.

The second analysis approach, which is shown on the right branch of the flow chart in figure 3.2b is based on my Pareto optimization approach for computing specific thresholds for each snapshot network (see also section 3.5). While it is possible to compute a consensus network in a time window surrounding each considered time grid snapshot network, a consenus module structure may be computed instead (see also chapter 2 section 2.9). Thereby, the considered time grid snapshot networks and all surrounding networks within the time window are soft thresholded individually using Pareto optimal thresholds. Using a weighted network module detection algorithm, the considered snapshot network and all its surrounding networks are partitioned into modules. The information of these separate module structures is combined and enhanced by the consensus clustering procedure, which outputs the final module structure at each considered time step. The second analysis approach reveals module structure in single snapshot networks with respect to only the connectivity structure at the given time step. After obtaining either a network partition of each consensus snapshot network or a consensus clustering of each snapshot network, my module matching algorithms presented in section 3.2 relate the changing modules at each time step to each other. I note that it would be also possible to use consensus clustering in combination with the global edge weight thresholds of the first analysis approach. Alternatively, consensus networks can be computed if the Pareto optimal thesholding is applied. The main point is that some form of aggregation and smooting of connectivity information of networks in the temporal vicinity of each considered time step, including the network at the time step, is likely to improve the results of module detection as it reduces a possible high temporal variation of the identified module affiliations. However, the presented workflow and the combinations of edge thresholding and connectivity structure aggregation yielded the most clear and robust results for the application study in chapter 7. Stability of the obtained network partition of each snapshot network could be further investigated with perturbation analysis [179, 243].

Part III

Applications and results

Chapter 4

Network sample-specific detection of motifs with pairwise-different node labels

THIS chapter deals with the filtering, description and comparison of the intricate topology of samples of node-labeled binary directed networks, which were obtained from EEG recordings during the neural processing of painful intracutaneous electrical stimuli in patients with major depression (MD) and healthy controls. For this task I applied both of my approaches for network sample-specific motif detection that are presented in section 3.1. The results of this study were reported in two different publications [97, 156]. The data generated by this EEG experiment [244] is scientifically interesting by virtue of the not well understood intertwined relationship between pain processing and depression.

In this respect, it is known that chronic pain and major depression are correlated since depression is a common comorbidity of chronic pain and often chronic pain is an additional symptom of depressed patients [245, 246]. It has been confirmed by some studies that thresholds for acute painful stimulation are lower in depressed patients than in healthy controls [247, 248], whereas other studies found the opposite, namely increased thresholds in depressed patients [249, 250, 251, 245, 252]. The physiological basis for pain perception, pain processing and the sensitivity to painful stimuli of depressed patients remains unclear. It is hypothesized that in depressed patients the processing of painful stimuli in the so-called neuromatrix of pain [253] and consequently the functional connectivity might be altered [250].

Analysis of the intricate wiring patterns found in the functional pain processing network samples by visual inspection and by way of qualitative description seems impossible. The aim of the two studies is to extract insightful, clinically relevant but not readily accessible information about elementary directed interaction patterns, i.e. motifs (see also chapter 3, section 3.1), in both groups that can subsequently be used for group comparisons. Such a network decomposition into motifs provides a means to simplify network structure analysis by focussing only on functionally important substructures of the network. Unification of this information for all networks of each investigated network sample is promising to shed light on the basic neural activity which occurs during the processing of painful stimuli in patients with major depression and in the healthy controls. In general, the new approach offers new ways for studying structure–function relationships and design principles of network samples.

4.1   Data description

The following description of the major points of the EEG data acquisition and connectivity analysis is taken mostly verbatim from my publication [97]. Data and connectivity analysis were first described in [244].

4.1.1   Subjects and connectivity analysis

Eighteen patients (10 women, 8 men) with major depression (mean age $\pm$ standard deviation: 38.9 ± 15.5 years) and 18 sex- and age-matched healthy control subjects (39.3 ± 14.8 years) participated in this study. The procedure was approved by the Ethics Committee of the Friedrich Schiller University (reference number 2282–04/08). Major depression was established by a staff psychiatrist according to DSM IV criteria using a structured interview. All subjects were right-handed. Nine patients were treated with antidepressive medication (5 patients received selective serotonin reuptake inhibitors SSRI; 4 patients norepinephrine and serotonin reuptake inhibitors NaSRI) while the remaining participants did not receive any medication. All subjects were electrically stimulated intracutaneously at the tip of the middle fingers of both the right and the left hand. Stimuli consisted of a bipolar rectangular pulse of $10\,ms$ duration. Participants were requested to rate each electrical stimulus on a scale ranging from 0 to 6 [254, 255]. The pain threshold was defined as the intensity yielding a sensation described as a sharp painful pinprick, corresponding to a rating of '3'.

The EEG was recorded continuously during the electrical stimulation from 60 electrodes, referenced to Cz, using a standard EEG cap based on an extended 10-20 system. Finally,

nine electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 (re–referenced to linked ears) that are situated above some of the important regions of pain processing, attention, and depression (frontal, central, and parietal brain regions) were used. In order to compare the pre- and post-stimulus condition, signal sections of 700 ms duration were extracted pre (700ms before onset to the onset of stimulus, i.e. -700ms to 0ms) as well as post stimulus onset (from stimulus onset to 700ms after stimulus onset, i.e. 0ms to 700ms). To assess the functional connectivity between the nine electrodes, the generalized partial directed coherence (gPDC) (see also section B.1.2) was applied. Binary functional brain networks were obtained by thresholding using significantly increased gPDC values and subsequent dichotomization of remaining edges. Statistical significance was assessed using the bootstrap method described in section 2.3.1 and Holm multiple test correction with a multiple significance level of $\alpha = 0.05$.

Connectivity analysis was complicated by the fact that during the Bootstrap procedure not all MVAR processes under H_0 could be realized due to MVAR parameter sets yielding non-stationary processes, so that many networks are partially unobserved. As a consequence, traditional analysis (listwise deletion of networks with missing values) would discard a huge amount of useful information of the data. Since there is no superior approach concerning all possible missing data mechanisms, a reasonable imputation strategy according to the specific data set has to be chosen. Because it seems implausible that observed variables contain useful information to predict missing values (i.e. data is missing at random), we performed extreme case imputation: First, all missing directed edges are imputed as no interactions, while in the second data set they are treated as interactions. Both data sets are analyzed independently, and only subnetworks that are significant in both cases are considered as motifs. This is a conservative approach to deal with the problem of missing values, and it keeps the effect of the imputations on topological pattern selection as small as possible.

4.1.2 Samples of functional connectivity networks

Due to the nature of the EEG data, the connectivity analysis and the clinical question eight different samples of unweighted directed functional connectivity networks have to be considered: They are defined by all combinations of the group assignment (MD – patients suffering from major depression vs. HC – healthy control subjects), the stimulated side (left vs. right) as well as the time window with respect to the stimulus conditions (pre – time window before noxious stimulation vs. post – time window directly following the stimulation, i.e. including the processing of the noxious stimulus). The nomenclature is: MD-pre-left, MD-pre-right,

(a)   **(b)**
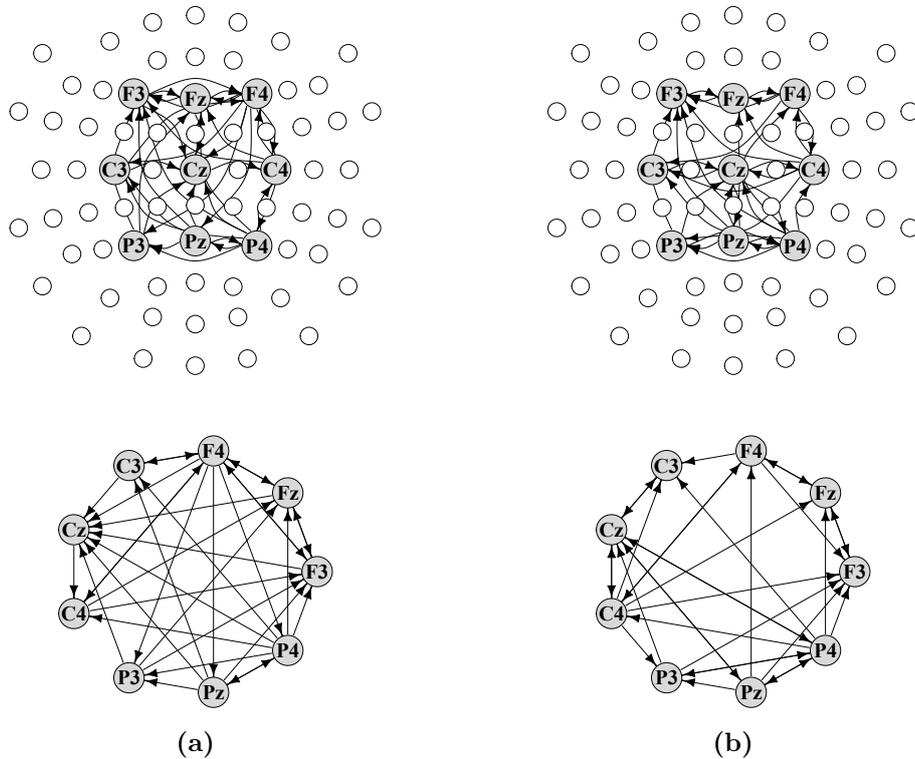
Figure 4.1 – **Functional networks of a healthy subject, pre- and post-stimulus.**
The network is shown within the EEG-electrode layout (top) and as a circle layout (bottom),
which yields a clearer representation of the interaction patterns. The pre-stimulus network is
shown in column (a), whereas the post-stimulus network is shown in column (b).

MD-post-left, MD-post-right, HC-pre-left, and so forth. The sample size for the MD-post-right
sample is fifteen, whereas the sample size is sixteen for all other samples.

These functional connectivity networks have a small size, each consisting of the same set of
nine nodes that are pairwise differently labeled with associated EEG-electrode identifiers. For
our approach to network motif detection in network samples it is crucial that, due to the node
labeling, all nodes are different. As shown in figures 4.1 and 4.2 these functional connectivity
networks exhibit dense and intricate patterns of directed interactions.

4.2 Results and discussion

Both methods to detect motifs, i. e. locatable characteristic topological patterns, were applied
to each of the eight network samples separately. Because of the spatial information associated
with the node labels, it makes sense to look even for 2-node motifs in order to find significant
interactions between two areas covered by the EEG electrode scheme. Furthermore, char-
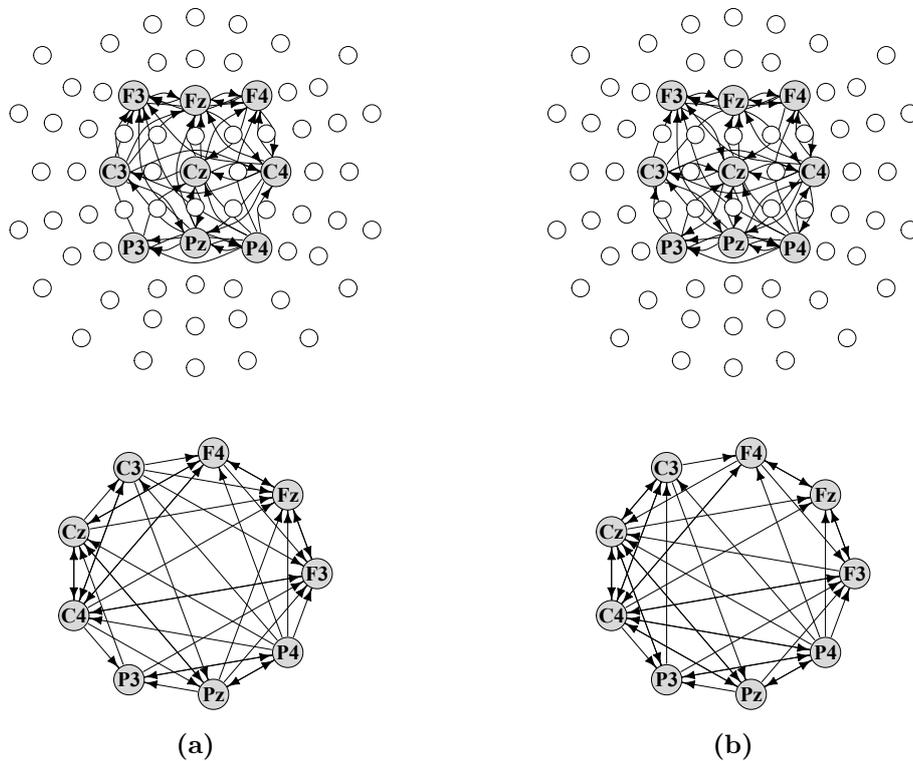
Figure 4.2 – **Functional networks of a patient with MD, pre- and post-stimulus.**
The network is shown within the EEG-electrode layout (top) and as a circle layout (bottom),
which yields a clearer representation of the interaction patterns. The pre-stimulus network is
shown in column (a), whereas the post-stimulus network is shown in column (b).

acteristic interaction patterns given by 3-node motifs were detected. Motifs of a larger size
were not detected, because clinical interpretation of neural processing patterns represented by
2-node motifs and 3-node motifs is already difficult. Hence, detecting larger motifs does not
seem to contribute much to the qualitative knowledge about functional connectivity patterns
in the investigated groups. However, from a theoretical point of view the detection of larger
motifs is straightforward given that sufficiently large samples are available.

As a result of dismissing interactions that are by definition less important, the large amount
of information contained in the intricate directed interactions of the investigated network
samples was reduced to only the characteristic interaction patterns. These were used for
subsequent interpretation with respect to the underlying neural information processing.

4.2.1 Results obtained with the analytical approach for assigning subnetwork significance

For the sake of keeping the presentation of results concise I omit the recapitulation of the results obtained with the analytical approach and refer instead to my publication, where a detailed account of the results and an assessment of their clinical relevance can be found [97]. The results of the randomization-based approach, which are presented in the following section can be seen as a prime example of the utitlity of motif detection for the analysis of functional brain networks. Moreover, despite its more accurate null model for subnetwork significance assessment, the results of the randomization-based approach are consistent to a large degree with the ones obtained with the fast but less accurate analytical approach.

4.2.2 Results obtained with the randomization-based approach for assigning subnetwork significance

A detailed account of the results can be found in my publication [156]. Prior to generating the null model network ensembles for each of the eight network samples, the simulation parameters Ω ('mixing' parameter that determines the number of edge swaps) and b^* (number of random realizations for every sample element network) were determined as described in chapter 3, section 3.1.3. To cope with the immense computational overhead, the 'mixing' parameter Ω was determined for all eight network samples at once. Thereby, the representative network needed for the determination of the parameter Ω was an element of the 'HC-post-left' sample. Random realizations of the representative network were generated using the MCMC edge-switching algorithm with different values of Ω and the independent results were then pooled together to yield a total of 190,400,001 networks. In this set we found 101,996,824 pairwise different networks with the given prescribed degree sequence. According to detailed statistical considerations given in [156], for the analyzed network data the edge-switching algorithm seems to sample networks with a non-uniform distribution for every Ω. Because no particular value of Ω could be identified with the test statistic of the exact chi-squared goodness-of-fit test the lowest value of Ω was selected for which the test statistic was minimal. As it turned out this was $\Omega = 120$. According to the procedure outlined in chapter 3, section 3.1.3 the required number of random realizations for every sample element network in all network samples was $b^* = 7500$. This number reflects a conservative choice, since for most investigated network samples 7000 realizations would have been sufficient to achieve convergence of relative subnetwork frequencies.
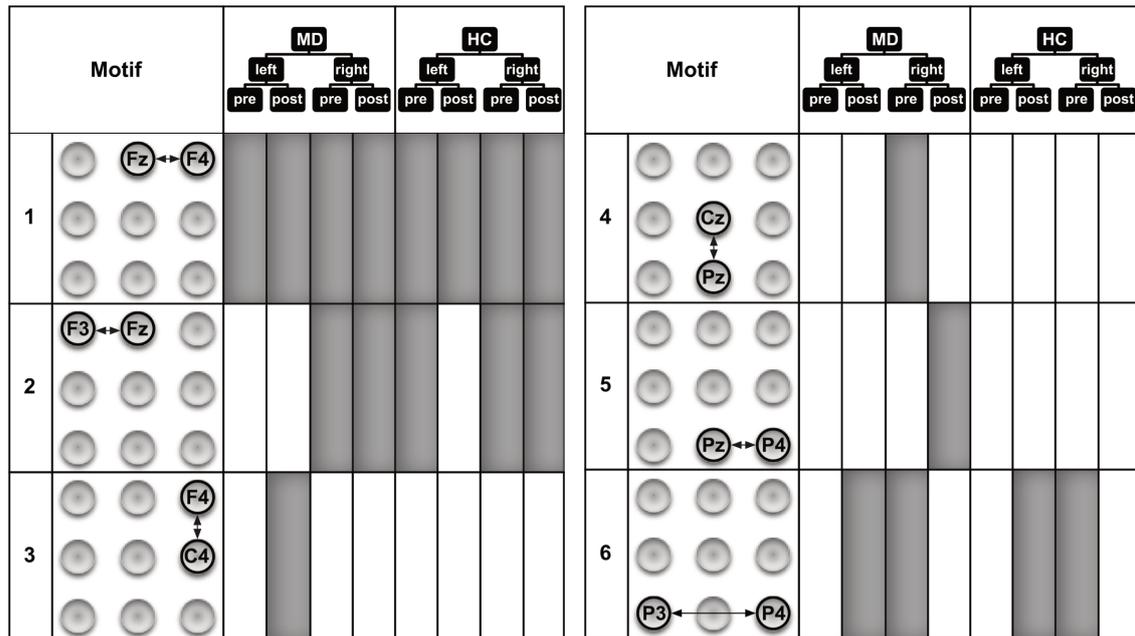
Figure 4.3 – **Motifs of size 2.** The motifs were detected in the eight functional network samples with the randomization-based approach. The occurrence of a 2-node motif in an network sample is indicated by filled areas. These motifs represent characteristic directed interaction patterns of brain activity recorded at three different EEG electrodes that occur before and during the processing of painful electrical stimuli.

The detected motifs of size 2 shown in figure 4.3 reveal several interesting points. MD patients show slightly more motifs of size 2 than HC subjects. However, 8 out of 12 motifs in MD and 8 out of 9 motifs in HC are similar with respect to the motif and the time period when it occurs. Overall this demonstrates that motifs of size 2 show strong communalities in processing between the groups. This shows that the method allows the identification of robust connections. One of these functional connections, Fz $\leftrightarrow$ F4, is present for all time windows for both sites of stimulation. This motif was also seen in the results of the analytical approach (section 4.2.1). It is likely that this connection represents a part of the background activity or attentional processes which are independent of group (MD, HC), time period (pre, post), or site of stimulation (left, right). Other motifs, e. g. F3 $\leftrightarrow$ Fz, are primarily found in association with the stimulation of the right hand. So this processing contralateral to the stimulation site might represent processes of preparing to and analyzing the nociceptive input. Interestingly, this motif is the only 2-node motif in HC that is also not present in MD. It occurs during the pre-stimulus period prior to the left hand stimulation in HC. This activity might represent a preparation in advance of the hand stimulation, e. g. the process of distributing attentional resources. The lack of the F3 $\leftrightarrow$ Fz motif in MD fits with additional motifs in MD which do
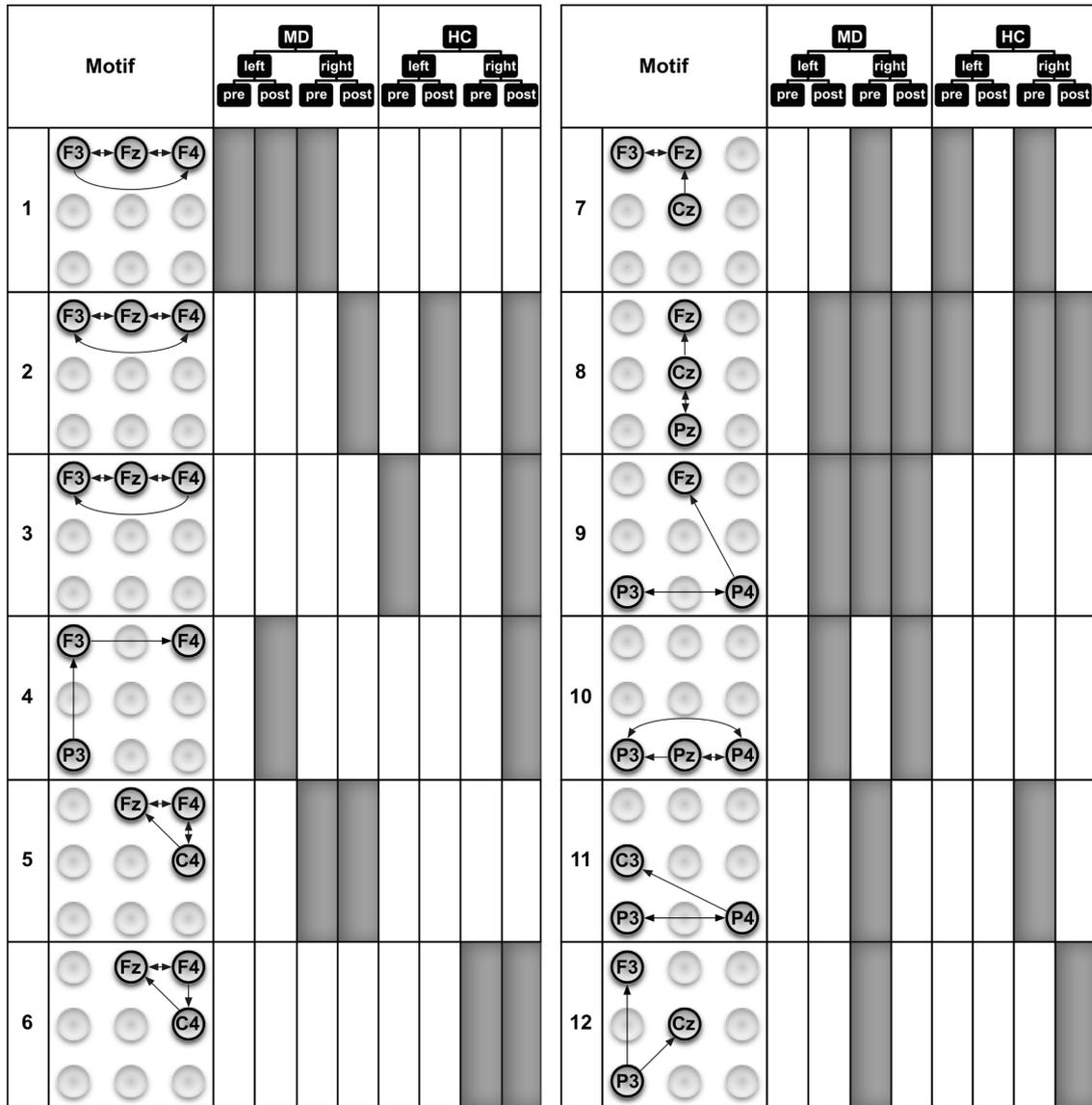
Figure 4.4 – **Motifs of size 3.** The motifs were detected in the eight functional network samples with the randomization-based approach. The occurrence of a 3-node motif in an network sample is indicated by filled areas. These motifs represent characteristic directed interaction patterns of brain activity recorded at three different EEG electrodes that occur before and during the processing of painful electrical stimuli.

not appear in HC. Similarly to previous results presented in section 4.2.1 all these additional motifs are located in the right hemisphere or midline. This might reflect the role of the right hemisphere in the processing of emotions and mood in MD patients [250, 256, 248, 257, 258].

Similarly to motifs of size 2, motifs of size 3 (figure 4.4) are also more often identified in MD than in HC subjects. However, the exact communalities are far less expressed for the motifs of size 3 (5 of 18 in MD; 5 of 13 in HC) compared to size 2 (8 of 12 in MD; 8 of 9 in HC). The results seem to indicate that some motifs in HC are replaced by different motifs in MD. For example, motifs 2 and 3 in HC seem to be replaced by motifs 1 and 2 in MD (including the communality of motif 2 for the processing after stimulation of the right hand). Interestingly, when comparing these motifs between groups, the principle difference lies in stronger activation of the right frontal areas in MD patients. This finding might be interpreted as agreeing with theories on the role of the prefrontal cortex (PFC) in the processing of emotions [256]. The left PFC has been demonstrated to be involved preferentially in processing associated with approach-related, appetitive goals, while the right PFC is more strongly involved in the processing of behavioral inhibition and withdrawal [259, 260]. This theory opens possible interpretations on pathophysiological mechanisms for MD, namely a hypoactivity of the left PFC or a hyperactivity of the right PFC [256]. Our data clearly points to a hyperactivity of the right PFC in our patients. Our data is also consistent with findings indicating the additional recruitment of prefrontal areas by MD patients [261]. While there are nearly as many motifs of size 3 in MD patients during the pre-stimulus period as in the HC subjects, a clear difference can be found in the pre-stimulus period with respect to the site that will become stimulated. Seven out of the 8 motifs of size 3 in MD patients were found before stimulation of the right hand, only one motif was found before stimulation of the left hand. In contrast, in the HC subjects we found 4 motifs of size 3 before stimulation of the right hand and 3 motifs before stimulation of the left hand. Obviously, there is a clear preponderance of motifs before stimulation of the right hand in MD patients. One reason for this preponderance might lie in the contralateral organization of somatosensory information processing. Thus it might be more demanding for MD patients to recruit resources for the analysis of the left hand stimulation because the resources had to be redistributed from the more active right to the left hemisphere. In line with this interpretation, most of the motifs active during the preparation to stimulation of the right hand in MD patients include directed information flow to or within the left hemisphere (i.e. motifs 1, 7, 9, 11 and 12). Another

somewhat surprising finding is that there are slightly more motifs in the post-stimulus period found in MD patients. It was previously found that MD patients compared to HC exhibit higher pain thresholds to external stimulation including electrical stimulation [245, 249], lower sensitivity to C-fiber activation [252], and/or lower sensitivity to experimental nociceptive stimulation [249, 262]. However, it should be mentioned that the stimulation was performed with stimuli that were adjusted for subjective pain ratings (i.e. moderately painful in both groups). This might be the reason that there is no obvious difference in the number of motifs found in MD vs. HC subjects. Nevertheless, there are clear differences with respect to the motifs themselves.

Thus, there exists an interesting phenomenon regarding differences in motif composition between the MD and HC group after stimulation. It seems that the connectivity due to stimulation shows an opposite direction compared to the pre-stimulus connectivity. With regard to left-hand and right-hand stimulation, MD showed an equal number of motifs (5), whereas the HC showed more motifs after right-hand stimulation (6) and only one after left-hand stimulation. This might again represent the nature of stimulation with a preferred contralateral processing of the information. So the noxious stimulation of the right hand will primarily (or, at least, more quickly) activate the left hemisphere, but nociceptive processing will activate behavioral inhibition and withdrawal. Thereby a need exists to transfer the information from the left hemisphere to the right and to activate the right PFC due to the noxious stimulation in the HC subjects. Indications for such a transfer might be seen in motifs 2, 4, or 8. In our MD subjects, there is already a clear preponderance of right hemisphere activation, as discussed in the previous paragraph. This might indicate that activation of the right PFC in MD to the same degree as in the HC, does not occur because it is already activated.

4.2.3 Discussion

An in depth discussion of my approaches for motif detection in network samples in the light of the pain processing functional network samples, motif detection in general and the strengths and limitations of such topological network decomposition approaches can be found in my publications [97, 156]. Here I revisit only a few of the most important points.

I applied both of my methods for motif detection in network samples with pairwise different node labels to the same functional network data. Since both approaches rely on different null models for contrasting the subnetwork counts obtained from the network samples they are

expected to identify –and in fact they do identify– non-identical, but overlapping sets of motifs. For the identified 2-node motifs the agreement is larger than for the 3-node motifs. Thereby, the randomization-based approach seems to be able to better discriminate variations in edge patterns for the same set of nodes, which might be explained by the fact that the design of the underlying null model incorporates more topological information of the input network sample.

It has to be noted that the two presented studies focused on motifs at the sensor level, which means that the current view on anatomical locations of motifs might only serve as a cautious hint with reference to anatomical sources.

The null model widely employed in motif detection preserves the in-degree and out-degree sequences of the input network, which is a basic property on the node level that ultimately affects many other properties of the network. For the assignment of subnetwork significance based on simulated null model networks I used the MCMC edge-switching alorithm (see also section 2.11.2) and applied new techniques for the determination of randomization parameters Ω and b^*. To the best of my knowledge, methods to determine these central parameters were never published before and in the literature on motif detection their choice is usually either not stated or not justified. Since the edge-switching algorithm is applied a large number of times to each sample element network of each of the eight network samples a very large data set of randomized networks was generated. These simulations are very time-consuming with compute times in the range of weeks. Also the subsequent counting of subnetwork appearances in the obtained random ensemble is costly. On the other hand my analytical statistical test is computed much faster, taking only seconds, but comes at the cost of simpler assumptions being made for the null model which accounts for the mean number of edges of the input network sample. Referring to the randomization-based approach, it seems that for small values of Ω the network most often yielded from the randomization process was the input network itself. This is expected, because for smaller values of Ω the MCMC edge-switching algorithm can cover only a small part of the network configuration space. Thus, in general it might be a good strategy to use the last generated network as input for subsequent randomizations instead of always using the investigated real-world network as input for the MCMC edge-switching algorithm. However, for motif detection one is not primarily interested in the particular instances of sampled random networks. Rather, the uniform sampling of networks with precribed degree sequences is an desirable property of the MCMC edge-switching algorithm. In this regard, at least for the investigated network data, the MCMC edge-switching algorithm did not seem to uniformly sample networks with the prescribed degree sequences of the representative

network. The extent of compensating effects given by the choice of the number of random realizations per input network is currently unknown. Thus, the null model random networks generated with the parameter values obtained by my determination techniques might still be 'good enough'. The difficulties arising from determining the value of the 'mixing' parameter Ω illustrate that one must not simply select an arbitrary value for it (as seems to be the case for most of the literature).

The design of a suitable null model formalizes a particular null hypothesis and defines the notion of randomness [208]. It is crucial for distinguishing regular topological effects from contingencies in the structure of network samples (cf. [113, 263, 264]) and thus is crucial for obtaining valid results [216]. At the present time there is no established theoretical background for choosing null models that fit to given network data and thus it is not clear which network properties might be incorporated into a good null model. Some progress was recently made in [209], where the problem of selecting the right base properties for null model networks was explored. Clearly, motif detection would miss any functional meaningful subnetworks that appear only infrequently. Conversely, subnetworks that appear with a statistically significant frequency are not necessarily important for the structure and functioning of their network. Whereas experimental validation of motifs in functional brain network seems to be beyond reach in the foreseeable future, isolated network motifs have been tested experimentally for their regulatory functions as recurring circuits in bacteria and yeast transcription networks [134, 265]. These experimental studies confirmed theoretical predictions and could assign specific modes of molecular information processing to distinct motifs in these networks, thereby demonstrating that some network motifs indeed appear to be building blocks of transcription networks. Yet despite some potential limitations on assigning information processing roles to motifs in functional brain networks both novel approaches for motif detection in network samples might be seen as tools that simplify the topology of all member networks in a network sample by thinning out interactions that are not characteristic for the network sample, thereby yielding a compact description of recurring important topological elements contained in the network sample. Finally, after this simplification step network samples can be compared with respect to their characteristic interaction patterns, which was as yet not attainable.

Chapter 5

Network sample-specific characterization of network topology

NETWORK structure quantification can be utilized to discriminate functional brain network samples and to evaluate them with respect to therapeutical treatment or other effects that can be assumed to be reflected in the topology of sample element networks. Such a description of network structure can be obtained on all three topological scales presented in chapter 2, section 2.7, that is to say the microscopic, mesoscopic and macroscopic scale. Quantification of network topology considerably contributes to the understanding of a network and it complements network decomposition approaches like motif dection, which was presented in the preceding chapter. If applied to functional brain network samples, the resulting characterization of all sample element networks might be the basis for deriving qualitative information about underlying recorded brain processes and their alterations in different investigated groups.

Following I present the network analysis results of an application study [266] where weighted edge-complete functional networks of cognitively impaired HIV-infected patients were investigated. For this work information obtained from structural and functional imaging data was combined to quantify the effects of lithium treatment-induced changes in brain microstructure on gPDC-derived functional connectivity patterns.

5.1 Data description

The following description of the fMRI data and the connectivity analysis is a concise recapitulation of the main facts. Detailed information may be found in a publication I co-authored [266].

5.1.1 Subjects and connectivity analysis

A cohort of seven HIV-infected individuals (4 male, age range 43–52 years, mean = 45.5) with diagnosed cognitive impairment was enrolled in a 10-week, open-label lithium study [267] at the University of Rochester. The study was reviewed and approved by the Institutional Review Board at the University of Rochester Medical Center and all subjects signed a written informed consent prior to undergoing study procedures. The open-label trial study design did not include a control group. Subjects were instructed to begin taking lithium carbonate 300 mg PO at approximate 12-h intervals. Neuroimaging was performed before and after lithium treatment, with a period of 10 weeks between the recordings. A series of BOLD EPI scans (GRE EPI sequence, TR/TE = 2000/30ms) was acquired while participants performed a working memory task. The task was based on Garavan *et al.* [268] and consisted of sequences of large and small squares presented visually for 1500 ms each and intermixed with 100 ms fixation trials. Each sequence of squares was considered a condition and labeled '1-switch', '2-switch' or '3-switch' based on how many times the size of the squares changed during the sequence. Participants were required to retain separate counts of small and large squares in memory and report it at end of the sequence. Each imaging run consisted of 15 randomly presented conditions, five of each 'switch' type. During each visit (pre or post treatment), participants performed three imaging runs, with the order of sequence presentation changing from run to run to avoid practice effects. The diffusion tensor imaging (DTI) results from [267] were used to select the regions of interest (ROI) for the connectivity analysis of this study. The following seven areas were selected as ROIs for the connectivity analysis: right cerebellum, right putamen, right medial frontal gyrus, right and left frontal orbital cortex, right lateral occipital cortex and right subcallosal cortex. The degree of directed information transfer between the seven nodes (ROIs) of the fMRI-based network was quantified by means of time-variant gPDC based on time-variant multivariate autoregressive (tvMVAR) processes (see also appendix B.1.2).

5.1.2 Samples of functional connectivity networks

Due to the way the vast amount of connectivity data was pooled together 600 functional connectivity networks were analyzed: For each HIV-infected patient in each of the three repeated measurements (runs), the three 'switch' conditions were repeated five times each, resulting in 45 functional connectivity networks for each one of the two treatment conditions. This results in 315 networks representing the functional connectivity patterns in all seven patients prior lithium treatment. Since for patients five and six only two instead of three runs from the post treatment visit were available for analysis, the number of functional connectivity networks in the post-lithium treatment condition was just 285. Connectivity patterns were represented as weighted edge-complete directed networks with seven labeled nodes representing ROIs and edge weights defined by the corresponding aggregated gPDC values. All networks of each of the two treatment conditions constitute a network sample. Example networks of both network samples are shown in figure 5.1.

5.2 Analysis of the network samples

The aim of the study was to investigate the network samples with respect to expected differences in the pre-treatment vs. post-treatment global connectivity structure. Contrary to the common approach to analyze thresholded binary networks (see also section 2.3) in this study I extracted topological features from the unaltered *weighted directed edge-complete* networks. This strategy enables to circumvent information loss and the threshold-dependency, i.e. the complications that arise from analyzing different samples of binary networks obtained from dichotomizations of the gPDC data for different, yet arbitrarily defined thresholds. Due to the investigated clinical question and its specific interpretation I primarily analyzed the global increase in interaction strength (average node strength), which is associated with positive effects of lithium treatment on brain function of the investigated patient group. Other topological characteristics were also investigated, including the characteristic path length and weighted clustering coefficients (see also section 2.7). To meaningfully calculate the characteristic path length, connection strengths given by gPDC values were mapped to connection costs using equation 2.8 as suggested in [67]. The obtained network characteristics were further analyzed using surrogate-assisted network analysis (see also section 2.12) so that mechanistic and stochastic effects can be segregated from the influence that nontrivial topological properties have on the assessed measures. With this additional information changes in functional brain

(a)

(b)
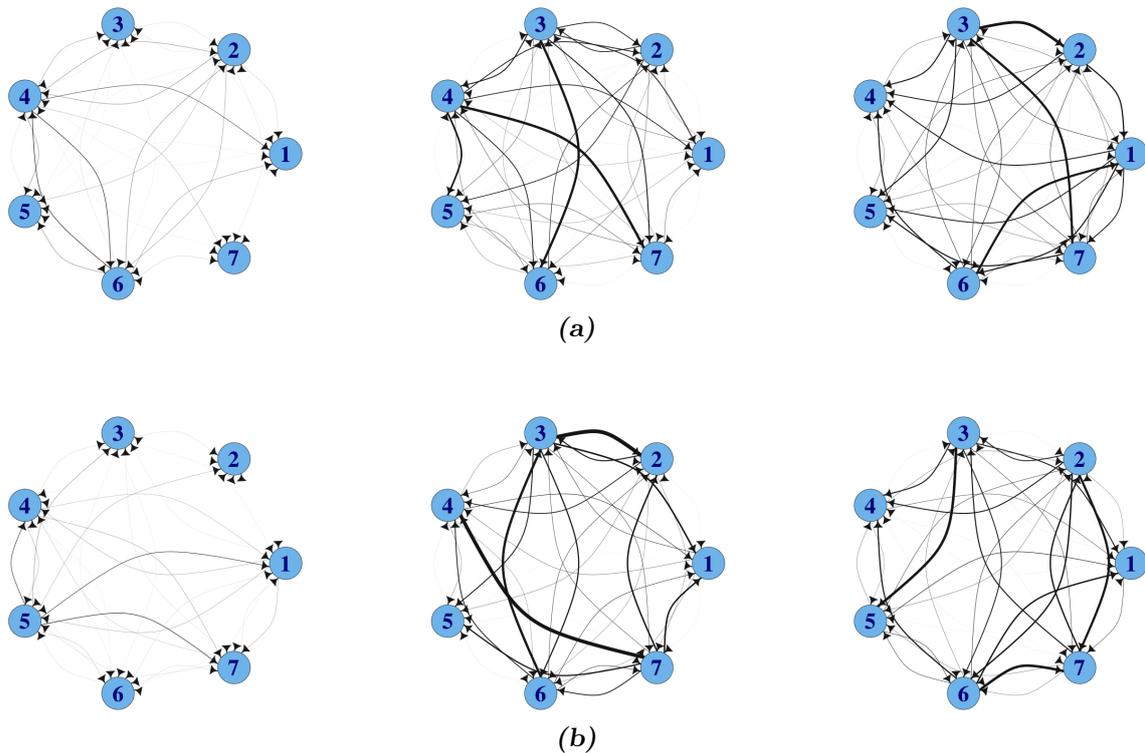
Figure 5.1 – **Pre-treatment and post-treatment functional connectivity structure of three different subjects.** Randomly selected weighted edge-complete networks of three subjects, pre- (a) and post-lithium (b) treatment in the same experimental condition. The weakest interactions are not shown. The following subjects and condition are depicted: subject 1, run 1, switch 1, block 1 — subject 3, run 1, switch 3, block 4 — subject 7, run 3, switch 3, block 15. Detailed information about the experiments may be found in [266]. It can be seen that the interaction strength tends to increase after lithium treatment. The nodes correspond to the following ROIs: 1 – right cerebellum, 2 – right putamen, 3 – right medial frontal gyrus, 4 – left orbital gyrus, 5 – right orbital gyrus, 6 – right lateral occipital cortex and 7 – right subcallosal gyrus.

network characteristics can be reported more conclusively. Thereby, for each functional brain network the observed network characteristics were normalized by dividing them with their mean value in a sample of 1000 surrogate null model networks with random topologies. This normalization rules out biasing effects that stem mainly from very basic topological properties, like the collection of edge weights and accounts for the degree of influence that nontrivial and nonrandom topological properties have on the measures. Additionally, statistical differences between values of unnormalized network characteristics in each sample of functional brain networks and their average values in their surrogate network ensembles were identified with the paired two-sided Wilcoxon signed rank test [269] and a significance level of $\alpha = 0.05$. A fully connected, edge-complete network is combinatorially very constrained with respect to randomization procedures, i.e. edges cannot be rewired. As surrogates I used an ensemble

of random networks that preserve the weights of the edge-complete (i. e. fully connected) weighted functional brain networks by means of global edge weight permutations (random shuffling of edge weights) as described in [94].

5.3   Results and discussion

Following I present the main results of this study. A detailed description of the results may be found in [266] and its supplement.

The main finding of this study is the increase in interaction strength (average node strength) after lithium treatment ($p = 4.7{\cdot}10^{-7}$). Such a lithium treatment effect has never been reported before. A global lithium effect on the characteristic path length was also found, which was decreased due to the lithium treatment ($p < 10^{-7}$). Raw values of all seven weighted clustering coefficients were significantly increased in the post-lithium treatment condition ($p < 10^{-4}$). These findings were obtained using linear mixed models [270] (this statistical analysis was performed by coworkers of my institute). An additional perspective on the network analysis results, in particular the weighted clustering coefficient, was obtained by surrogate-assisted analysis. For some individual weighted clustering coefficients statistical differences between the functional brain networks of both network samples and their randomized surrogate counterparts were found. These weighted clustering coefficients had both, statistically significantly increased and decreased values. Since high values of clustering coefficients are commonly associated with good local information processing capabilities, this result hints at non-trivial effects of lithium treatment on brain function. The results are represented in figure 5.2. To summarize both statistical analyses, the weighted clustering coefficients of all nodes were statistically significantly increased post lithium treatment and for at least some of the nodes it cannot be entirely ruled out that this increase can be explained, at least in part, by nontrivial local changes in the underlying networks. However, since the effect sizes were small one should refrain from emphasizing the increase in weighted node clustering too strongly in the interpretation of the results. To give an idea of typical effect sizes, I state the statistical results of the surrogate-assisted analysis of weighted clustering coefficients for the post-treatment network sample. As can be seen in figure 5.2, ROIs 3, 6, 7 have increased weighted clustering coefficients compared to the respective randomized surrogate network ensembles. According to the paired two-sided Wilcoxon signed rank test this increase is statistically significant, with p-values $p < 10^{-16}$, $p = 1.33 \cdot 10^{-12}$ and $p < 10^{-16}$, respectively. For these nodes the

median values (and interquartile range) of the normalized weighted clustering coefficients were 1.070 (0.142), 1.062 (0.163) and 1.060 (0.157), respectively. No statistical differences were found for the weighted clustering coefficients of ROI 2 ($p = 0.868$; 0.997 (0.145)) and ROI 4 ($p = 0.985$; 0.998 (0.195)). The weighted clustering coefficients of ROI 1 ($p < 10^{-16}$; 0.914 (0.164)) and ROI 5 ($p = 4.46 \cdot 10^{-6}$; 0.967 (0.165)) were decreased in the functional brain networks as compared to the surrogate networks. For the pre-treatment network sample, the situation of the effect sizes is similar. Contrary to the situation for the weighted clustering coefficients, the value of the characteristic path length was not changed in the surrogate network ensembles. The reason for it is the edge-completeness of the analyzed functional brain networks and the fact that edge weights obtained after the necessary strength–to–cost transformation are homogenous enough so that the direct connections between node pairs are always the shortest ones. Since the collection of edge weights is preserved in the surrogate networks, the characteristic path length does not change. This illustrates a drawback of not filtering out low weight edges prior to the network analysis. In addition, there is the possibility for biasing effects that spurious interactions might have on the results.

(a)   *(b)*
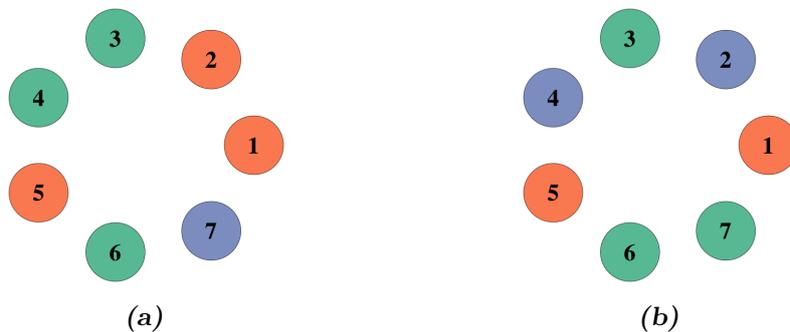
Figure 5.2 – **Network sample-specific changes of the weighted clustering coefficient pre- and post-lithium treatment.** While all weighted clustering coefficients were significantly increased after lithium treatment, surrogate-assisted network analysis, as a supplemental statistics layer with different emphasis, indicates that for some nodes this increase is likely caused solely by mechanistic effects of the global increase in interaction strength and cannot be associated to any additional non-random topological changes. Thus, surrogate-assisted analysis gives a more detailed picture of the network characteristics in the pre-treatment (a) and post-treatment network sample (b). ● nodes had significantly larger clustering coefficients in the functional brain network sample than in the surrogate networks. For ● nodes there was no statistical difference in clustering coefficients between the real and surrogate networks. ● nodes had significantly decreased clustering coefficients as compared to the surrogate networks. Functionally relevant clustering is increased after treatment, as node 7 and node 2 are improved with respect to their local clustering. However, all effect sizes were small. The nodes correspond to the following ROIs: 1 – right cerebellum, 2 – right putamen, 3 – right medial frontal gyrus, 4 – left orbital gyrus, 5 – right orbital gyrus, 6 – right lateral occipital cortex and 7 – right subcallosal gyrus.

The found global increase of interaction strengths might be indicative of a positive over-all effect of lithium treatment on cognitive performance [271] of HIV-infected patients with cognitive impairment. A decrease of the characteristic path length post treatment could be associated with the benefit of lithium treatment in restoring or building efficient connections for information transfer between nodes in the underlying brain network. It seems that lithium treatment of HIV-infected individuals induces changes in brain microstructure (as assessed by DTI) that are associated with improved performance related features of brain functional network connectivity (as assessed by fMRI). Considering the results of surrogate-assisted analysis, it remains largely unresolved to which extent the changes in the characteristic path length and weighted clustering coefficients are a direct byproduct of the global increase in edge weights post-lithium treatment or if additional functionally relevant effects on the network topology are at work, too.

Chapter 6

Analysis of module structure quality in large scale functional brain networks

S ETS of strongly interconnected nodes called network modules (see also chapter 2, section 2.7.2), or communities represent a defining topological feature of many network data sets [70, 148]. They indicate relatively independent indivisible and cohesive substructures of a network that play an important role in the organization of network structure. Typically, the subnetworks induced by network modules are characterized by a larger number of internal interactions and stronger internal interaction patterns that allow for more information flow between affiliated vertices as compared to interactions between these subnetworks. Network module structure might be associated with specific domains of behavior and cognition [3] and plays a role in human learning [15] and the organization of human brain structural connectivity [272]. Thus, obtaining a functional segmentation of a brain network node set yields a picture of brain activity at a specific instant of time. Network module detection at the outset is inherently connected with two uncertainties: It is usually a priori not clear which module detection algorithm is suited to partition the network data at hand, i.e. how good its detection strategy fits the network topology and the underlying processes it encodes. After having uncovered a network partition, its quality with respect to topological information remains to be determined, as well as its compatibility with the specific concept and notion the user has about network modules. However, prior to obtaining a functional segmentation in the form of identified module structure in functional brain networks the underlying cognitive processes

and the information flow among and within spatially distinct regions of the brain has to be
quantified [3]. High dimensional data, such as fMRI recodings, in combination with a low
temporal resolution imposes computational limits on classical Granger Causality analyses with
respect to obtaining a representation of functional interactions in the brain at a large and
highly resolved scale.

The work described in this chapter was published in two publications about a new ap-
proach, the large scale Granger Causality Index (lsGCI) (see also appendix B.1.1) that much
improves the quantification of directed information transfer in very high-dimensional systems
and that is less restricted by dimensionality of the underlying time series than other approaches.
The preservation of a high spatial data dimensionality that is possible with this new large
scale Granger Causality Index (lsGCI) may result in the computation of unprecedentedly
large functional brain networks, representing functional connectivity patterns consisting of
several thousands of nodes connected by millions of edges. This approach was developed by
my colleagues at the institute. To evaluate the practical effect of the lsGCI approach I applied
several state-of-the-art module detection algorithms to lsGCI-derived functional networks
computed from time series that were realized on the basis of ground truth network data,
for which I devised a simulation algorithm (see also chapter 3, section 3.4). Subsequently, I
analyzed module structure quality using the analysis strategy outlined in chapter 3, section
3.3 to examine the tradeoff between computing increasingly higher dimensional functional
connectivity networks and the loss of topological information caused by the approach. Finally,
lsGCI-derived spatially highly resolved functional networks obtained from fMRI brain scans
were investigated with respect to their module structure (functional segmentation). For ob-
taining a visualization of a subject's functional segmentation, the network module affiliations
of all voxels can by directly projected back onto the neuroimaging data slices that were used
for the lsGCI computations. The first of these two publications, which will be referred to as
study I [61], is basically a feasibility analysis that demonstrates the usefulness of the large
scale Granger Causality Index. The data used for this study was one instance of a ground
truth network model ($|\mathcal{V}| = 50$, see also 3.4 on page 71) and clinical data in the form of a
functional brain network computed from one slice of an fMRI scan ($|\mathcal{V}| = 1031$). The other
study, referred to as study II [273], builts upon the results of the first study and enhanced our
methodology towards a representation of almost full-brain fMRI functional interactions. For
the analyses involved in this study, a sample of 100 simulated ground truth networks of size

$|\mathcal{V}| = 100$ and several fMRI-based functional brain networks of size between $|\mathcal{V}| = [5723, 6007]$ were used.

6.1 Study I - Data description

The following concise description of the first study and the involved data analysis is taken from my publication, where a detailed account of the application study may be found [61].

6.1.1 Subjects and connectivity analysis

Data from one male healthy subject was used in the first study, which was conducted as a proof of principle study to investigate whether our proposed techniques for analyzing functional brain connectivity at the large spatial scale yields useful results. The acquisition was approved by the Ethics Committee of the University of Rochester Medical Center (reference number RSRB00042912), and the individual gave his written consent. Resting state [11, 227] (EPI-BOLD) fMRI images were acquired. Each volume consisted of 30 axial slices with an in-plane resolution of 4 mm x 4 mm. The inter-slice distance was 4mm. During the scan, the subject was instructed to stay still and keep eyes closed. To aid in localization and registration of functional data, a high-resolution T1-weighted MPRAGE sequence was acquired. 192 slices were acquired in the sagittal direction with an in-plane resolution of 1 mm x 1 mm and a slice thickness of 1 mm. Functional MRI data was then preprocessed using FSL v4.1.9. Data volumes were motion corrected, brain extracted, temporally filtered with a high-pass filter of cut-off frequency 0.005 Hz, and normalized to MNI152 brain atlas. To demonstrate the value of the lsGCI methodology for obtaining functional segmentations of brain connectivity data, we used one slice (slice number 10) for network module structure identification, which contained $N = 1031$ voxels associated to the brain. The MVAR model order was set to 6. The weighted directed lsGCI-derived networks were binarized using edge weight percentiles (see also chapter 2, section 2.3.2) that were defined in advance as thresholds (40th, 60th and 80th percentile of lsGCI values). All binary directed networks remained weakly connected after thresholding. This is important, because network fragmentation causes problems for some module detection algorithms and would not represent a good model of (functional) brain connectivity from a neurophysiological point of view.

6.1.2   Synthetic networks with ground truth module structure

To investigate the reliability of the lsGCI approach, i. e. to investigate the effect of dimension
reduction on network module structure and for comparing the lsGCI approach with the
conventional Granger Causality Index I simulated an artificial data set consisting of binary
directed ground truth (ad hoc) networks with known, pre-defined module structure (see also
chapter 3, section 3.4). The module structure's degree of definiteness depends on the chosen
parameterization and should be identifiable by module detection algorithms. The algorithm
and the chosen parameterization used in study I [61] to simulate such networks differs slightly
from the one used in study II, which is described in the next section. For the analysis, one
generated network instance was used to realize different time series as input to the lsGCI
approach and subsequent functional network computation. The size of the simulated ground
truth network was $D = 50$. Given a predefined number of modules and their sizes (two modules
of size 12, two modules of size 13), nodes are selected and accordingly assigned membership to
non-overlapping modules. The edges connecting these nodes in the ground truth networks are
placed randomly under constraints that define a notion of module structure, as designated by
the simulation algorithm outlined in section 3.4, the variant that does not vary the network
generation parameters for each generated network. The chosen parameterization was $p_{int} = 0.7$,
$p_{ext} = 0.02$, $\kappa_{int}^{in} = \kappa_{int}^{out} = 6$ and $\kappa_{ext}^{in} = \kappa_{ext}^{out} = 4$.

Stationary first order multivariate autoregressive processes ($N_t = 1000$ and $N_t = 3000$)
were simulated on the basis of the resulting adjacency matrix $\mathcal{A}$, the details of which can
be found in the appendix B.2 on page 150. Interactions between components of these time
series were identified with the lsGCI approach and used to construct edge-complete and
edge-weighted functional connectivity networks. To dichotomize and to remove spurious edge
weights given by lsGCI values, statistical significance of lsGCI values was determined using a
Monte Carlo method (see also chapter 2, section 2.3.1), which estimates the distribution of
lsGCI values under the null hypothesis of no directed interactions between respective nodes
v_i and v_j. An analytical distribution under H_0 [274] was used for the dichotomization of
synthetic networks based on the classical GCI. These statistical analyses were performed with
a type-I error of 1% adjusted for multiple comparisons.

6.2   Study II - Data description

The following concise description of the data used in the second study is taken from my publication, where a detailed account of the application study may be found [273].

6.2.1   Subjects and connectivity analysis

Five HIV positive subjects (four males, one female; mean age: 41 years; age range: 28-53 years) participated in this study. The study was approved by the Ethics Committee of the University of Rochester Medical Center (reference number RSRB00042912). Brain scans of the subjects were acquired at the Rochester Center for Brain Imaging. High resolution structural imaging was performed using T1-weighted magnetization-prepared rapid gradient echo sequence (MPRAGE). Resting state fMRI scans were acquired using a gradient spin echo sequence. Four independent runs were recorded for each subject, where the acquisition of each run lasted 6 minutes with 250 volumes each. A total of 25 slices, each 5 mm thick, was acquired for each volume. During acquisition, the subject was asked to lie still with closed eyes. The first 10 volumes were deleted to allow the signal to reach equilibrium. The volumes were then subjected to slice timing and motion correction as well as brain extraction. Linear detrending was performed by high pass filtering (0.01 Hz). These were then registered to the standard MNI152 template (2 mm isotropic). For subsequent analyses, time series from ventricles were masked out using the standard ventricle mask based on the MNI152 template available in FSL [275]. All preprocessing steps were carried out using FEAT (FMRI Expert Analysis Tool), which is part of FSL and its respective subroutines.

Connectivity analysis of clinical data was limited to every third voxel, in sagittal, frontal and transverse direction, still resulting in $|\mathcal{V}| = 5723$ to $|\mathcal{V}| = 6007$ voxel time series to be processed. The MVAR model order was set to 5. The lsGCI approach utilizes an embedded dimension reduction during the MVAR modeling process. In the studies described in this chapter the Principal Component Analysis was applied, introducing a parameter, the variance explanation, into the lsGCI approach. The amount of variance explanation was variably chosen between 80% and 90%. The weighted directed lsGCI-derived networks were binarized using edge weight percentiles (see also chapter 2, section 2.3.2) that were defined in advance as thresholds (90th, 95th and 98th percentile of lsGCI values). Relatively high thresholds were chosen to preferentially obtain networks with a reduced fixed edge density [37]. All binary

directed networks remained weakly connected after thresholding.

6.2.2   Synthetic networks with ground truth module structure

For the simulation of edge-directed networks with ground truth module structure I used my algorithm described in chapter 3, section 3.4 on page 71 with the following parameter settings: 100 different instances of ground truth networks were generated for each network size $|\mathcal{V}| \in \{100, 200, ..., 800\}$. These numbers of nodes were chosen to obtain networks that can be still processed with standard GCI methods for the purpose of quantitative comparisons. The generating algorithm scales the number of modules with the network size, so that for every increase of 100 nodes eight additional clusters of nodes are simulated. All column sums in the simulated adjacency matrices were restricted to be at most fifteen (an explanation for it is given in appendix B.2). The probability for intra-module edges in both directions between pairs of nodes was given by $p_{int} = 0.5$, whereas the probability of directed inter-module edges depends on the network size so that the constraint on the column sums holds true. This probability was given by $p_{ext} = 3/(N - 15)$. Underlying this probability for each node v_i is the conservative assumption that v_i has on average only 3 in-going connections[1] from at least $N - 15$ possible nodes outside the module of v_i, with 15 being the maximum size of any module. The minimum intra-module ("internal") in- and out-degree and the maximum inter-module ("external") in- and out-degree of all nodes were set to $k_{int}^{in} = k_{int}^{out} = k_{ext}^{in} = k_{ext}^{out} = 4$. An example of a ground truth network is shown in figures 6.6 (a) and 6.7 (a).

Similar to study I, for each network size $|\mathcal{V}| \in \{100, 200, \cdots, 800\}$ stationary first order multivariate autoregressive processes with 1000 temporal samples each were simulated on the basis of the adjacency matrices of 100 instances of ground truth networks (see also appendix B.2 on page 150). Interactions between components of these time series were identified with the lsGCI approach and used to construct edge-complete and edge-weighted functional connectivity networks. These networks were dichotomized using network-specific thresholds that were chosen so that the resulting binary networks had maximum similarity to their associated ground truth network according to a maximum Cohen's kappa [197, 276]. Results are presented for ground truth networks of size $|\mathcal{V}| = 100$.

---

[1]It was found that with increasing network size too many outlier nodes with either more than 4 inter-module edges or with in-degree (column sum) larger than 15 were generated, which causes frequent restarts of the simulation. This problem is slightly more pronounced for nodes whose module size is smaller than 15, since they potentially can make slightly more interactions with external nodes.

6.3   Module structure quality analysis for the assessment of network topology alterations

As mentioned above, the analysis of the extent to which the dimension reduction step of the lsGCI procedure negatively affects edge patterns —the interaction structure between the recorded time series of brain activity— and degrades the recoverability (preservation), definiteness and quality of network modules was based on samples of simulated networks with ground truth module structure and the corresponding realized multivariate time series, which comprise the ground truth connectivity patterns. Network module structure was detected in the resulting samples of lsGCI-derived networks and, for comparisons, in GCI-derived networks. For it, different module detection algorithms for binary directed and binary undirected (i.e. symmetrized) networks were used. They are described in section 2.8 on page 31. The effect of varying degrees of lsGCI dimension reduction on the quality, definiteness and recoverability of network module structure was assessed by contrasting detected network partitions of (ls)GCI-derived networks with the detected and known module affiliation of nodes in the ground truth network using several network structure characteristics. The structural and information theoretic measures that were used for the quantification of network module structure are described in section 2.10 on page 38.

6.3.1   Results of study I

In study I [61] network partitions were identified with the algorithm of Blondel *et al.* for directed networks (see also section 2.8). Tables 6.1, 6.3, 6.2 and 6.4 show the results of the involved comparisons. Thereby, *module recoverability* in the networks identified by (ls)GCI approaches was quantified by the following measures: ratio of correctly classified nodes with regard to their known module membership classification, normalized variation of information as a partition distance and normalized mutual information between two network partitions. Additional information to these measures of network partition similarity was obtained by assessing the *quality of the partitions* to distinguish "clear-cut" from "weak" network partitions. The following quality measures were evaluated: modularity measure, performance, coverage and overall average silhouette width.

As expected, for time series length $N_t = 1000$ the impact of dimension reduction on network topology is noticeable, but remains manageable as the assessed module structure characteristics

are reduced in comparison to the ground truth network, yet they are not conflicting with the presence of a marked network module structure (tables 6.1 and 6.2). It can be seen in panels (d) and (e) in figures 6.1 and 6.2 that intra-module and inter-module edge patterns are thinned out (false negatives) as compared to the ground truth network. Still, for $N_t = 1000$ and different degrees of dimension reduction the recoverability of ground truth nework modules is acceptable, with up to 98% correctly identified module affiliations of nodes, which is a larger rate than for the GCI network and it explains the only moderately reduced similarity of obtained network partitions with a low variation of information (partition distance) and sufficiently large mutual information between lsGCI and ground truth network partitions. Surprisingly, certain levels of variance explanation seem to yield networks in which the preservation of module structure is particularly good, even better than for the classical GCI. Modularity values, that measure the quality of module structure inherent in the network, are already high. It can be seen that higher modularity values do not always translate to a higher ratio of correctly classified nodes. This means that topological alterations result in a new, slightly different module structure as compared to the ground truth network. Overall average silhouette width values were somewhat low, indicating that nodes were either not always placed in their best fitting module with respect to their dissimilariy with comember nodes or that due to edge pattern alterations the differences betweenn the average dissimilarity of nodes to comember nodes and to nodes of other modules were not marked enough. There might also be a bias introduced by the choice of the node dissimilariy measure and perhaps another such measure could have been used (see the discussion in chapter 2, section 2.10).

Table 6.1 – **Network partitioning: module quality assessment (study I, $|\mathcal{V}| = 50$, $N_t = 1000$).** Network module structure definiteness can be quantified and compared using several measures that account for module membership assignments of nodes and how this classification is backed up by topological properties of identified, putative modules. For lsGCI-derived networks, the degree of variance explanation is stated.

Network		Modularity	Performance	Coverage	Avg. silhouette
ground truth		0.651	0.942	0.903	0.395
GCI		0.529	0.890	0.731	0.090
lsGCI	93%	0.478	0.725	0.839	0.090
lsGCI	86%	0.555	0.842	0.818	0.103
lsGCI	78%	0.441	0.900	0.602	0.059
lsGCI	69%	0.482	0.831	0.737	0.070

For $N_t = 3000$ the quality of (ls)GCI network module structure is similar to the ground

Table 6.2 – **Network partitioning: module recoverability assessment (study I,** $|\mathcal{V}| = 50$, $N_t = 1000$**).** The effect of the lsGCI dimension reduction step on recoverability of ground truth network module structure can be quantified using several measures that account for the similarity of identified network partitions in (ls)GCI networks and the ground truth network. For lsGCI-derived networks, the degree of variance explanation is stated.

Network		Correctl. class. nodes	Variation of inf.	Mutual inf.
GCI		0.800	0.116	0.852
lsGCI	93%	0.760	0.089	0.857
lsGCI	86%	0.940	0.073	0.897
lsGCI	78%	0.640	0.169	0.807
lsGCI	69%	0.980	0.036	0.949

truth network (table 6.3), with the performance measure being slightly smaller as compared to the case of the shorter time series. High performance and coverage values indicate that the identified module structure fits the network topology of the lsGCI-derived networks with respect to large numbers of node pairs of the same module being connected by edges and only comparatively few edges falling between nodes of different modules. In the case of $N_t = 3000$ all characteristics that quantify the recoverability of module affiliations of all nodes take on their optimal value for all analyzed (ls)GCI-derived networks, which means that the ground truth module affiliations of all nodes could be fully recovered (table 6.4). The expected positive effect of longer time series on the reproducibility of network topology by lsGCI and GCI approaches was thus observed in the data. Contrary to the case of $N_t = 1000$, for $N_t = 3000$ it can be seen in figure 6.1b, 6.1c and figure 6.2b 6.2c that there are spurious, false positive interactions between nodes of different modules, given by increased numbers of inter-module edges, while intra-module interactions remained largely stable as compared to the ground truth network.

To further evaluate the effect of the embedded dimension reduction step on binary edge patterns beyond visual inspection (figures 6.1 and 6.2) we considered Cohen's kappa (figure 6.3) for measuring the agreement between a ground truth adjacency matrix (figure 6.1a) and corresponding adjacency matrices identified by lsGCI and classical GCI approaches (figure 6.1b - 6.1e). As before, in the case of the longer time series, $N_t = 3000$, the agreement with the ground truth is substantial up to very good. For $N_t = 1000$, the agreement may be considered as moderate for the GCI-derived network and lsGCI-derived networks for which the variance explanation was at least 78%.

Table 6.3 – **Network partitioning: module quality assessment (study I, $|\mathcal{V}| = 50$, $N_t = 3000$).** Network module structure definiteness can be quantified and compared using several measures that account for module membership assignments of nodes and how this classification is backed up by topological properties of identified, putative modules. For lsGCI-derived networks, the degree of variance explanation is stated.

Network		Modularity	Performance	Coverage	Avg. silhouette
ground truth		0.651	0.942	0.903	0.395
GCI		0.593	0.915	0.845	0.332
lsGCI	92%	0.572	0.898	0.824	0.258
lsGCI	84%	0.571	0.898	0.824	0.237
lsGCI	75%	0.591	0.899	0.844	0.218
lsGCI	66%	0.588	0.887	0.839	0.202

Table 6.4 – **Network partitioning: module recoverability assessment (study I, $|\mathcal{V}| = 50$, $N_t = 3000$).** The effect of the lsGCI dimension reduction step on recoverability of ground truth network module structure can be quantified using several measures that account for the similarity of identified network partitions in (ls)GCI networks and the ground truth network. For lsGCI-derived networks, the degree of variance explanation is stated.

Network		Correctl. class. nodes	Variation of inf.	Mutual inf.
GCI		1	0	1
lsGCI	92%	1	0	1
lsGCI	84%	1	0	1
lsGCI	75%	1	0	1
lsGCI	66%	1	0	1
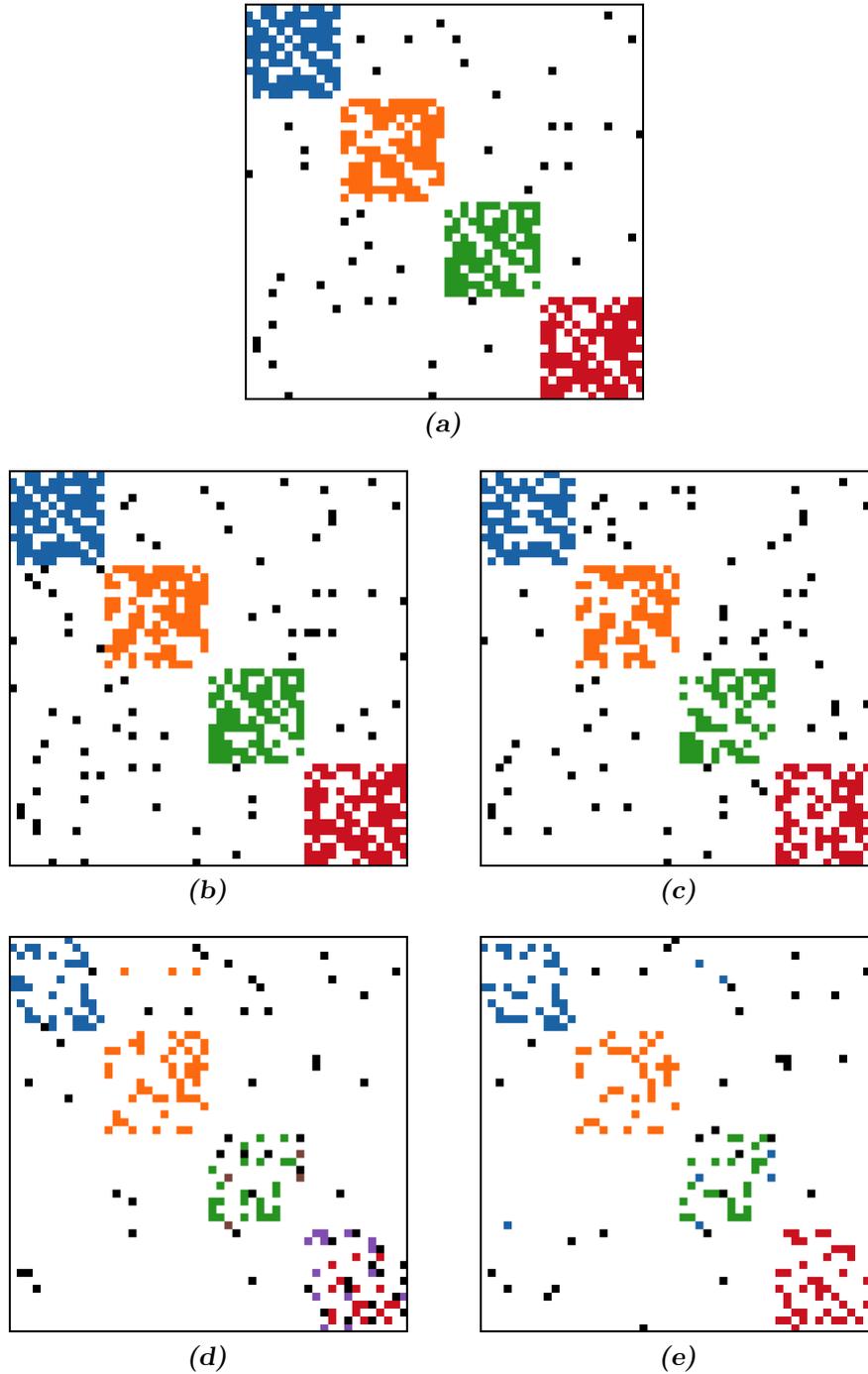
Figure 6.1 – **Edge pattern alterations in the adjacency matrices of GCI and lsGCI networks in comparison to the ground truth network (study I).** Networks (b) - (e) are based on the simulated network with ground truth module structure (a). Interactions are color-coded with respect to the module affiliation of their nodes as detected by the algorithm of Blondel *et al.* ("Louvain"). Inter-module edges that represent interactions between nodes of different modules are shown in black. (a) ground truth network ($|\mathcal{V}| = 50$, ground truth module affiliations were recovered 100%), (b) GCI computed network ($N_t = 3000$), (c) lsGCI computed network ($N_t = 3000$, 84% variance explanation), (d) GCI computed network ($N_t = 1000$), (e) lsGCI computed network ($N_t = 1000$, 86% variance explanation). The four network modules are visible as blocks of dense edge patterns centered on the main diagonal of the adjacency matrices. As expected, the preservation of ground truth edge patterns depends on time series length and degree of dimension reduction. However, even though edge patterns are altered and intra-module edges are thinned out, the module structure is still apparent in all cases.
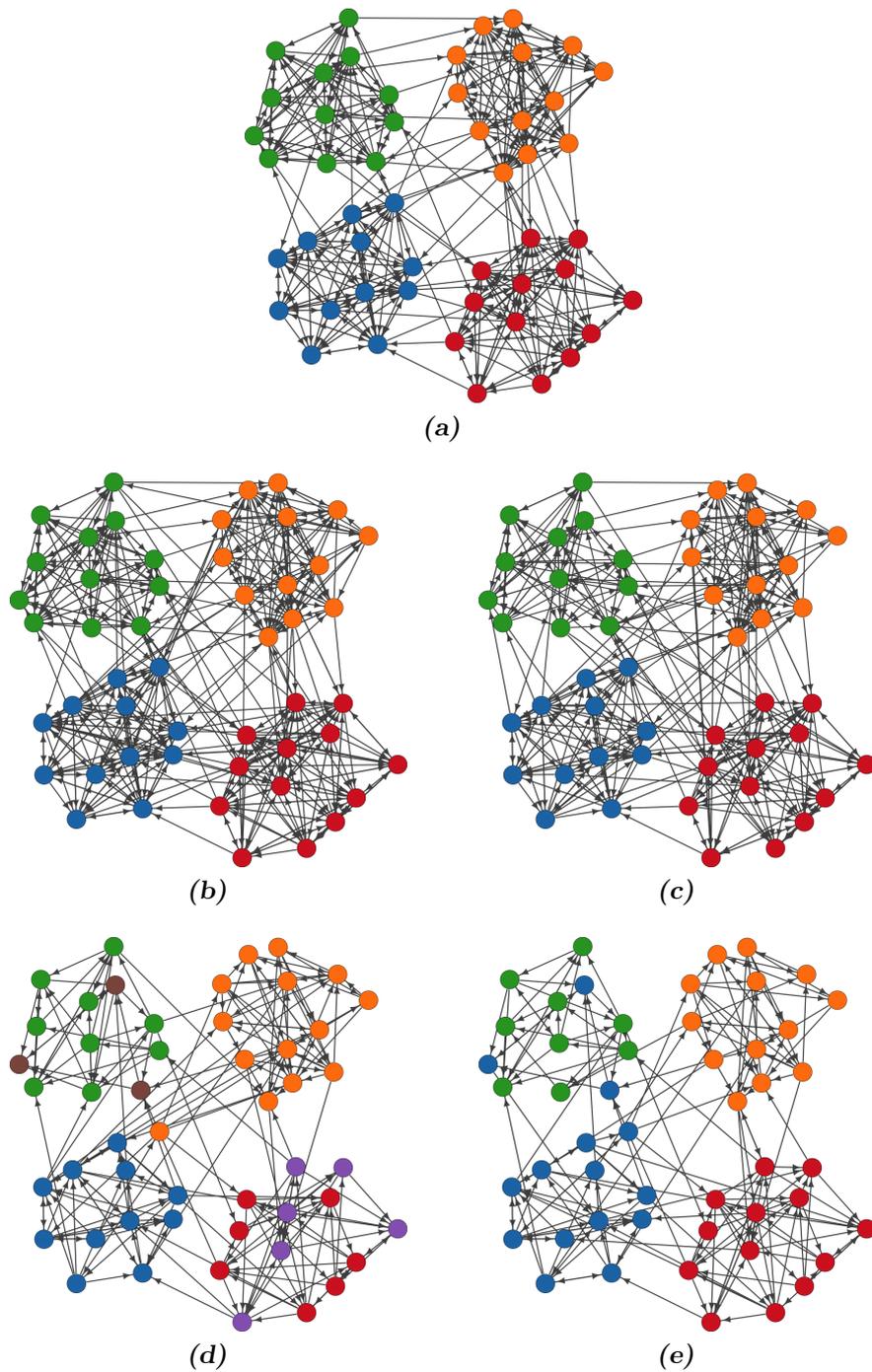
Figure 6.2 – **Network layouts for the adjacency matrix images of figure 6.1 (study I).** Two-dimensional node configurations were computed for the ground truth network using the algorithm of Kamada and Kawai [141]. Thereby obtained node coordinates were applied to the layout of all other networks. Nodes are color-coded with respect to their module affiliations (cf. figure 6.1) as detected by the algorithm of Blondel *et al.* ("Louvain"). (a) ground truth network ($|\mathcal{V}| = 50$, ground truth module affiliations were recovered 100%), (b) GCI computed network ($N_t = 3000$), (c) lsGCI computed network ($N_t = 3000$, 84% variance explanation), (d) GCI computed network ($N_t = 1000$), (e) lsGCI computed network ($N_t = 1000$, 86% variance explanation). The four network modules are visible as regions with high density of interactions. It can be seen again that the preservation of ground truth edge patterns and ground truth module structure depends on time series length and degree of dimension reduction. Modules display higher interaction densities in all cases and can be recovered to a large degree (cf. tables 6.2 and 6.4).
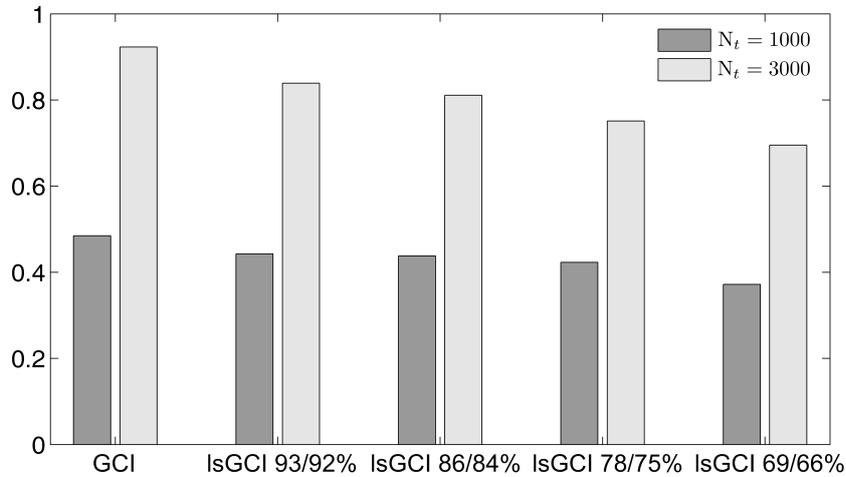
Figure 6.3 – **Assessment of edge pattern alterations in terms of Cohen's kappa (study I).** The agreement between edge patterns of a ground truth adjacency matrix and corresponding adjacency matrices identified by lsGCI and classical GCI approaches indicates the amount of edge pattern alterations caused by the lsGCI dimension reduction step. The ratios at the abscissa denote the proportion of retained components. As expected, the agreement of adjacency matrix entries depends on time series length and degree of dimension reduction. In particular for the case of the longer time series and lower degrees of dimension reduction the agreement ranges between high and adequate.

(a)   *(b)*   *(c)*
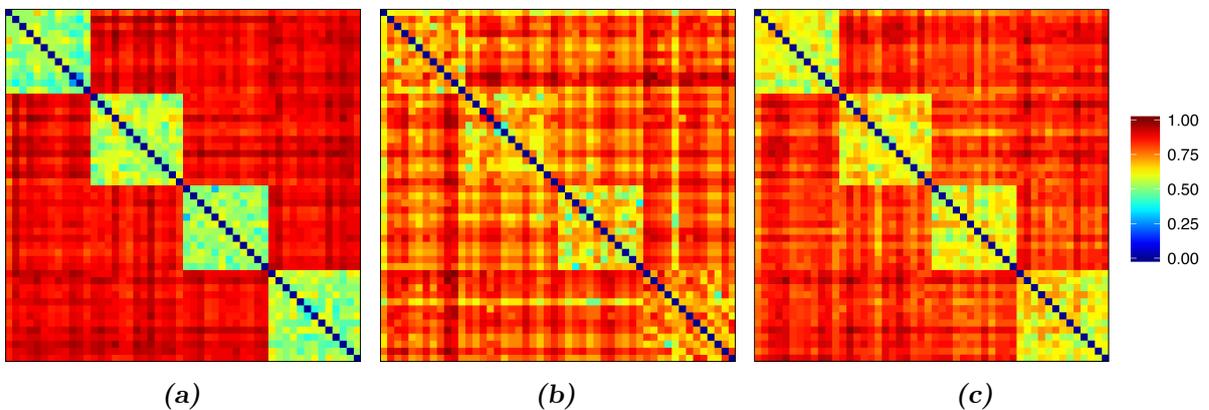
Figure 6.4 – **Pairwise adjacency relationship-based node distances (study I).** (a) Ground truth network ($N = 50$). Clear network module structure is reflected in small distances between comember nodes and large distances between nodes of different modules. (b) lsGCI network ($N_t = 1000$). Degradation of module structure is apparant by larger distances between nodes of the same modules and less pronounced differences in distances between nodes of flanking modules. (c) lsGCI network ($N_t = 3000$). With increasing length of simulated time series the module dissimilarities become similar to the ground truth network, hinting at an increasingly more clear-cut module structure that resembles the one of the ground truth networks, although inter-module node distances are reduced. These findings are consistent with the silhouettes calculated for the same data (cf. figure 6.5).
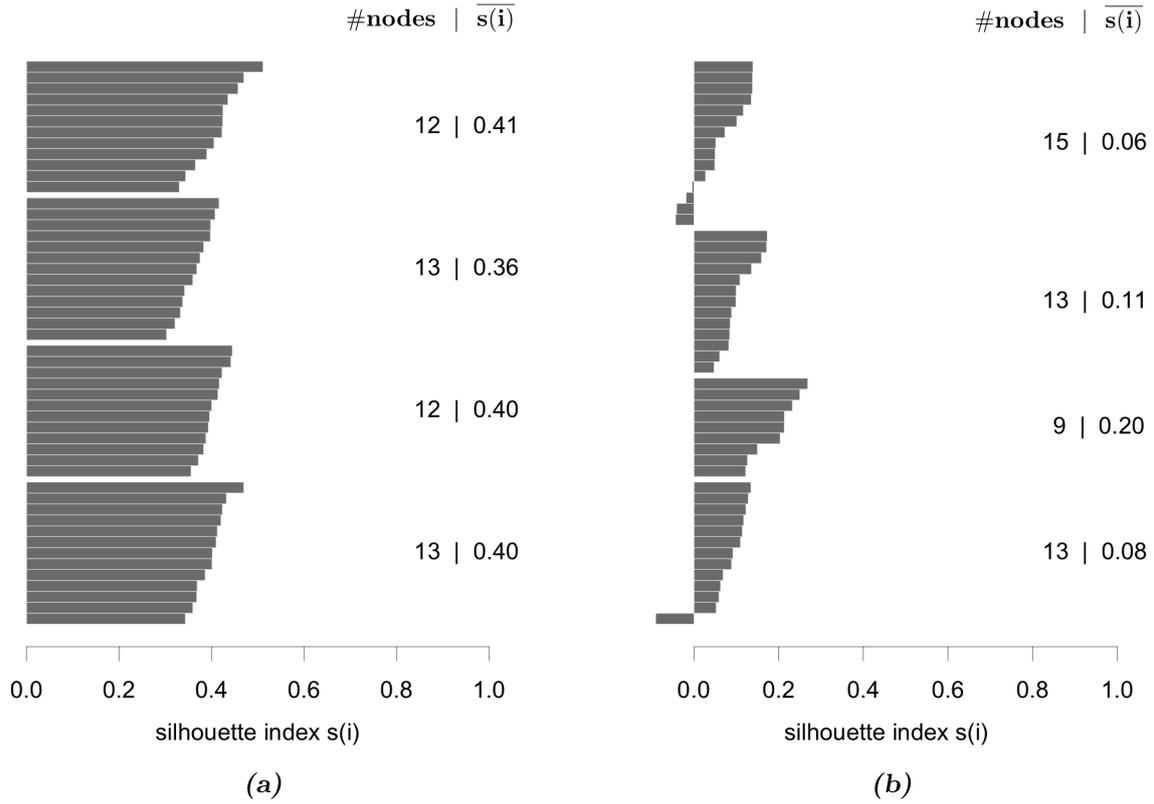
Figure 6.5 – **Silhouettes of identified module structure (study I).** The four silhouettes in each plot show which nodes are qualitatively fitting well into their assigned modules, which nodes are not showing a strong association with their module and are merely lying in between different modules and which ones are misclassified. Narrow and negative silhouettes are indicating the degree of artificial clustering of nodes to define modules. The adjacency relationship node distance measure of section 2.10 is used for calculating the silhouette widths of single nodes. (a) Ground truth network ($|\mathcal{V}| = 50$). (b) lsGCI network ($|\mathcal{V}| = 50$, $N_t = 1000$, 86%). (c) lsGCI network ($|\mathcal{V}| = 50$, $N_t = 3000$, 84%). The silhouette plot reveals the influence of dimension reduction and the length of underlying time series on the quality and recoverability of the four ground network modules. A relatively weak module structure that is less pronounced than the one of the ground truth networks is apparent in (b) for short time series length (cf. figure 6.4). Also note the different module sizes and the different average silhouette widths $\overline{s(i)}$ of modules.

116

The effect of dimension reduction on node degree sequences can be quantified by means of the L_1-metric (Manhattan distance) separately for in-degrees and out-degrees. Using the ground truth network of study I for comparison the degree sequences were found to be altered profoundly. The L_1 distance for $N_t = 1000$ and lsGCI-computed networks was in the range of 268 (93% variance explanation) to 287 (69% variance explanation), whereas for the respective GCI-computed networks it was 253. The L_1 distance for $N_t = 3000$ and lsGCI-computed networks was in the range of 66 (92% variance explanation) to 153 (66% variance explanation), whereas for the respective GCI-computed networks it was just 33. Again, it can be seen that the influence of different degree of dimension reduction on alterations of network topology are smaller than the effect of the chosen time series length.

In summary, the module preservation is very good, and the influence of the dimension reduction on module recoverability and module quality is much lower than one might assume by comparing edge patterns only (figure 6.3) or comparing graphical representations of adjacency matrices (figure 6.1).

6.3.2 Results of study II

In study II [273] network partitions were identified with the following algorithms:

1. Algorithm of Leicht and Newman [154] ("leading eigenvector", BD networks)

2. Algorithm of Blondel *et al.* [158] ("Louvain", BD networks)

3. Random walk algorithm of Pons and Latapy [162, 163] (BU networks)

4. Greedy algorithm of Clauset *et al.* [160] (BU networks)

5. Algorithm of Newman [152, 153] ("leading eigenvector", BU networks)

6. Potts spin glass based algorithm of Reichardt and Bornholdt [167] (BU networks)

7. Algorithm of Blondel *et al.* [158] ("Louvain", BU networks)

8. Algorithm of Rosvall and Bergstrom [168] ("Infomap", BD networks).

The results of the algorithm of Rosvall and Bergstrom were excluded, because for the investigated fMRI data, uncovered network partitions frequently consisted of one module only. All algorithms are described in chapter 2, section 2.8.

Edge pattern alterations caused by different degrees of dimension reduction and their ef-
fect on module structure can be seen in the adjacency matrix plots in figure 6.6. The
corresponding network layouts are shown in figure 6.7. Again, network module structure
is directly recognizable in the images of the adjacency matrices as diagonal blocks of dense
edge patterns. Due to dimension reduction, a number of intra-module edges are lost, whereas
spurious, false positive interactions between nodes of different modules are gained. As expected,
the preservation of ground truth edge patterns depends on the degree of dimension reduction.
Similar to the case of the smaller ground truth networks of study I, with respect to the edge
pattern alterations, an intermediate range of variance explanations seems to yield the best
results. Large deviations from ground truth module structure can be seen for dimension
reductions with high variance explanations (figure 6.6e and to some extent also evident from
figure 6.6d). In the case of 100% explained variance (which corresponds to the classical GCI)
two ground truth modules are fused into one module, inter-module edges connect original
ground truth comember nodes of the same module, while at the same time some ground truth
non-comember nodes are linked by intra-module edges. For dimension reductions within a
range of lower variance explanations the situation is not very different and similar assertions
hold true (figure 6.6b). Figure 6.7 shows edge pattern alterations from the perspective of
two-dimensional network layouts.

For a more objective assessment, edge pattern alterations in ground truth network ensembles
were quantified with several measures that combine information on network topology with the
classification of nodes, or measures that compare node partitions directly (see also chapter 2,
section 2.10). The various network characteristics that were evaluated cover many different
aspects of network module structure and give a coherent picture of ground truth network and
lsGCI network topology. Despite the information loss inflicted by the dimension reduction
step and its associated edge pattern alterations, we found that the module recoverability as
well as the quality of the identified network partitions was still good in the lsGCI networks,
as compared to their respective ground truth networks, even when considerably reduced.
This is particularly true for the range of explained variance that is of practical relevance.
In this analysis the *recoverability of module structure* was investigated with the following
characteristics that quantify the similarity between the detected partitions of the ground truth
networks and the ones detected in lsGCI-computed networks: the ratio of correctly classified

nodes, Rand index and adjusted Rand index, normalized variation of information, normalized mutual information and split-join distance. Network *module structure quality* was assessed with the following characteristics: partition edit distance, modularity, performance measure, coverage and overall average silhouette width. The boxplots in the appendix, figures B.1 - B.11 show the results of the analysis of the effects of dimension reduction on ground truth network module structure. Notably, for most considered module identification algorithms, we found that the percentage of correctly classified nodes in all lsGCI networks was adequately high, even though a reduction in comparison to the ground truth networks was noticeable. Depending on the module detection algorithm used and the amount of explained variance, the median percentage of correctly classified nodes for lsGCI networks was between 47% and 87.5%, whereas for the the ground truth networks it was in the range between 77% and 100%. The results for the ratio of correctly classified nodes are presented in figure B.1. As explained in section 2.10 on page 39, the ratio of correctly classified nodes potentially yields a distorted picture of the module detection results if the number of identified modules does not coincide with the number of modules in the ground truth network. The Rand index mitigates against this effect. It measures the similarity of the module structure detected in the ground truth networks with the one in the lsGCI networks. The boxplots in figure B.2 show that a large fraction of node pairs are either clustered together or are separated into different clusters in an identical fashion in the ground truth networks and lsGCI networks. Depending on the module detection algorithm used and the amount of variance explanation the median Rand index is in the range between 0.76 and 0.96, which is close to its maximum. For the chance-adjusted Rand index (figure B.3) the interquartile ranges are increased and mean values are much smaller. Like the ratio of correctly classified nodes and the Rand index, variation of information (figure B.4), mutual information (figureB.5) and split-join distance (figure B.6) are measures that depend only on the classifications of nodes in the two network partitions that are compared with each other. The results show that their values were still in line with the recoverability of node module affiliations, which means that the lsGCI networks were sufficiently similar to their ground truth networks from the perspective of module identification.

To further contrast the module structure of the ground truth networks with the one of the lsGCI networks I consider network characteristics that take into account features of network topology, as well as the module affiliations of nodes. Such characteristics can be expected to yield a particularly accurate picture of the degree to which the lsGCI network structure was impaired by the dimension reduction step of the lsGCI computation. The (partition)

edit distance of intra-module edges (figure B.7) was relatively high, reflecting the changes
of intra-module edges patterns in the lsGCI networks. I found that the values of network
characteristics like modularity (figure B.8), performance measure (figure B.9) and coverage
(figure B.10) were noticeably reduced in comparison to the ground truth networks. Using the
example of coverage values, it can be seen that a negative influence of increasing degrees of
dimension reduction on module structure exists (as shown in the boxplots from panels (E) to
(B)), but over entire network samples this influence turned out to be smaller than expected.
This confirms the initial findings obtained from inspecting the adjacency matrix plots. The
overall average silhouette width values (figure B.11) were surprisingly low. This can be in
part attributed to the somewhat short time series length ($N_t = 1000$). As shown in study
I, silhouette values greatly improved for $N_t = 3000$. In addition, due to stricter constraints
on the synthetic ground truth network model the ground truth module structure was less
pronounced as compared to study I. For the ground truth network data of study II the node
dissimlarity measure might not have been appropriate (see the discussion in section 2.10 on
page 45).

6.4 Application of the lsGCI approach to functional resting state MRI data — detecting large scale module structure

In both studies [61, 273] clinical data was analyzed in addition to simulated ground truth
networks. Thereby, lsGCI-derived functional networks were analyzed that describe interaction
pattens in spatially highly resolved fMRI data with $|\mathcal{V}| > N_t$ that cannot be handled by
classical GCI approaches. Network module structure in dichotomized networks was detected
using different community detection algorithms. Thereby, best results were frequently obtained
with the algorithm of Blondel *et al.* for directed binary networks (see also section 2.8). The
integer module affiliations of nodes that were returned by the detection algorithms were
matched optimally to ensure best possible comparability across data sets (see also chapter 3,
section 3.2) and mapped to unique colors and projected to the fMRI slice voxel masks. The
resulting images yield a visualization of functional segmentation in the fMRI recordings.

6.4.1 Results of study I

As can be seen in figure 6.8 that depicts one fMRI slice with projected module structure, the
affiliation of spatially distributed voxels to modules did not occur unsystematically. Therefore,

(a)

(b) *(c)*

(d) *(e)*
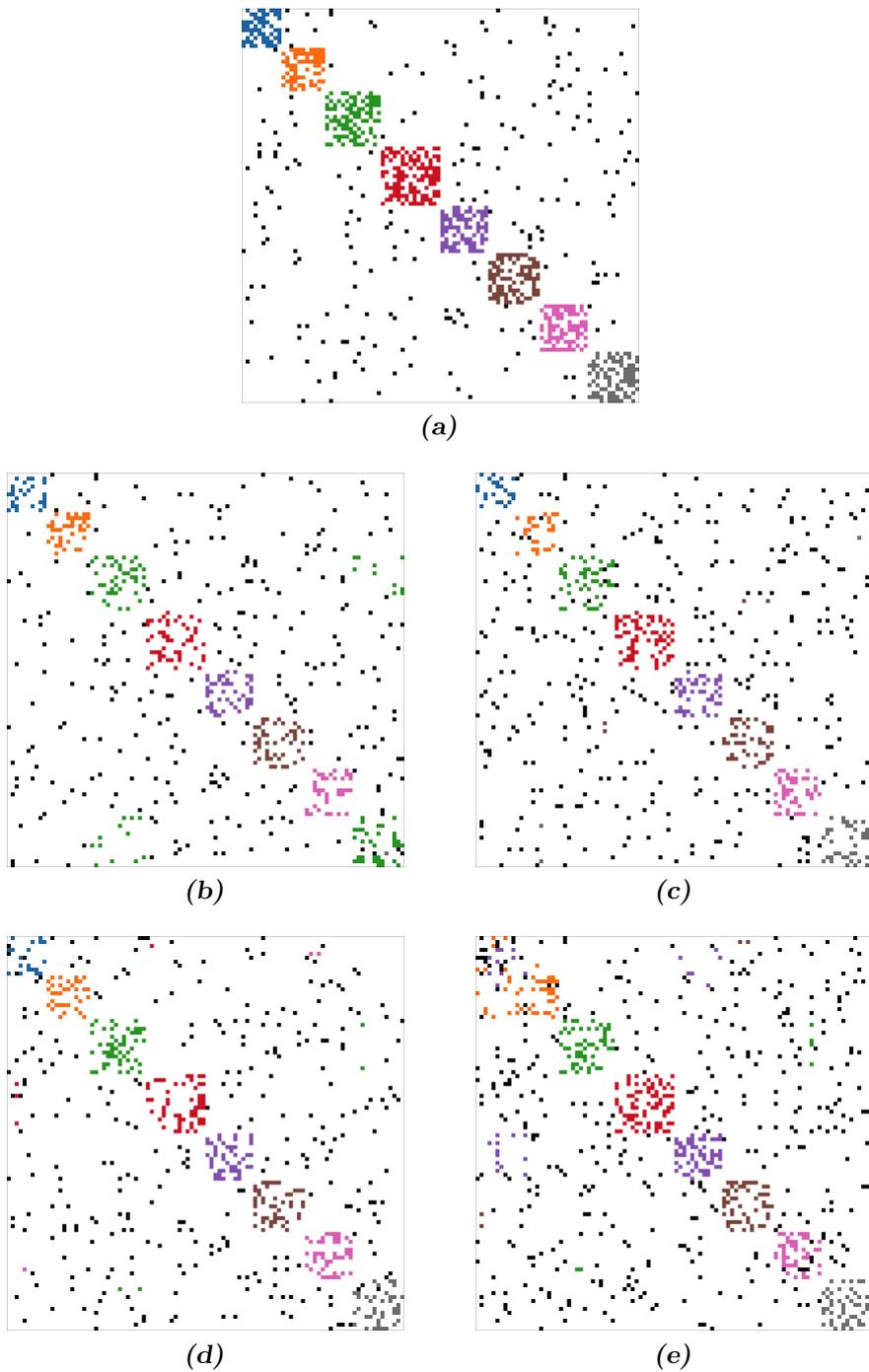
Figure 6.6 – **Edge pattern alterations in the adjacency matrices of (ls)GCI computed networks with different degrees of dimension reduction in comparison to an exemplar ground truth network (study II).** For all networks $|\mathcal{V}| = 100$. Networks (b)-(e) ($N_t = 1000$) are based on the network (a). Interactions are color-coded with respect to the module affiliation of their nodes (Potts spin glass algorithm [167]). Inter-module edges are shown in black. (a) ground truth network, (b) lsGCI computed network, 70% variance explanation, (c) lsGCI computed network, 80% variance explanation, (d) lsGCI computed network, 90% variance explanation, (e) GCI computed network, 100% variance explanation. Network modules are visible as blocks of dense edge patterns. In (b)-(e) inter-module edges are increased and intra-module edges are thinned out. As expected, the preservation of ground truth edge patterns depends on the degree of dimension reduction. For some of the (ls)GCI networks module structure alterations are visible by inter-module interactions between ground truth comember nodes (black pixels in colored blocks) and intra-module interactions between ground truth non-comember nodes (colored pixels outside blocks).
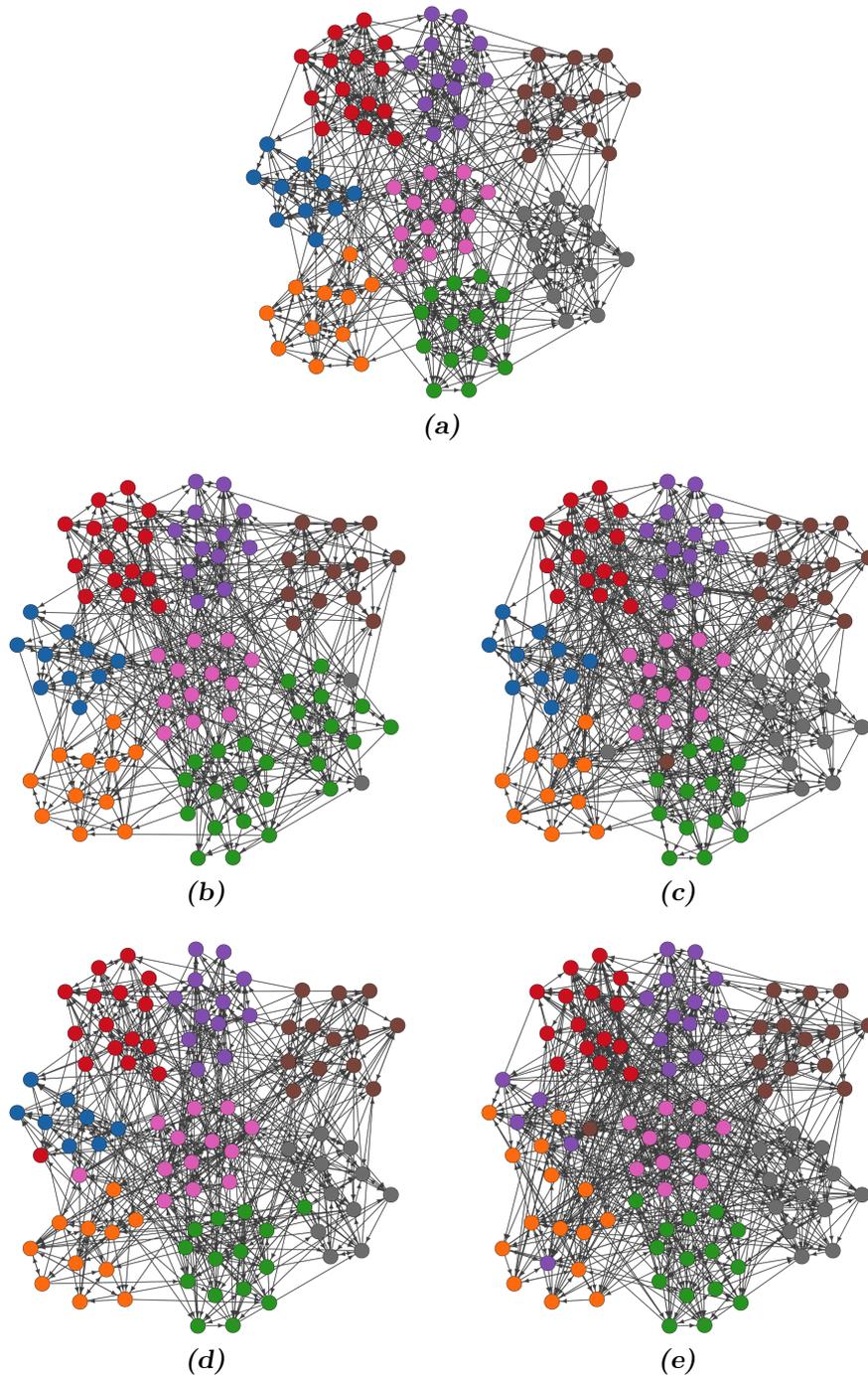
Figure 6.7 – **Network layouts for the adjacency matrix images of figure 6.6 (study II).** For the ground truth network the node configuration on the plane was found by the graphopt force-directed network layout algorithm [277]. To improve visual clarity (retain separation of modules) these node coordinates were used for networks (b)-(e), too. Nodes are color-coded with respect to their module affiliations (Potts spin glass algorithm [167], cf. figure 6.6). All networks have size $|\mathcal{V}| = 100$. Time series length for the lsGCI computed networks was $N_t = 1000$. (a) ground truth network, 100% correctly classified nodes (CCN), (b) lsGCI computed network, 70% variance explanation, 89% CCN, (c) lsGCI computed network, 80% variance explanation, 89% CCN, (d) lsGCI computed network, 90% variance explanation, 97% CCN, (e) GCI computed network, 100% variance explanation, 88% CCN. Network modules are visible as regions with a high density of interactions. The preservation of ground truth edge patterns and module structure depends on the degree of dimension reduction. In all cases modules can be recovered to a large degree (cf. figures B.1 to B.11).

similar to the situation of the simulated data set used in study I, it seems that the effect of
the embedded dimension reduction on the module structure is small enough so that module
affiliations can be recovered. Module structure identification results seem to be more robust
with respect to variations of network dichotomization thresholds than expected, although
the network topology is directly affected. Most likely this can be attributed to intra- and
inter-module edges being similarly influenced by threshold alterations. Also, it is remarkable
that the dimension reduction with the highest variance explanation does not necessarily result
in the most pronounced segmentation. This effect was already described for artificial data
[278] and can be observed here again. Variance explanations of around 80% were found to
result in similar identified module structures (figure 6.8).

6.4.2   Results of study II

Even more conclusive are the findings of study II. In this study five functional networks
of high spatial resolution that cover multivariate interactions in several slices of fMRI data
were analyzed. An example functional segmentation for the entire functional network sample
is given in figure 6.9, which shows the projection of identified module affiliations onto the
original fMRI voxel masks. Remarkably, for all subjects the 'voxel modules' closely followed
the conventional classification of the lobes of the brain and the arrangement of voxels with
the same color (nodes with the same module affiliation) was almost symmetric with respect
to the left and right hemispheres of the brain. The module structure projections of all five
analyzed subjects displayed a demarcation of the area of the precentral gyrus (primary motor
cortex) and postcentral gyrus (primary sensory cortex) to the frontal lobe and parietal lobe,
respectively. This module (colored in red) was particularly prominent in the images for subjects
E and C. In summary it can be stated, that the resulting module structure (derived based on
functional connectivity patterns) is very alike to anatomical structures for all five analyzed
subjects.

6.5   Discussion

In both studies network module structure was detected in synthetic networks and in real-world
functional networks that represent the interaction structure in spatially high-resolution fMRI
recordings. The synthetic networks were used to evaluate the feasibility and utility of the
new lsGCI approach for calculating multivariate interactions in high-dimensional time series
data, such as functional neuroimaging data, and for obtaining an informative functional
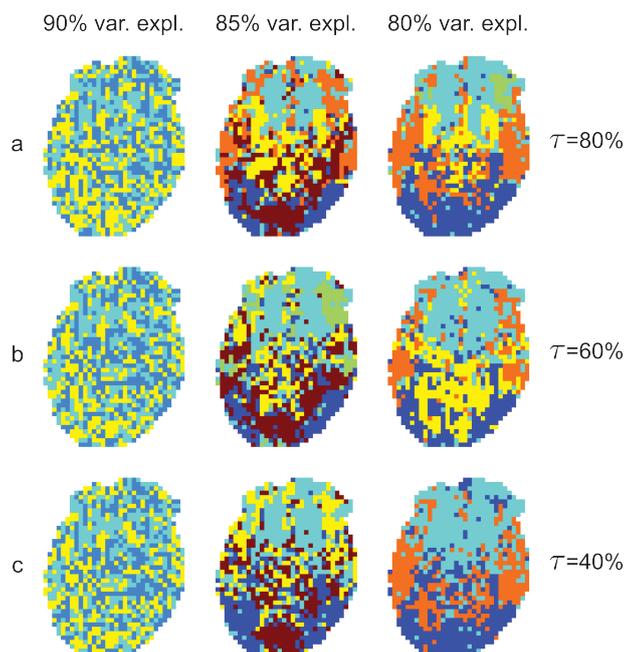
Figure 6.8 – **Functional segmentation of one resting state fMRI slice (study I).** Network module structure was identified in binary directed networks computed for various degrees of dimension reduction and threshold levels using the algorithm of Blondel *et al.* for directed networks. The ventricular system of the brain was not excluded from the fMRI voxel mask. The columns represents module structure of lsGCI networks with different levels of variance explanation (from left to right, 90%, 85% and 80% variance explanation). The rows show segmentation in dependence on different threshold levels used for network dichotomization (a: 80th, b: 60th and c: 40th percentile of lsGCI value distribution). For all thresholds, dichotomized networks consisted of one connected component. Module identities are color-coded. Despite cost-optimal matching of the node's module affiliations, there are few differences across networks due to the individual network partitions being unique. Thus, module location information has to be additionally considered for interpretation of the network sample.

segmentation of the resulting large scale binary directed networks. For this evaluation, the definiteness and quality of identified modules was quantified using a comprehensive analysis concept. The results of both studies indicate the existence of a range of appropriate levels of dimension reduction and network dichotomization thresholds, which preserve main features of network module structure in the resulting high-dimensional lsGCI functional networks, despite profound alterations of in-degree and out-degree sequences of nodes. This is of particular interest when the detection of functional similar, strongly interacting nodes (voxels in the case of fMRI data) and tracing changes of entire regions with similar connectivity characteristics (functionally segmented brain areas in the case of fMRI data) is the primary objective. For it, unfortunately, the interpretation of module structure projected back to fMRI slices is difficult, as it depends on prior physiological knowledge, definiteness and localization of identified
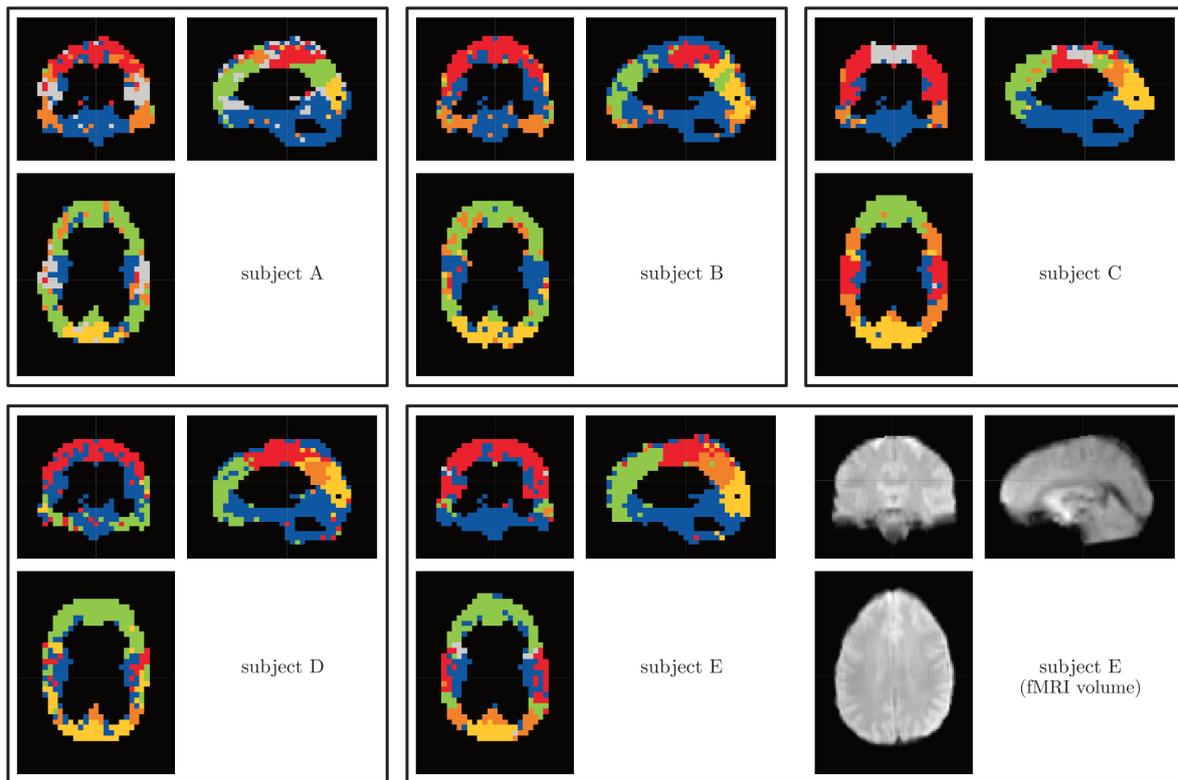
Figure 6.9 – **Functional segmentation of virtually full brain resting state fMRI voxel interactions of five different subjects (study II).** Network module structure was identified in binary directed networks computed with 85% variance explanation and dichotomized with the 95th percentile of the edge weight distribution as threshold for binary connections. All dichotomized networks consisted of one connected component. The algorithm of Blondel *et al.* for directed networks was used to obtain a functional segmentation of strongly interacting nodes. Module identities are color-coded. Despite cost-optimal matching of the node's module affiliations, there are few differences across networks due to the fact that individual network partitions are unique. Module location information has to be considered for interpretation of the network sample. As reference, one fMRI volume is shown bottom right (registered to the standard MNI152 template).

modules.

Like all other functional networks considered in this thesis, the investigated networks were initially edge-weighted and edge-complete. As a result, the biases and problems described in section 2.3 on page 16 affect the analysis and its results. For the two presented studies the strategy to analyze binary network instances obtained for a range of pre-defined percentiles of the edge weights as thresholds seems to give a good overview of the data and the robustness of its features with respect to the thresholding influence. It turned out that the choice of the global threshold parameter had a stronger influence on uncovered module affiliations than the choice of explained variance of the lsGCI computation. For appropriate percentages of

(a)

(b)

(c)

(d)
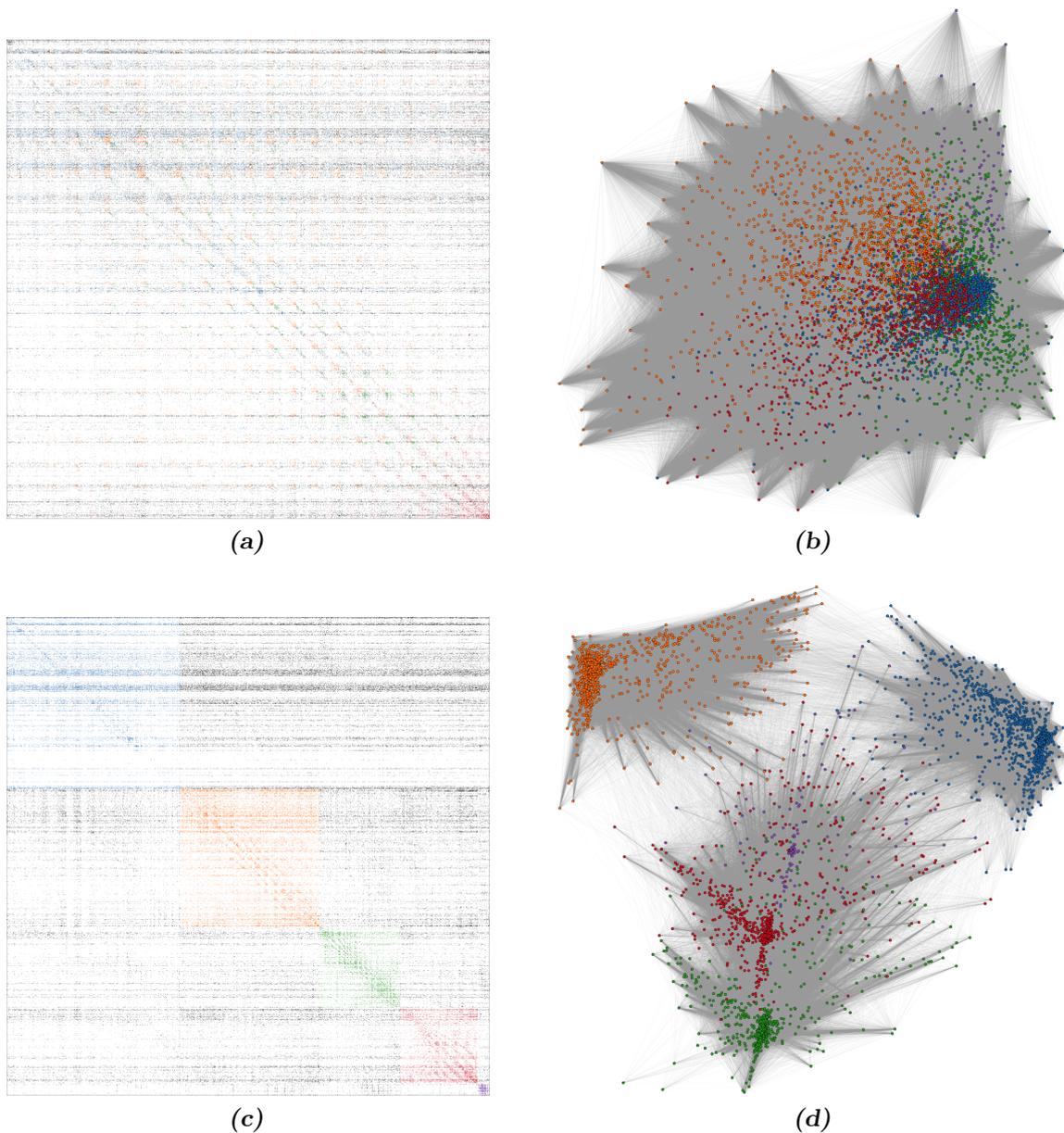
***Figure 6.10** – **Images of a binary directed lsGCI network of a HIV infected patient (study II).** Variance explanation after dimension reduction was 85%. The threshold for the dichotomization of the complete edge-weighted lsGCI network was the 95th percentile of the edge weights. The resulting network is connected and represents functional interactions between recorded voxel time series. Module structure was detected using the algorithm of Blondel *et al.* ('Louvain'). (a) Presentation of the original, unsorted adjacency matrix. The shown lsGCI network consists of 5,779 nodes (voxels) and 1,669,842 edges (voxel interactions). Black pixels indicate inter-module edges. Intra-module edges are color-coded with respect to module identity. (b) A 2D network layout of the unfiltered lsGCI network further reveals the high complexity of its interaction patterns. To compute this layout a metric multidimensional scaling algorithm (MDS) [279, 178] was applied to the shortest path matrix, which is used to define dissimilarities between nodes. Nodes are colored according to their module affiliation. (c) Image of a reordered version of the original adjaceny matrix in which comember nodes of each module are occupying subsequent positions in the rows and columns. Consequently, the interactions of comember nodes are centered around the main diagonal of the adjacency matrix and network module structure becomes apparent by the (color-coded) block-diagonal structure of pairwise interactions. It can be seen that the lsGCI network has a large number of inter-module interactions, which represents a demanding situation for module detection algorithms. (d) A filtered version of the original lsGCI network. Network filtering improves visual clarity and allows to show relevant structures. This particular network filtering retained all nodes, all intra-module edges and edges with largest betweenness values (99.5th percentile threshold) that are linking modules. Again, the MDS algorithm was used to compute node configurations in two-dimensional Euclidean space.*

explained variance it seems that intra-module edges and inter-module edges are altered by the dimension reduction step in a balanced way, so that information about the original module structure is preserved to a large degree. However, in study II for different reasonable settings of variance explanation and dichotomization threshold, visual assessment did not reveal large differences for the classification of the lobes of the brain given by identified module structure. Due to the lack of an available ground truth for real-world data, evaluation of the results is mainly based on plausibility arguments. In the case of our analyzed data, this plausibility was provided by a similarity between identified network modules and anatomical lobe classification.

In study II [273] several algorithms for module structure identification (see also section 2.8 on page 31) were applied to the data. The quality of the results was varied and depended directly on the used algorithm. Frequently, the 'leading eigenvector' algorithms of Leicht & Newman [154] and Newman [152, 153], the 'Louvain' algorithm of Blondel *et al.* [158] for directed and undirected networks and the 'Potts spin glass' algorithm [167] yielded clearly outlined network modules for our data. Although these four algorithms utilize structural information differently to identify node partitions into modules, they all yield plausible and comparable network partitions. On the other hand, some of the applied algorithms (random walk algorithm of Pons and Latapy, 'Infomap', greedy algorithm of Clauset) identified either only one module or only an implausible module structure consisting of a large number of small modules that were split and scattered across their network. This demonstrates the value of exploratory data analysis using different methods. One of the last remarks in this discussion section concerns the color coding of network module structure accross a set of networks. It is clearly desirable that modules with similar topological properties, composition or location are color-matched. However, as explained in section 3.2 on page 64 this objective cannot be achieved if modules occupying specific areas in one network are disaggregated into several smaller modules in another network. To correctly interpret the functional segmentation obtained by identified modules, spatial information, i. e. the location of modules in the brain should be considered in addition to the color codes. In fact, the spatial arrangement of modules contains more information than the color coding alone.

The finding that the module structure identified based on lsGCI functional connectivity patterns resembles anatomical structures, as seen in particular in the spatially high-resolution functional networks of study II, constitutes an encouraging result that might offer new perspectives for future research on applications that trace functional segmentation dynamics, i. e. the

changing composition of strongly interacting areas of the brain. Revealing such changes of functional segmentation over time and at a spatially large scale as the underlying functional connectivity patterns change in response to stimuli, medical treatment or as an accompanying symptom of disease certainly will lead to an improvement of our understanding of brain function.

In this context, the next chapter deals with the last application study that is presented in this thesis, in which a novel concept for dynamic module structure extraction in time-evolving EEG-derived networks is introduced.

Chapter 7

Extraction of dynamic module structure in time-evolving networks

I NFORMATION on the temporal evolution of functional brain network structure is assumed to give insights into time-variant neural processing and the dynamics of brain function. Sequences of consecutive snapshot networks, i. e. longitudinal networks, can be computed from the functional interactions of the brain recordings at selected time steps. As already mentioned, module structure constitutes an intuitive characterization of network organization, as it gives information about strongly interacting group of nodes. This information enables and facilitates an interpretation of underlying brain activity based on the functional connectivity data. In this regard it was shown that module structure in functional brain networks can be associated with cognitive performance and plays a role in human learning [15]. Therefore, dynamic module[1] structure is identified and analyzed in this application study to gain a better understanding of the time-variant changes[2] of the functional connectivity data in response to the execution of cognitive tasks. For each individual time step and its corresponding snapshot network [281, 172] a network partition can be computed, which represents the functional organization of network nodes, i. e. brain regions, in several distinct functionally segmented modules. Then, the network structure reconfigurations and the resulting dynamics in the module structure from time step to time step reflect how functional interactions between brain regions are

---

[1]Motifs are another mesoscopic structural feature that can be tracked over time, too [126].

[2]Network dynamics in general can be displayed on many levels. For example, qualitatively different nodes can be gained or lost, like bridging nodes [55] that mediate connections between different areas of the network and that have access to non-redundant information. Nodes can affect network topology by acquiring edges at different rates due to their intrinsic capacity, fitness and ageing [280, 281]. On the level of edges, new edges can introduce triadic closures that close structural holes, introduce important weak ties [52, 53, 54] or reciprocity of interactions.

redistributed in the course of the neural processing. In particular, module structure changes in response to stimuli and during solving cognitive tasks promises to be informative and to improve the understanding of normal and pathological brain function. Also, investigating the inverse situation is interesting, where the aim is to use the topological information of the network sequence to identify certain events in the time evolution, i.e. the time a certain stimulus sets in and is processed. Thus, analysis of sequences of time varying networks offers new perspectives on brain function and can reveal information that remains largely absent from the analysis of static, single network representations of functional interactions between brain regions. At present, network science applications are just beginning to see a shift from static network representations of systems to embrace the dynamic nature of interaction patterns in many real-world networked systems that constantly evolve over time. Contrary to the static version of the problem of identifying network modules, detecting and tracking network modules in series of network snapshots is still mostly uncommon in present research and is only recently being considered for network analysis [70, 282, 172, 281]. In particular in the field of computational neuroscience analysis of time-variant networks seems to be the next logical step, as brain recordings naturally have a temporal aspect and timestamped (functional) network data can commonly be made available.

Given the usually intricate and dense interaction structure, identification of network module structure in functional networks obtained from EEG data is a challenge in itself. Sophisticated pre-processing of network data is essential to uncover patterns in the otherwise inaccessible network structure that usually features high interaction density 'noise'. A main concern is the interpretability of obtained results with respect to neural information processing across the considered time steps. In chapter 3 section 3.6 I present a network data processing strategy that is capable to solve these problems, which I demonstrate by applying it to a sequence of snapshot networks that represent brain activity before and during the compensation of a major balance perturbation. It is known that motor activities that sustain balance are predominantly controlled by neural circuits in the spinal cord, the brainstem and the cerebellum. Strong evidence suggests additional contributions are made by the cerebral cortex [283, 284, 285]. The compensatory reactions of the cerebral cortex in response to a sudden and unpredicted balance perturbation to regain postural stability and to avoid a fall is currently only imprecisely understood [241]. Improving the understanding of cortical balance control and the identification of activated cortical regions and cortical adaptation effects in the context of balance recovery

is a topic of ongoing research and constitutes the background of the experiments on which the application study of this chapter is based.

7.1 Data description

Following, I state the main facts about the data used in this study. The experiment and the data acquisition is described in detail in [241].

7.1.1 Subjects and connectivity analysis

EEG data was recorded from 37 healthy male subjects, (mean (SD) age: 24.7 (3) years; body weight: 77.3 (8.1) kg). The recording was obtained from 32 electrodes according to the 10-10 international system. The electrical reference was located at FCz and the ground electrode was located at AFz. The experiment on which the study is based consisted in ten trials of transient unpredictable balance perturbations using a passively oscillating platform. Each perturbation trial, lasting 30 s in total, started immediately after the subjects positioned themselves on their dominant leg in the center of the platform. The task for the subjects during which the EEG was recorded consisted in balancing on this platform and keeping platform oscillations to a minimum after sudden perturbation, while having their hands at the iliac crest and focusing on a fixation cross at eye level. The platform was laterally deflected by 2.2 cm and fixed by an anchor in this position. In each trial the investigator randomly released the anchor, which induced a medial movement of the supportive platform. The perturbation onset was determined as the time point following anchor release at which the platform motion exceeded its mean oscillation level in medial-lateral or anterior-posterior direction during the last 5 s prior to anchor release by five standard deviations.

Next, I state the main features of the functional connectivity analysis, which was performed by members of my group. The multi-trial EEG data was fitted using the general linear Kalman filter in order to obtain time-variant MVAR model parameters (see also appendix B.1). The connectivity analysis based on the tvMVAR parameters was performed by means of direct DTF (dDTF) (see also appendix B.1.3). The dDTF was chosen because of its good frequency resolution that was able to clearly separate the specific frequency bands of interest, the theta and alpha band. This translates into frequency selective networks that can describe the connectivity related to different oscillations. For group analysis, the dDTF time-frequency maps are averaged over all subjects. The resulting dDTF interactions are in a narrow range

of very low values, which makes module detection in the weighted functional brain networks challenging. Therefore, using equ. 2.5 a rescaling of the values was performed to enhance the values of the strong interactions (see also section 2.5).

7.1.2   Samples of functional connectivity networks

For each subject 800 snapshot networks were available. For group analysis, the dDTF time-frequency maps of each interaction are averaged over all subjects. Such a map represents the strength of an directed interaction among two electrodes at each frequency bin and at each point in time. From these time-frequency maps, we extract the median value of a specific frequency band of interest at each time step to obtain weighted directed edge-complete functional brain networks. Thereby, the theta (5 Hz to 7 Hz) and alpha (9 Hz to 11 Hz) band were selected, because in these frequency bands the strongest activity was observed. Consequently, two sequences of snapshot networks were obtained for the entire group of subjects, one for the theta band and one for the alpha band. The time grid was defined with respect to time-frequency analysis results, i. e. with respect to characteristic amplitude changes in the time-frequency maps. Thereby, seven time steps and corresponding snapshot networks were preselected: 2 s, 3 s, 4 s, 5 s, 6 s, 7 s and 7.8 s. The last time step was determined by the width of the consensus data aggregation time window (see also section 3.6 on page 79). In the preprocessed EEG data, the stimulus occured at 3 s. Due to heavy artifact contamination four electrodes, that acted solely as sources of interactions (Fp1, Fp2, P7, P8) were excluded resulting in networks with 28 nodes. These nodes had minor influence on the tvMVAR model when comparing with a reduced model (that excludes these electrodes) but a strong influence on the detection of network modules. The resulting EEG recording sites used in this study are shown in figure 7.1.

7.2   Analysis of the network samples

The analysis concept outlined in chapter 3 section 3.6 was utilized to analyse the two network samples given by the group averaged connectivity in the theta and the alpha frequency bands. A global threshold based on pooling all 800 snapshot networks was computed using the 90th percentile. To aggregate and smooth connectivity information in the vicinity of investigated snapshot networks at preselected time steps, a time window length of 41 networks corresponding to 400 ms was selected in an exploratory fashion. This means, for each investigated snapshot network the connectivity information of 20 networks directly preceding it and 20 networks

directly succeeding it in the network sequence is combined with the investigated snapshot network itself to yield either a consensus network or a consensus partition, depending on the analysis strategy. Module structure in these aggregated interaction patterns at the predefined time steps was identified using the algorithm of Blondel *et al.* for weighted directed networks (see also section 2.8).
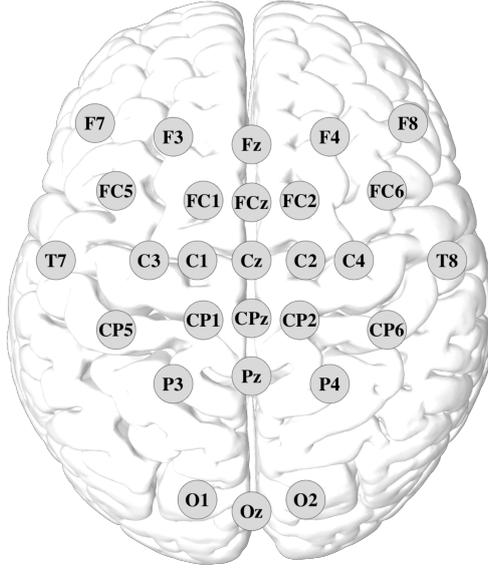
Figure 7.1 – **EEG electrode positions.**

7.3 Results and discussion

The main results of the application study can be summarized as follows. The global threshold computation using pooled edge weights of all 800 available networks leads to an identified module structure evolution that is associated with the temporal evolution of interaction strengths. In doing so, deviations from this overall information on interaction strengths are highlighted. By contrast, Pareto optimized network-specific thresholds take the temporal reconfiguration of connectivity structure at each considered time step into account. Therefore, by using different threshold definitions, different perspectives on the module structure dynamics are given. As expected, after the stimulus onset, i.e. after the transition to a more demanding balancing task, module structure starts to constantly reconfigure in both frequency bands, with no dramatic changes being observed between any two subsequent time steps. In particular, in the theta frequency band network sequence the module structures between 5 s and 8 s (2 s to 5 s after stimulus onset) remains relatively stable in spite of an increasing number

of threshold-exceeding interactions strengths (figure 7.2). The general localization and the time-evolution of the modules mostly corresponds to the expected neural processing for the balancing task.

In the *theta band* network sequence obtained by the global threshold an increase in laterality can be observed following the stimulus onset (figure 7.2a), whereas laterality is a consistent property of the module structure for the Pareto optimal thresholds (figure 7.2c). It is known that balancing tasks are associated with significant activation of the fronto-central and centro-parietal cortical regions. Specifically, when balance tasks became more challenging a significant increase in cortical theta activity was observed in the aforementioned regions [286]. The localization of the modules of the theta network (figure 7.2) demonstrate that these regions are involved in information processing during balancing on the unstable surface, i. e. after stimulus onset. In particular, the central electrodes C1, Cz, C2 and the centro-parietal electrodes CP1, CPz and CP2 (cf. EEG electrode layout in figure 7.1) act as the core of these modules for both thresholding approaches. Thereby, for both thresholding approaches the centrally located theta network module around CPz (module 4, ●) is already established after the stimulus onset and remains stable until the end of the analyzed period. The somatotopic representation of the feet and legs comprises the medial part of the primary motor cortex M1 (Cz) and that of the neck and trunk comprises the more lateral parts (C1, C2). Coinciding with the beginning of interactions related to M1, a frontal module (module 2, ●, electrodes FC1, FCz, FC2, Fz, F4) evolves over the supplementary motor area (medial part) and the premotor area (lateral part, area extrapyramidales). Associated functions are e. g. motor sequencing and planning as well as movement initiation and inhibition [287]. The corresponding alluvial diagram (figure 7.2b) shows that the module around CPz (module 4, ●) shows the highest degree of dynamics during balancing on the unstable surface. The application of the Pareto-optimal threshold to the theta band network sequence yields stable central modules around CPz, Cz (module 4, ●) and a widespread module which covers frontal and occipital regions (module 2, ●) (figure 7.2c). The main feature that stands out is the stability of the central module and of the modules, which are associated with the temporal lobes of the cortex (module 1, ● and module 3, ●).

Laterality of module structure is less pronounced in the *alpha band* network sequence (figure 7.3) as compared to the theta networks, where it is again more prominent in the Pareto thresholded networks (figure 7.3a). In the alpha networks, the occipital nodes O1 and O2 (cf. EEG electrode layout in figure 7.1) act as a source of interactions, which propagate predominantly in the directions from occipital to parietal and to centro-parietal areas. In the

alpha band network sequence for the global thresholding it can be seen that two occipital modules do already exist pre stimulus (figure 7.3a). These two occipital modules cannot be seen in the theta band, wheras in the alpha band they unite and the fusion module continues to exist after the 'alpha drop' at 4 s (reduced number and strengths of interactions). The alpha network for the global threshold (figure 7.3a) is characterized by strong interactions between the occipital electrodes (O1, Oz, O2, primary visual cortex) during the less demanding balancing task before the stimulus onset. The interaction strengths, the number of interactions to other electrodes and the number of modules slowly increases after the stimulus onset, i. e. the module structure gradually expands from occipital to frontal electrodes until the entire network is involved in the processing of the stimulus. The Pareto-optimal threshold reveals the existence of a large module (module 4, ●), which covers a high percentage of all network nodes, including the occipital region (figure 7.3c). Starting with the stimulus onset this module is constantly contracting, which can be seen in the corresponding Sankey diagram (figure 7.3d). This could mean that at later time steps the visual control has lost some of its relevance for the balancing task. Particularly notable is the co-existence of lateral and central modules which encompass large parts of the ventral and dorsal pathways of visual information processing [288, 289, 290], where the main direction of interactions is clearly from the occipital to the temporo-parietal cortical areas. This finding indicates that both the ventral and the dorsal visual system might be involved the optimization of balancing control [291]. Such an interpretation is supported by the results based on the application of the Pareto-optimal threshold.

The results demonstrate the usefulness of the proposed module tracking approach, which can also be directly applied to functional brain networks obtained with different recording modalities, e. g. fMRI, or to longitudinal networks from entirely different domains. Information about dynamic module structure can be complemented with an analysis of time-variant network characteristics considering all investigated and aggregated time step networks. The quality and stability of resulting dynamic network modules and the amount of volatility between time steps, i. e. the amount of short-lived module structure changes are depending on the choice of investigated time steps and the time step window size. In addition to the proposed workflow (see also chapter 3 section 3.6), the stability of the module structure under minor perturbations of the network data could be analyzed [292, 179, 243], e. g. by using the module structure quality characteristics presented in chapter 2 section 2.10. This information could be beneficially used to complement the selection of investigated time steps, which was based
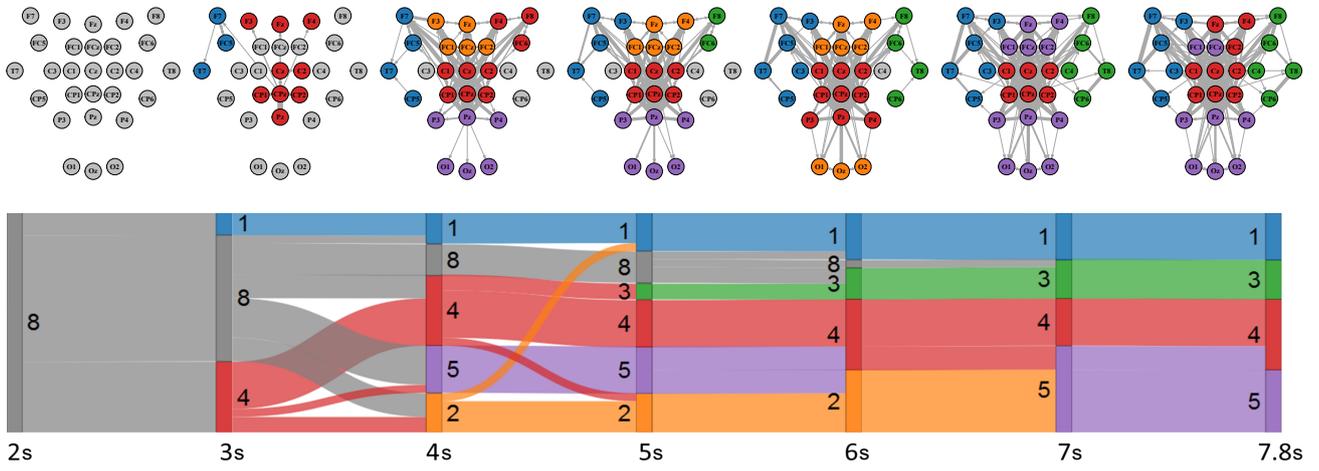
(a) Global threshold (90th percentile)
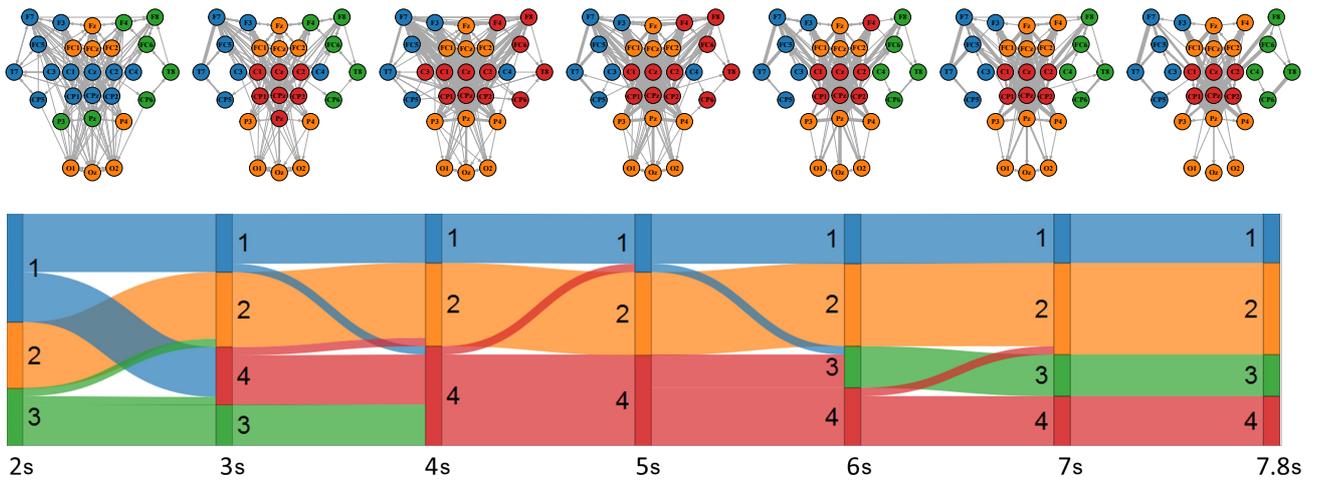
(b) Pareto optimal thresholds

Figure 7.2 – **Dynamic module structure of the *theta band* networks.** The spatial location of module structure changes is shown in the network plots according to the EEG electrode layout, whereas the dynamics, i. e. the flow of node membership, are shown in the Sankey diagrams. Thus, changes of information processing in the brain can be identified from its spatio-temporal module structure changes. Results for the global threshold (90th percentile, consensus networks) are shown in (a). Results for the Pareto optimal thresholds (with consensus clustering) are shown in (b).
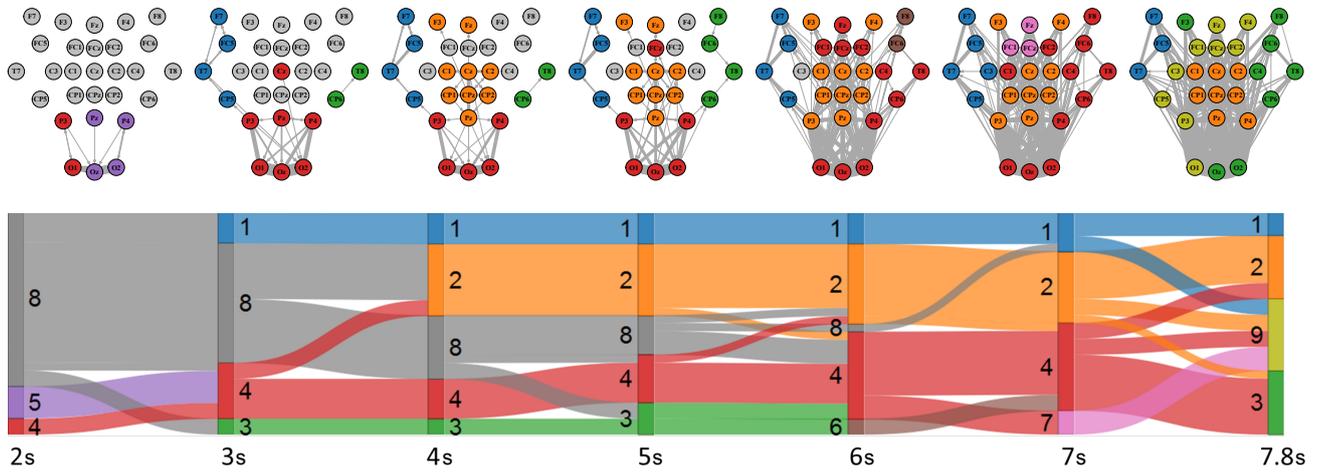
(a) Global threshold (90th percentile)
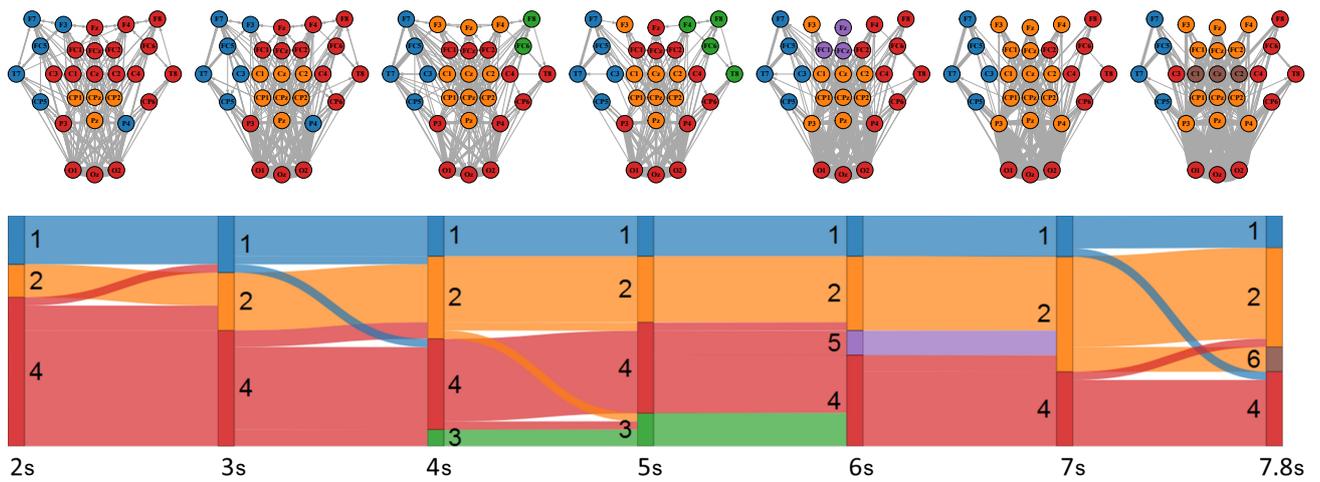
(b) Pareto optimal thresholds

Figure 7.3 – **Dynamic module structure of the *alpha band* networks.** The spatial location of module structure changes is shown in the network plots according to the EEG electrode layout, whereas the dynamics, i. e. the flow of node membership, are shown in the Sankey diagrams. Thus, changes of information processing in the brain can be identified from its spatio-temporal module structure changes. Results for the global threshold (90th percentile, consensus networks) are shown in (a). Results for the Pareto optimal thresholds (with consensus clustering) are shown in (b).

on time-frequency analysis results. Such a perturbation analysis might be used to devise a method for automatic time step selection. In a line of thought related to the perturbation analysis a bootstrap [57] method was proposed to identify significantly clustered module cores in weighted directed networks that might be tracked from time step to time step [159]. Tracking significantly clustered module cores is supposed to distinguish real trends in the time evolution of the module structure from noisy data and random fluctuations that act to insignificantly associate some nodes to their modules. The utility of this otherwise powerful approach for EEG-derived noisy functional networks is unclear, because a too large number of nodes from the already small networks might be found to be non-significantly clustered and excluded from the module tracking. Dynamic module structure might also be identified in the framework of multislice or multilayer networks [293], where the network configuration at each time step corresponds to one slice and multiple slices (adjacency matrices) are coupled by edges that connect each node with itself in neighboring slices [294, 295]. Despite the significantly increased computational costs for the optimization of the associated multislice modularity as compared to standard modularity optimization, this approach or related approaches might represent a potentially interesting future avenue of research. Finally, in the experimental settings typical for computational neuroscience in which group data is analyzed, sets of individual time-varying network sequences have to be analyzed if one does not want to consider grand mean networks of a network sample, as was done in this study. This potentially generates a large amount of information that has to be dealt with, in particular high temporal variations of identified module affiliations. The complexity of the results would further increased if potentially interesting overlapping module structure is considered, where individual nodes can be members of more than one module. Thus, taking all these considerations into account, it can be said that the proposed workflow strives a good balance between extracting important information about the brain activity and strongly interacting cohesive brain regions in the form of network modules on one side and on the other side speed of computation and simplification of analysis results, which have to remain interpretable.

Part IV

Concluding remarks and perspective

Chapter 8

Conclusions and future research directions

THIS thesis investigated cognitive processes and brain function on the basis of functional connectivity analysis from a network perspective. The inherent network structure of functional connectivity patterns naturally suggests the role of network analysis approaches as one of the most suitable ways of gaining an improved understanding of functional connectivity. In dependence on the research aims and the experimental background of the network data, various adapted preprocessing techniques and analysis methods were used to extract relevant information from the investigated network samples. The topology of the sample element networks encodes the different ways information is transferred and processed among and within neural structures of the brain. This suggests the importance of functional connectivity analysis for gaining an deeper understanding of cognitive processes and brain function. The methodological developments described in this thesis, although being generally applicable to networks from different domains, were tailored towards the analysis of functional brain networks and in particular samples of such networks. These network samples can basically be analyzed on three different topological scales with respect to the level of topological details considered for the analysis: the microscopic scale, mesoscopic scale and macroscopic scale. For the work presented in this thesis, network characteristics of all three topological scales were covered. With the clinical research questions for each application study in mind, a particular focus was put on the mesoscopic scale of network topology, which is typically accessed by network decompositions into functionally important recurring or unique substructures, like building block subnetworks (motifs) or important node sets (e. g. modules). If longitudinal

network data for the sample element networks is available, time-variant analysis is promising to yield additional valuable information about the underlying recorded neural processing, e. g. reconfiguration of connectivtiy patterns in response to a stimulus. In this regard, one of the presented application studies taps into the potential of time-variant network analysis by extracting dynamic module structure and revealing stimulus-dependent information about the neural processing that would otherwise not have been available. In this work I described several new methodological developments whose main ideas and potential improvements are:

- *Motif detection in samples of binary directed networks with pairwise different node labels.* This approach makes it possible to detect functionally relevant small subnetworks that occur significantly often within a network sample. Such subnetworks, i. e. motifs, are deemed as characteristic interaction patterns. I demonstrated that they can be used for interpretations of network function and to distinguish groups of networks. Possible improvements of this approach relate to the involved network null models. Open questions remain with respect to the design of null model networks that fit given network data. In particular, the analytical calculation of subnetwork significance has to be improved.

- *Cost optimal matching of module affiliation labels and a fuzzy matching extension.* To analyse different network partitions, e. g. the partitions of sample element networks, and to improve the visual representation of module structure the module affiliation labels of nodes have to be matched. This problem seems to be overlooked in the literature, supposedly because analysis of single networks is still predominant in research. A few further improvements seem currently conceivable. One of them has to do with the handling of special cases with respect to the selection of the reference network partition. For example, it is possible that the reference network has particularly many modules and a poor module structure quality, resulting in bad (albeit cost optimal) matchings. In such a case, another reference network should be selected. Currently, it is possible to manually exclude networks from the matching process. In this regard, an algorithmic approach for the quantification of the quality of the matchings would be desirable. This could pave the way for selecting optimal subsets of partitions to be matched that could further improve the quality of the whole matching.

- *Comprehensive analysis strategies.* Analysis of network data usually requires suitable combinations of diverse techniques for preprocessing the data, network type conversions and the actual characterization of network structure. Consequently, for all application

studies several established preprocessing techniques and analysis approaches were integrated with my own developments to yield an application-specific analysis workflow. Most likely in some cases different workflows and combinations of approaches could have been successfully used, too. Future methodological developments should be incorporated into existing workflows.

- *Simulating networks with pre-defined module structure.* The presented algorithm enables the simulation of directed benchmark networks with known module structure for a range of different parameter settings, which define different aspects of the resulting module structure. Networks that had been generated with this algorithm were used for testing of module detection algorithms and for MVAR time series generation to evaluate a method for functional connectivity computation (lsGCI). In the future, this algorithm should be extended to generate weighted directed networks with ground truth module structure.

- *Computing edge weight thresholds using a multi-objective optimization approach.* This approach allows to compute objective thresholds for filtering edges so that the resulting non-complete binary or weighted networks exhibit their most strongly marked module structure. For it, a multi-objective optimization of module structure quality vectors obtained at different global threshold levels is performed. Defining optimal edge weight thresholds is still an open research problem, as is the case for the quantification of thresholding-induced biases on the resulting network topology. However, with respect to the underlying module structure, this approach might represent a step into the right direction. Finding ways for obtaining local thresholds for individual regions inside a network would likely further improve the situation.

With the help of these approaches several network samples were investigated in four major studies of different clinical background. It seems that one of the most limiting factors with regard to the applicability of network science approaches for functional brain network analysis lies in the ability to correctly interpret the findings with respect to the observed neural processing, for which additional knowledge about neurophysiological processes is essential. This illustrates the central problem of neuroscience that often there is no good way to evaluate a theory or a model [296]. Functional connectivity measures offer only a limited view on the underlying neurophysiology and networks constitute only models that approximate a real system. For the case of networks derived from EEG sensor space data, another source of inaccuracy is given by the presence of noise and volume conduction effects, i.e. the fact that

the exact neural signal sources are not known, as the recorded brain activity is a superposition of different source activities, which limits interpretations of functional connectivity patterns [297]. Consequently, much work remains in the area of relating functional network analysis results to cognition and neural information processing [3]. On the part of network approaches, major future advancements for the investigation of time-variant networks and network samples in the form of multilayer network [293] modeling and analysis can be anticipated. The network analysis approaches presented in this thesis can principally be applied for the analysis of network from different domains. This generality with respect to the investigated system is a great advantage of thesse methods. Natural examples of networks from other domains are social networks, which represent social contacts among individuals or social groups, or metabolic networks. Metabolic networks represent directed metabolic pathways in two ways. The first one is given by modeling the metabolites as nodes and the enzymes as edges that connect nodes if the enzyme catalyzes their conversion. The second way of representation is given by modeling the enzymes as nodes that are connected by directed edges if the product of one enzymatic reaction forms the substrate for the other enzyme. However, other metabolic network representations and analysis approaches than the structural ones covered in this thesis seem to have received more attention. Examples are the stoichiometry constraint-based analysis framework [298] and metabolic models, e. g. elementary flux pattern analysis for the prediction of novel biochemical pathways [299].

Overall, despite some inevitable limitations I think the appeal and the utility of network analysis techniques for revealing features of functional connectivity and monitoring their alterations in health and disease, in different neurophysiological states and during different modes of neural processing became apparent in this thesis. Representing functional connectivity as networks and trying to gain a deeper understanding of brain function by analyzing these networks with suitable approaches likely constitutes the currently most powerful and most elegant way of accessing brain activity.

Appendices

Appendix A

List of publications

This thesis is based on the following peer-reviewed journal publications and conference proceedings.

- C. Schmidt, T. Weiss, C. Komusiewicz, H. Witte, and L. Leistritz, "An analytical approach to network motif detection in samples of networks with pairwise different vertex labels," Computational and Mathematical Methods in Medicine, vol. 2012, no. 395, pp. 1–12, 2012. [97]

- C. Schmidt, T. Weiss, T. Lehmann, H. Witte, and L. Leistritz, "Extracting labeled topological patterns from samples of networks," PLoS ONE, vol. 8, no. 8, p. e70497, 2013. [156]

- C. Schmidt, B. Pester, M. Nagarajan, H. Witte, L. Leistritz, and A. Wismueller, "Impact of multivariate Granger Causality analyses with embedded dimension reduction on network modules," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2797–2800, 2014. [61]

- M. E. Tivarus, B. Pester, C. Schmidt, T. Lehmann, T. Zhu, J. Zhong, L. Leistritz, and G. Schifitto, "Are structural changes induced by lithium in the HIV brain accompanied by changes in functional connectivity?," PLoS ONE, vol. 10, no. 10, p. e0139118, 2015. [266]

- K. Schiecke, C. Schmidt, D. Piper, P. Putsche, M. Feucht, H. Witte, and L. Leistritz, "Assignment of empirical mode decomposition components and its application to biomedical signals", Methods of Information in Medicine, vol. 54, no. 5, pp. 461–473, 2015. [231]

- C. Schmidt, B. Pester, N. Schmid-Hertel, H. Witte, A. Wismüller, and L. Leistritz, "A multivariate Granger Causality concept towards full brain functional connectivity," PLoS ONE, vol. 11, no. 4, p. e0153105, 2016. [273]

Appendix B

Supplemental information for functional brain network analysis

B.1   Methods for functional brain network identification

This section gives a review of the functional connectivity estimation methods that were used to compute the networks whose algorithmic analysis is the foundation of this thesis.

B.1.1   Granger Causality Index and large scale Granger Causality Index

The concept of Granger Causality encompasses various approaches for the investigation of directed interrelations between time series. A popular Granger Causality [300] approach is based on the principle of mutual predictability, i.e. the notion that a process variable affects another process variable, if the knowledge of the former variable helps improving predictions of the latter variable [301]. More specifically, an univariate time series Y_i (e.g. the recording at a specific EEG electrode or one specific fMRI voxel time series) is said to Granger-cause another univariate time series Y_j of the same multivariate process if the knowledge of the past of Y_i leads to a significant improvement in the prediction of Y_j. Therefore, Granger causality measures directional (unreciprocal) dependencies between the pair of time series Y_i and Y_j and can be applied to identify directed functional connectivity in neural time series data [8] where it can be interpreted as the existence of some underlying information flow from the brain area corresponding to time series Y_i to the brain area related to Y_j [58]. Linear multivariate autoregressive (MVAR) time series models, which gives a quantification of prediction errors, are most commonly used for the identification of Granger causality. A p-th order $|\mathcal{V}|$-variate

autoregressive model is formally given by

$$\boldsymbol{Y}(n) = \sum_{r=1}^{p} \boldsymbol{A}^r \cdot \boldsymbol{Y}(n-r) + \boldsymbol{E}(n), \quad n = p+1, \cdots, \mathrm{N}_t \tag{B.1}$$

with state vectors $\boldsymbol{Y}(n) \in \mathbb{R}^{|\mathcal{V}|}$ for N_t available temporal samples, AR model parameter matrices $\boldsymbol{A}^r \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$ and a zero mean uncorrelated noise process $\boldsymbol{E}(n) \in \mathbb{R}^{|\mathcal{V}|}$.

Given an observed multivariate time series $(\mathbf{y}(n))_{n=1,\dots,\mathrm{N}_t}$, which consists of all $|\mathcal{V}|$ observed univariate time series $(\mathbf{y}_k(n))_{n=1,\dots,\mathrm{N}_t}$, in practice the model order p may be determined according to the procedure outlined in [302], which consists of applying the general AIC criterion [303] for the calculation of p and subsequent fine-tuning to match the AR-related, parametric spectrum with the Fourier power spectrum. To estimate the AR parameters the MVAR model is fitted to the observed time series $(\mathbf{y}(n))_{n=1,\dots,\mathrm{N}_t}$ subject to minimizing the sum of squared model residuals [301]. This yields an estimated multivariate time series $(\hat{\mathbf{y}}(n))_{n=1,\dots,\mathrm{N}_t}$ and corresponding model residuals given by $\hat{\mathbf{e}}(n) = \hat{\mathbf{y}}(n) - \mathbf{y}(n)$. For the case of a time-variant MVAR process [302], where the value of the autoregressive coefficients $\boldsymbol{A}^r(n)$ depends on n, the model parameters can be estimated e.g. by a Kalman filter approach [304].

To quantify the influence of the obscured time series y_i on the time series y_j the MVAR model is also fitted to a restricted time series $(\mathbf{y}^{i-}(n))_{n=1,\dots,\mathrm{N}_t}$ obtained by excluding all data on y_i, which yields restricted model residuals $\hat{\mathbf{e}}^{i-}(n)$. With the covariance matrices $\mathbf{C} = \mathrm{cov}(\hat{\mathbf{e}})$ and $\mathbf{C}^{i-} = \mathrm{cov}(\hat{\mathbf{e}}^{i-})$, a Granger Causality Index (GCI) that quantifies the directed influence from y_i to y_j is defined by

$$\gamma_{i \to j} = \ln\left(C_j^{i-}\right) - \ln\left(C_j\right), \tag{B.2}$$

where C_j^{i-} and C_j denote the j-th diagonal entry of $\mathbf{C}^{i-}$ and $\mathbf{C}$, i.e. the variance of the model residuals associated with y_j in the modified and in the original system. GCI takes on nonnegative values.

Due to a limited availability of temporal samples and the resulting problem of estimation equations in MVAR model fitting becomming under-determined, classical multivariate Granger Causality analyses, as described above, are commonly restricted to spatially low-dimensional data, which requires a pre-selection or aggregation of time series as a preprocessing step [273]. With this, only a reduced amount of the available spatially distributed information can ulti-

mately be exploited. Even approaches that are geared towards reducing spatial dimensionality by a suitable coordinate transformation, such as principal component analysis (PCA) [305] or independent component analysis (ICA) [306] do not constitute a fully adequate replacement for unfavorable time series subset pre-selection or time series aggregation approaches, because identified interactions between a few principal or independent components cannot be readily transferred back into the original high-dimensional space, which severely limits the interpretation of functional connectivity.

With the large scale Granger Causality (lsGCI) index [278, 61, 273] spatially high-dimensional data is incorporated into the connectivity analysis. This is achieved by an embedded PCA data dimension reduction step, which is followed by standard Granger Causality connectivity analysis, i.e. the MVAR model parameter estimation, is based on low dimensional time series that were obtained by PCA. Subsequently, the estimated low dimensional MVAR model is projected back into the original high dimensional space by an orthogonal transformation. The back projected model is then used to calculate residuals with respect to the original high-dimensional time series for defining pairwise standard GCI interactions between network nodes. The preservation of high spatial data dimensionality may result in functional networks that consist of several thousand nodes that are connected by millions of edges, yielding a representation of brain connectivity phenomena that span many spatially distinct brain regions.

B.1.2 Partial directed coherence (PDC) and generalized PDC

Partial directed coherence (PDC) [307] can be seen as a frequency domain counterpart of Granger Causality, which works in the time domain, and has become a very prominent tool for the quantification of functional connectivity in EEG recordings [58]. It yields a linear frequency-selective quantifier of the multivariate relationship between simultaneously observed time series [308]. The PDC can be derived based on Fourier transformed MVAR parameters. With a $|\mathcal{V}|$-dimensional identity matrix $\mathbf{I}$ the AR parameter matrix is given by

$$\mathbf{A}(f) = \mathbf{I} - \sum_{r=1}^{p} \mathbf{A}^r \cdot e^{-2\pi f r \sqrt{-1}} \tag{B.3}$$

The PDC, which quantifies the degree of connectivity, i.e. the relative interaction strength, from Y_i to Y_j at a normalized frequency f, is defined as follows

$$\pi_{i \to j}(f) = \frac{\mathbf{A}_{ji}(f)}{\sqrt{\sum_{d=1}^{|\mathcal{V}|} |\mathbf{A}_{di}(f)|^2}} \tag{B.4}$$

Thus, $\pi_{i \to j}(f)$ accounts for the ratio of the coupling strength of the interaction of Y_i with regard to Y_j and the coupling strengths of all interactions of Y_i. The PDC takes on values in the $[0, 1]$ interval [307].

The generalized PDC (gPDC) is an extension of the PDC that can be suitably applied in the presence of unbalanced predictive modeling errors if signal amplitudes differ to a large extent as it is more robust in estimating imprecisions associated with finite time series samples [308]. It is given by

$$\pi_{i \to j}^{(g)}(f) = \frac{\frac{1}{\sigma_j} \mathbf{A}_{ji}(f)}{\sqrt{\sum_{d=1}^{|\mathcal{V}|} \frac{1}{\sigma_d^2} |\mathbf{A}_{di}(f)|^2}} \tag{B.5}$$

where σ_d^2 refers to the variance of the d-th MVAR model residual. Weighting the AR coefficients with the variances of the model residuals $\boldsymbol{E}$ (MVAR prediction errors at frequency f) reduces the influence of particular signal amplitude amplifications.

B.1.3 Direct directed transfer function (dDTF)

A modification of the directed transfer function (DTF) was proposed in [309]. The so-called direct DTF (dDTF) is a measure to estimate direct causal relations between signals using the transfer matrix of the system in the frequency domain. The method combines information from the partial coherence function with information about the direction of influence. One advantage of the dDTF is that the resulting time-frequency maps have a good frequency resolution so that specific frequency bands can be separated with high accuracy. As a drawback, (d)DTF is susceptible to the influence of indirect interactions.

B.2 Simulation of MVAR time series that mirror ground truth network connectivity patterns

Publications [61] and [273], which I coauthored, rely on extensive simulation studies, in which lsGCI (see also section B.1.1) connectivity analysis was performed and validated on multivariate time series that comprise a known network module structure. Here I briefly describe the simulation of such time series based on simulated ground truth networks.

The multivariate time series corresponding to each ground truth network and comprising their connectivity structure were realized on the basis of first order $|\mathcal{V}|$-variate autoregressive models formally given by equation B.1. Thereby, the number of available temporal samples was kept constant with $N_t = 1000$ [273]. The corresponding AR matrices $\boldsymbol{A}^1$ of equation B.1 were separately defined for each ground truth adjacency matrix as follows: if there is no connection from node v_j to node v_i, the corresponding AR parameter coefficient $\boldsymbol{A}^1_{ij}$ was set to zero (note the different order of indices as compared to the notation typically used for adjacency matrices). Otherwise, it is $\boldsymbol{A}^1_{ij} = \phi \cdot (0.99/\eta)$ where ϕ is selected uniform at random from $\{-1, 1\}$ and η is the maximum in-degree of all nodes. This scaling ensures the stationarity of the resulting multivariate process [301]. The maximum column sums of all adjacency matrices were restricted to be at most fifteen, which yields $\eta \leq 15$ (see also section 3.4 on page 71). Thus, this column sum constraint ensures similar coupling strengths for all dimensions $|\mathcal{V}|$. As a consequence, the process of generating ground truth networks was inevitably controlled by the requirements of the autoregressive parameters $\boldsymbol{A}^1_{ij}$, which directly depend on η. Since the column sum constraint needs to be satisfied for all nodes, adjustment of the other network simulation parameters to a narrow range of values was required (see also section 3.4 on page 71, Simulating networks with pre-defined module structure). The added noise terms $\boldsymbol{E}(n)$ are zero mean normally distributed random numbers with an identity covariance matrix.

B.3 Supplemental module structure recoverability and quality analysis results

The following boxplots give a detailed overview of the module structure recoverability and quality analysis results for the second study presented in chapter 6.
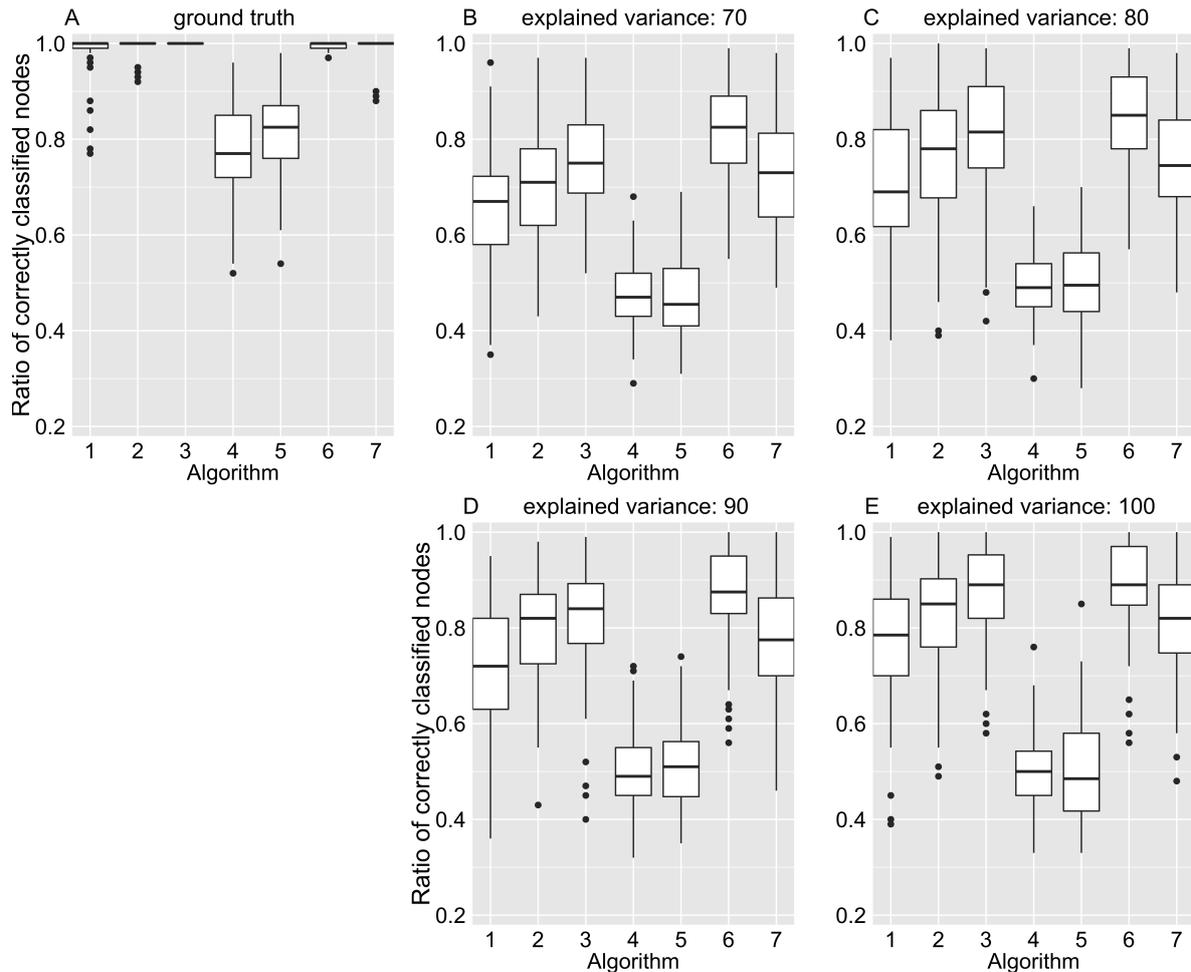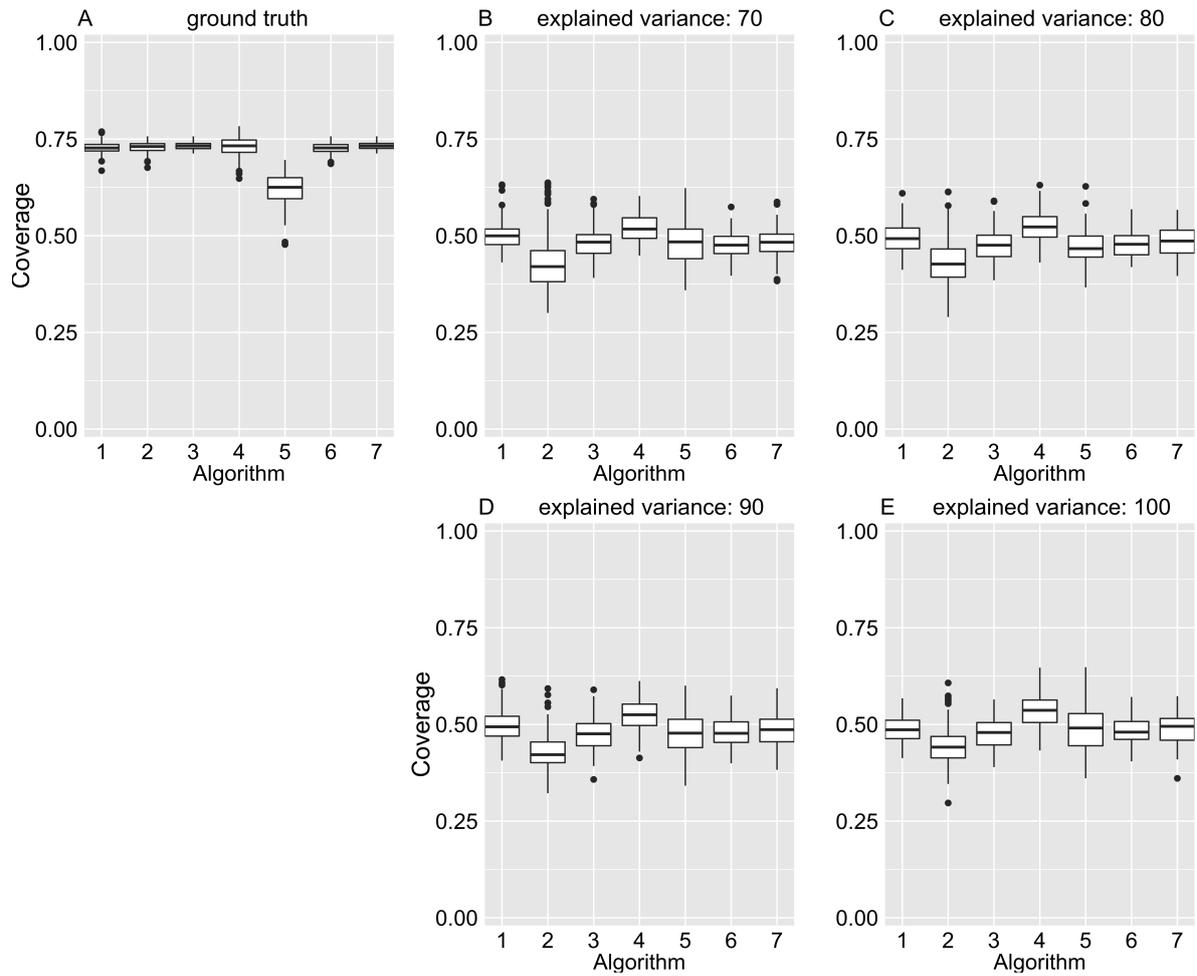

Figure B.1 – **Ratio of correctly classified nodes: Evaluation of module recoverability in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A) ground truth network, (B-D) lsGCI network with variance explanations from 70%-90%, (E) GCI network. Taken from [273].
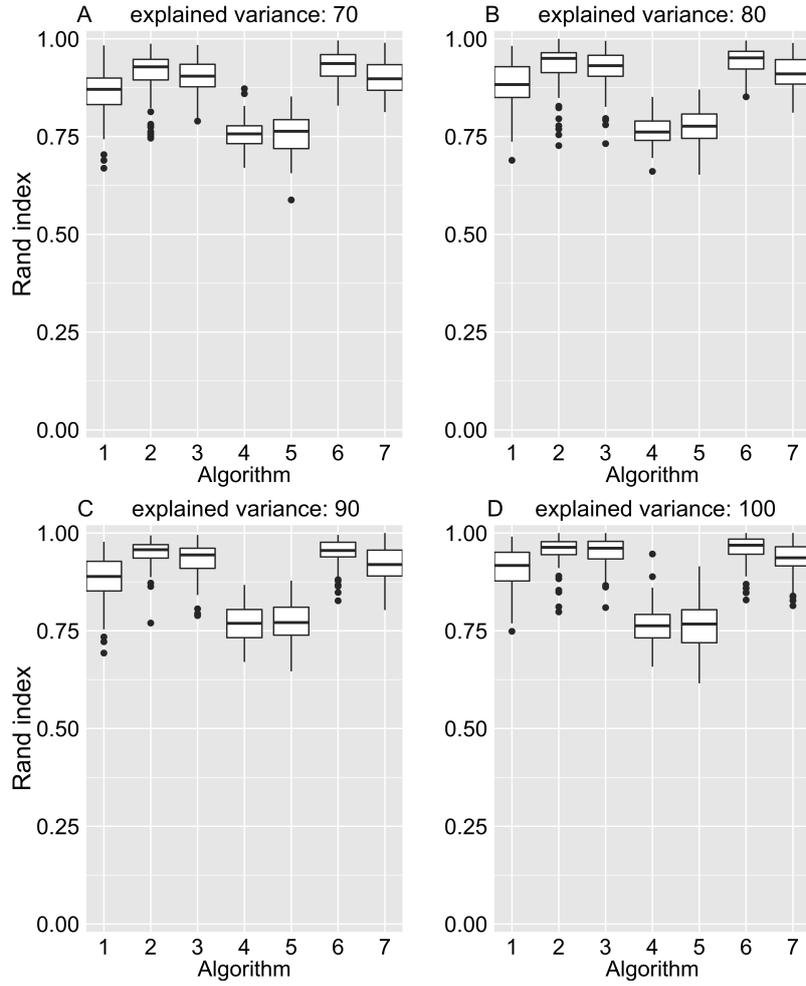
Figure B.2 – **Rand index: Evaluation of module recoverability in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A-C) lsGCI network with variance explanations from 70%-90%, (D) GCI network. Taken from [273].
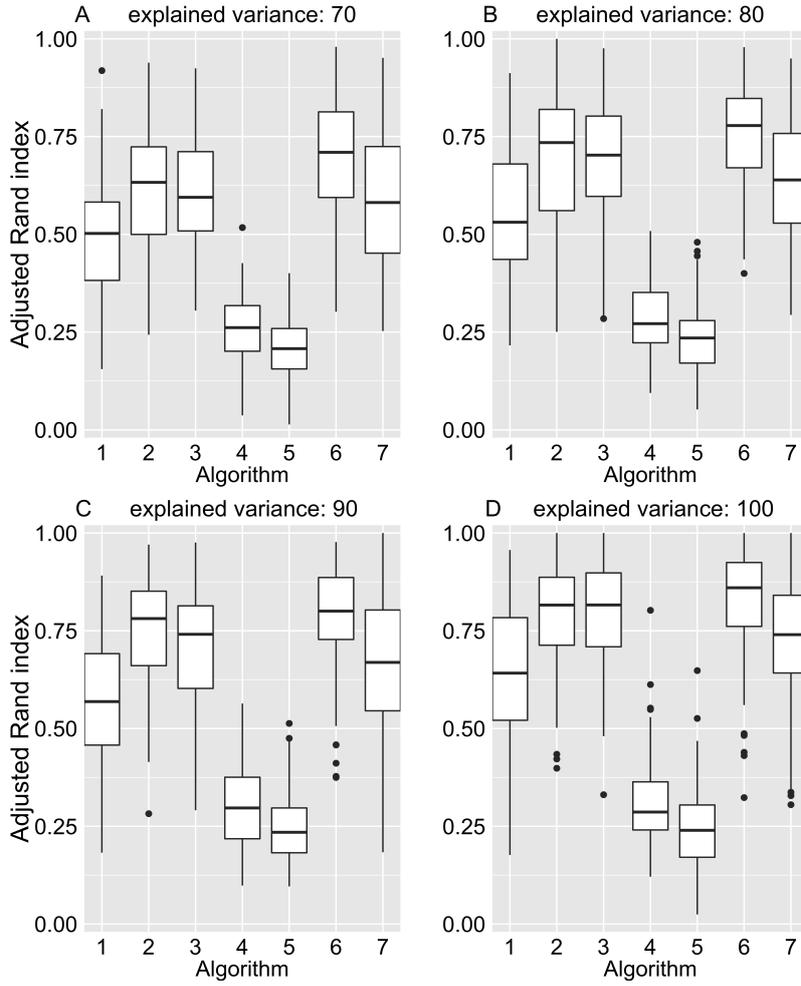
Figure B.3 – **Adjusted Rand index: Evaluation of module recoverability in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A-C) lsGCI network with variance explanations from 70%-90%, (D) GCI network. Taken from [273].
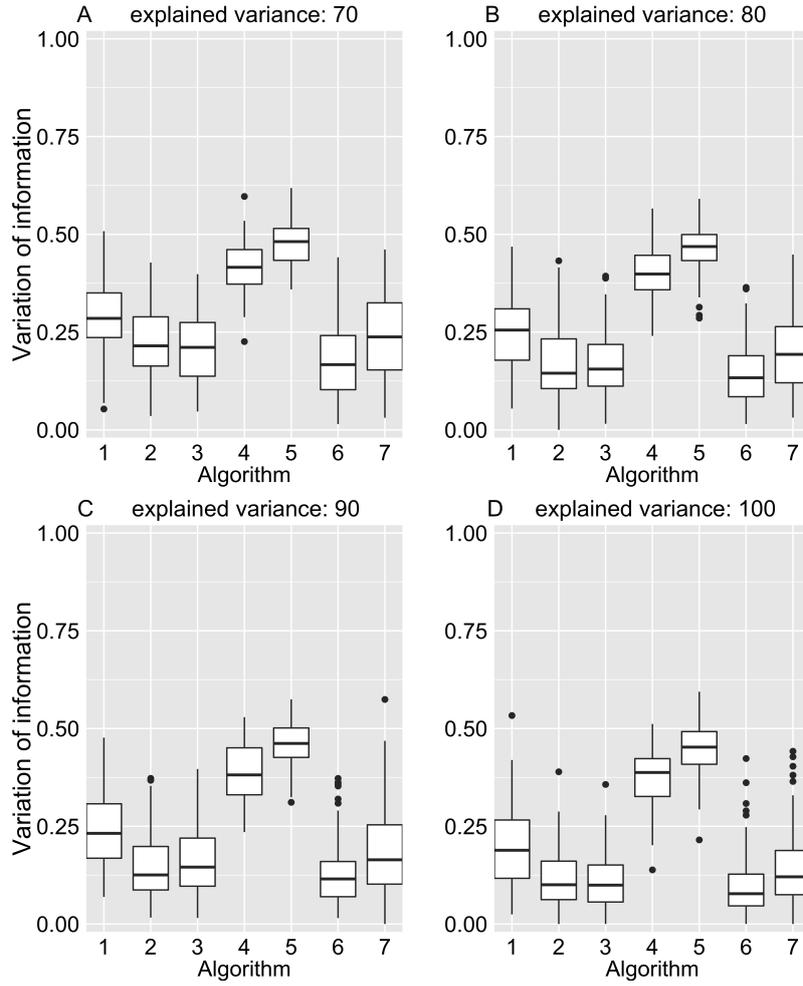
Figure B.4 – **Variation of information: Evaluation of module recoverability in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A-C) lsGCI network with variance explanations from 70%-90%, (D) GCI network. Taken from [273].
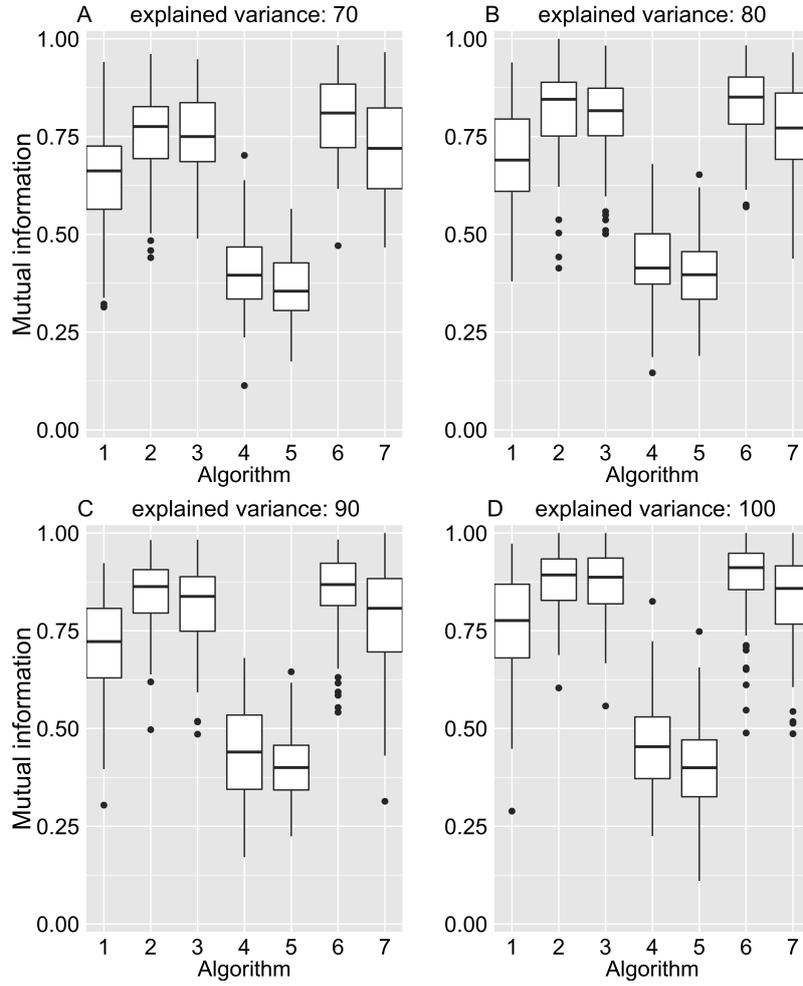
Figure B.5 – **Mutual information: Evaluation of module recoverability in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A-C) lsGCI network with variance explanations from 70%-90%, (D) GCI network. Taken from [273].
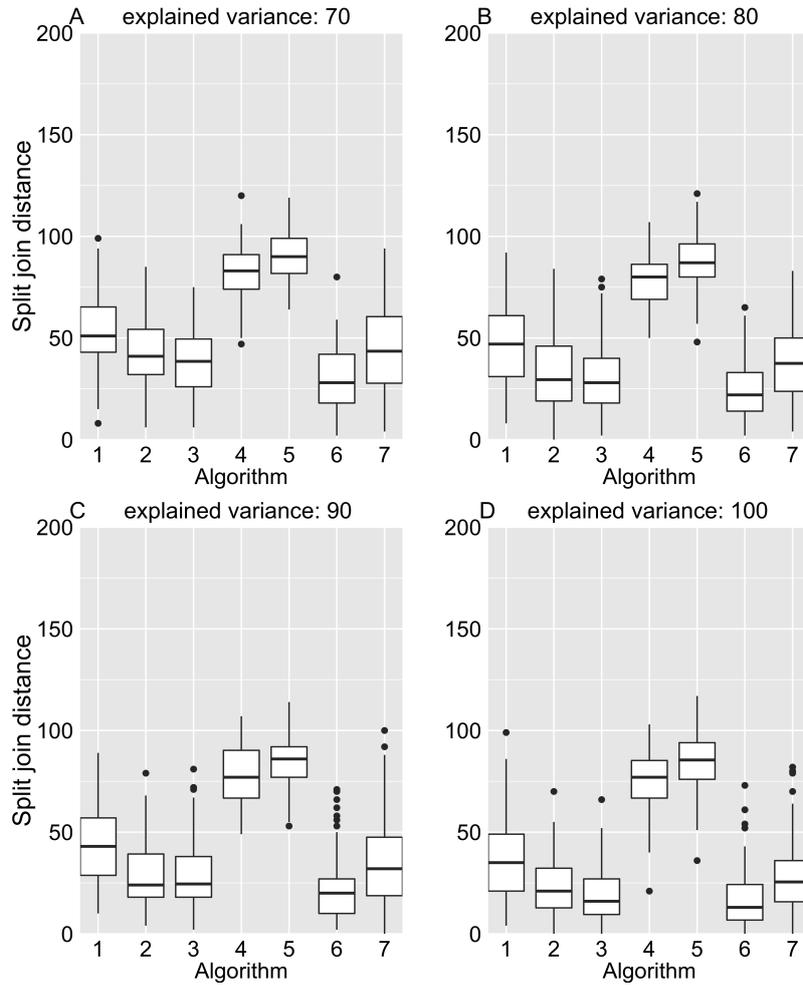
Figure B.6 – **Split-join distance: Evaluation of module recoverability in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A-C) lsGCI network with variance explanations from 70%-90%, (D) GCI network. Taken from [273].
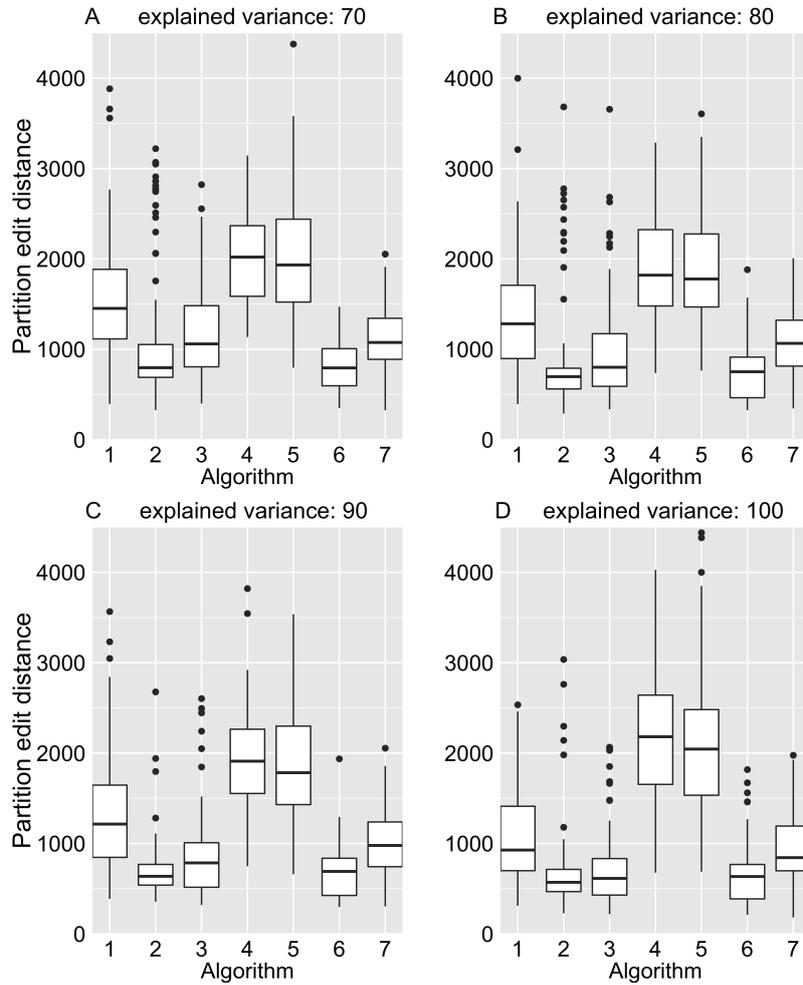
Figure B.7 – **Partition edit distance: Evaluation of module structure quality in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A-C) lsGCI network with variance explanations from 70%-90%, (D) GCI network. Taken from [273].
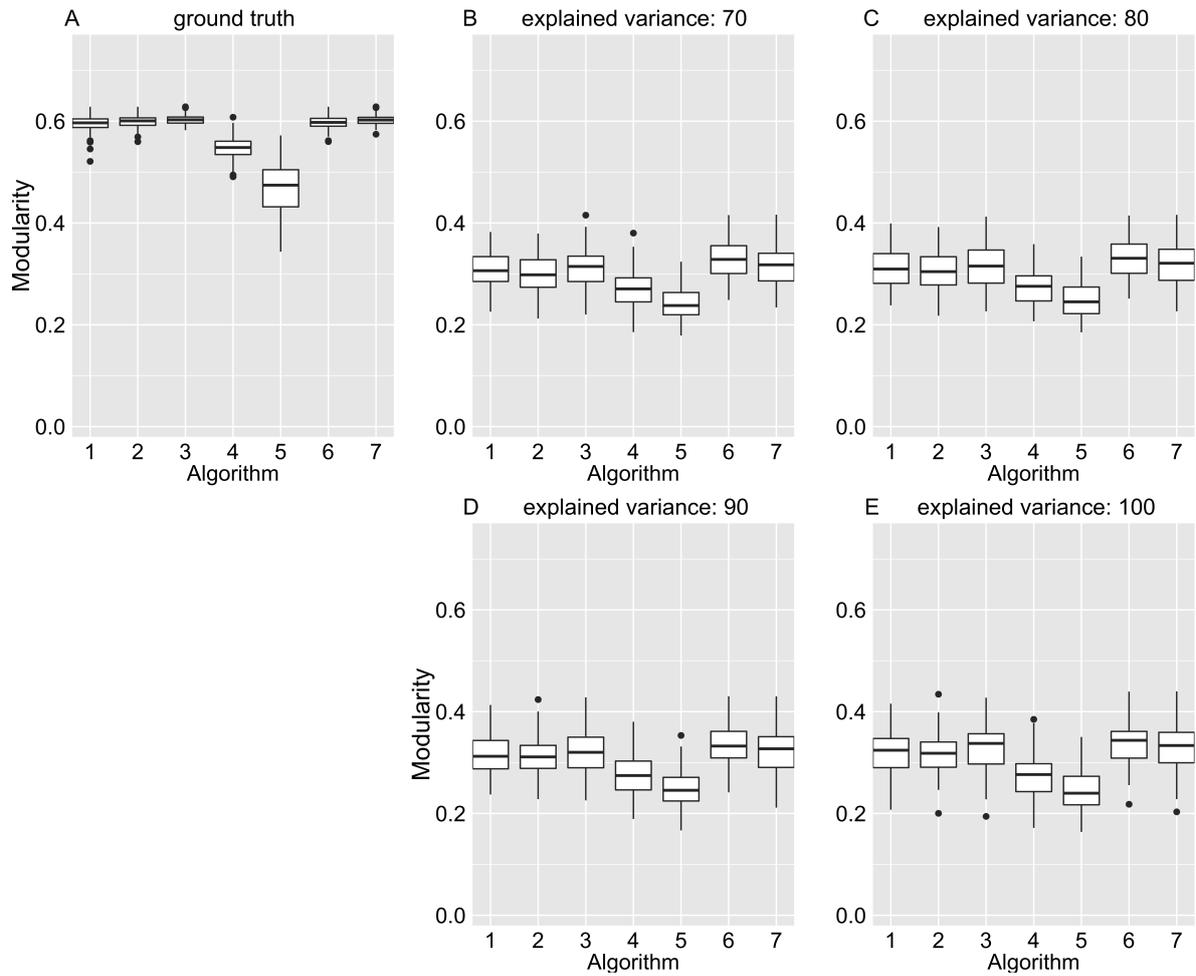
Figure B.8 – **Modularity: Evaluation of module structure quality in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A) ground truth network, (B-D) lsGCI network with variance explanations from 70%-90%, (E) GCI network. Taken from [273].
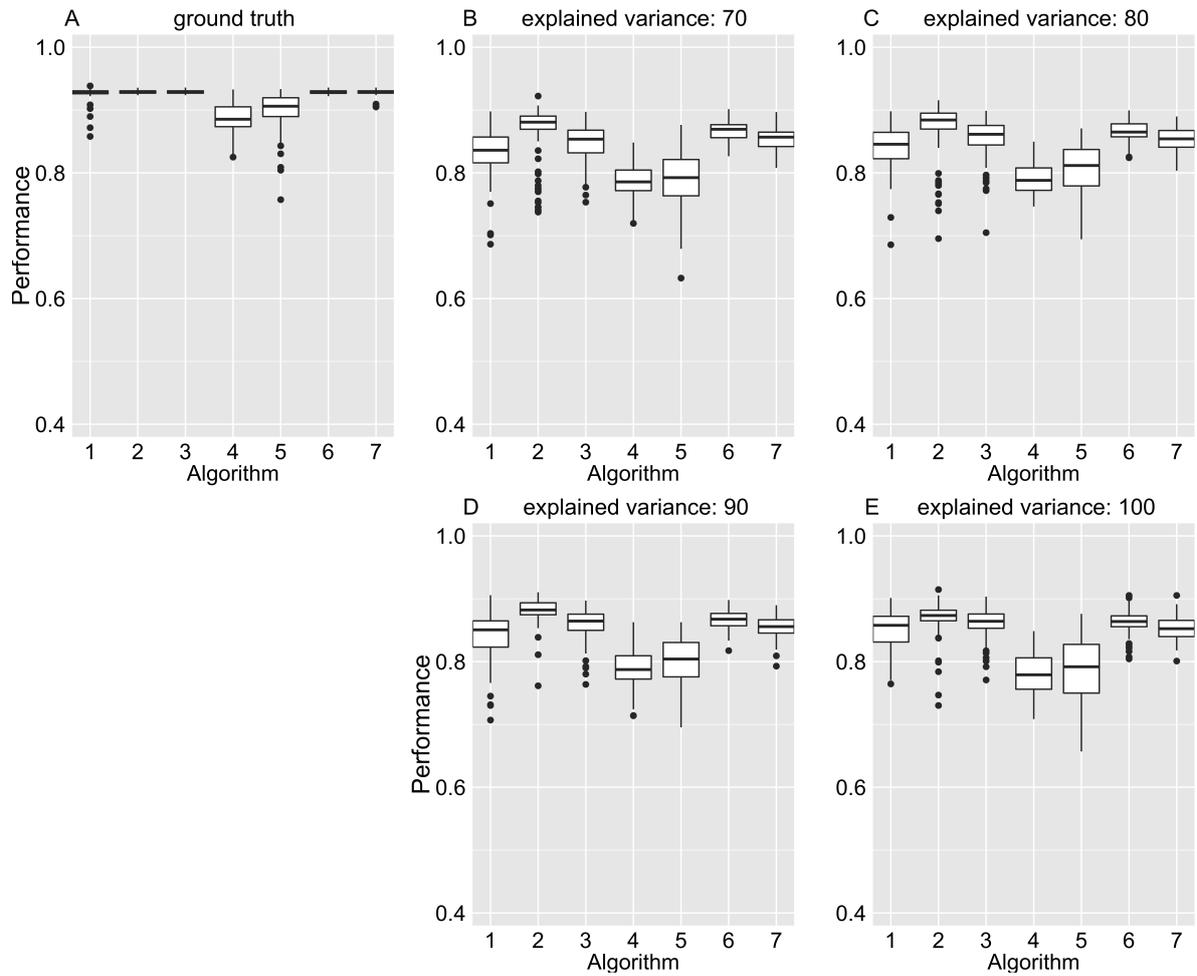
Figure B.9 – **Performance: Evaluation of module structure quality in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A) ground truth network, (B-D) lsGCI network with variance explanations from 70%-90%, (E) GCI network. Taken from [273].

Figure B.10 – **Coverage: Evaluation of module structure quality in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A) ground truth network, (B-D) lsGCI network with variance explanations from 70%-90%, (E) GCI network. Taken from [273].
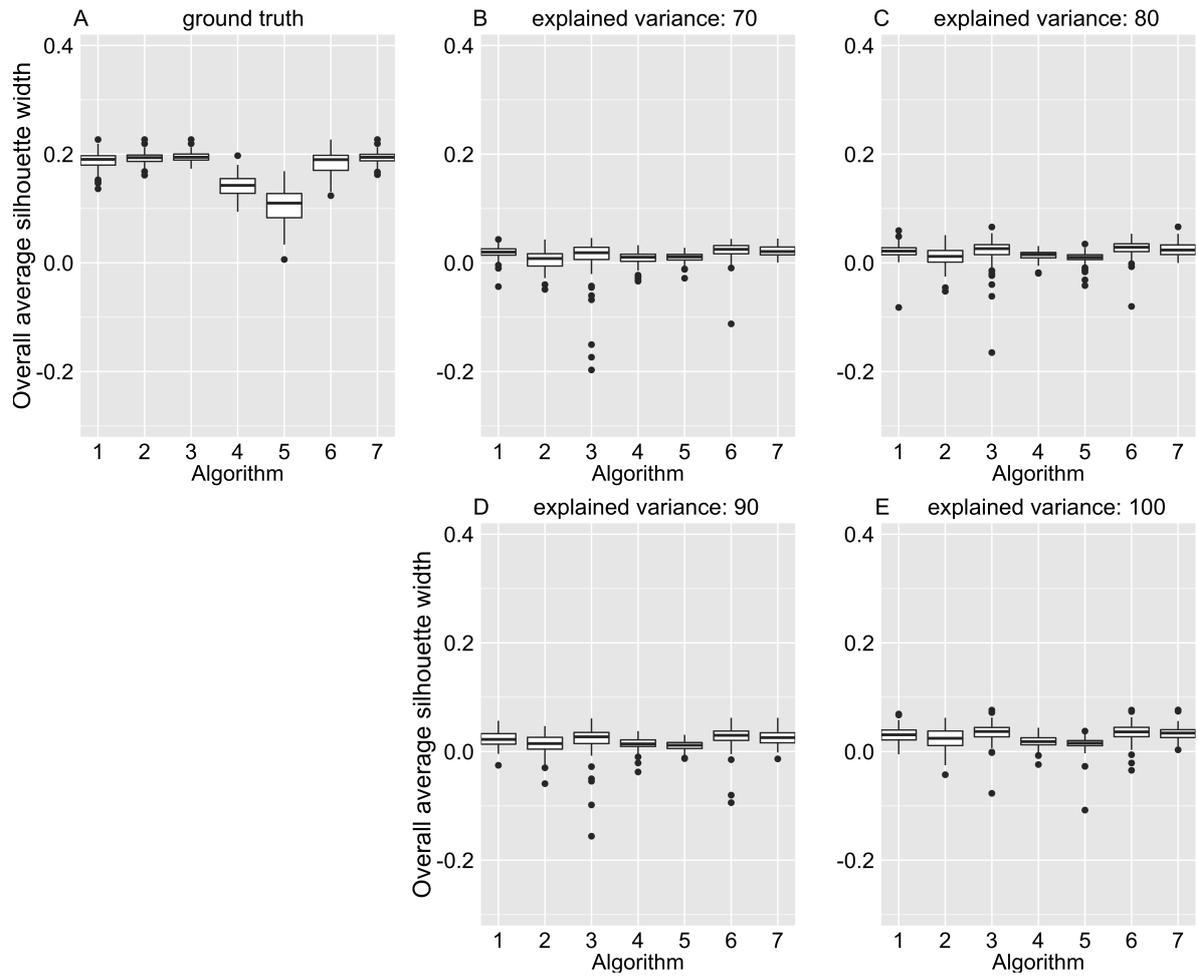
Figure B.11 – **Overall average silhouette width: Evaluation of module structure quality in ground truth and (ls)GCI networks (study II).** $N = 100$, $N_s = 1000$. The following algorithms for network module identification were used: (1) "leading eigenvector" [BD], (2) "Louvain" [BD], (3) Random walk algorithm of Pons & Latapy [BU], (4) greedy algorithm of Clauset *et al.* [BU], (5) "leading eigenvector" [BU], (6) "Potts spin glass" [BU], (7) "Louvain" [BU]. (A) ground truth network, (B-D) lsGCI network with variance explanations from 70%-90%, (E) GCI network. Taken from [273].

B.4   Time evolution of module structure

As networks change over time, so do the network modules they contain. The 'life cycle' of modules[1], the topic of this chapter, is given by formation events, growth processes, contraction processes, mergers, splitting events, continuation events and dissolution events [233]. A module *formation* event is being observed if a module k emerged at time n, $\mathcal{M}_k^n$, for which no corresponding (similar) module exists at time $n-1$ (or, if also non-subsequent time steps are considered, at any previous time step). During module *growth* or *contraction*, the module gains or loses a considerable amount of member nodes. *Mergers* of modules occur between successive time steps if at least a minimum percentage $\mathcal{K}$ of nodes of k previously separated modules $\mathcal{M}_1^{n-1}, \ldots, \mathcal{M}_k^{n-1}$ have aggregated and are fused together to form a new module $\mathcal{M}_m^n$ [281]. In a *splitting* event, the opposite situation appears, since from on time step to another, a module $\mathcal{M}_m^{n-1}$ is fragmenting into different smaller detached modules $\mathcal{M}_1^n, \ldots, \mathcal{M}_k^n$ if a sufficiently large percentage of its nodes are members of the new separated modules. Module splits signify that a considerable fraction of interactions between comember nodes is lost between two succesive time steps. A mathematical treatment of the conditions for merging and splitting events may be found in [281]. A *continuation* event is present if a module observed at one time step $n-1$ remains largely unchanged at the following time step n with respect to its node composition ($\mathcal{M}_k^{n-1} = \mathcal{M}_k^n$, $\mathcal{V}_{\mathcal{M}_k^{n-1}} \sim \mathcal{V}_{\mathcal{M}_k^n}$) so that both observations of the module can be trivially matched one-to-one. A module *dissolution* event occurs if a given module has not been observed from one time step $n-1$ to another time step n (or, in a different notion, if also non-subsequent time steps are considered, for at least d consecutive time steps [172]). If also non-subsequent time points are considered for the module tracking analysis, it is possible to classify modules as *intermittent* if they are not found during all observed time points. To discern the intermittence of module existence from a dissolution event, a maximum number of time steps in which an intermittent module is required to reoccur has to be defined.

---

[1]A more fine-grained analysis, which adopts a node-centric perspective to consider and track the dynamic interactions between nodes and the influence individual nodes have on dynamic network modules, is possible, too. For it, several node related indices and node activity functions were proposed in [281].

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die Promotionsordnung der Medizinischen Fakultät der Friedrich-Schiller-Universität bekannt ist,

ich die Dissertation selbst angefertigt habe und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben sind,

mich folgende Personen bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts unterstützt haben: Prof. Dr. Herbert Witte, Prof. Dr. Stefan Schuster, Dr. Lutz Leistritz, Dr. Britta Pester, Dr. Diana Piper, Dr. Thomas Lehmann, Prof. Dr. Thomas Weiß, Prof. Dr. Giovanni Schifitto, Dr. Nicole Schmid-Hertel, Prof. Dr. Axel Wismueller, Dr. Madalina Tivarus,

die Hilfe eines Promotionsberaters nicht in Anspruch genommen wurde und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,

dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe und

dass ich die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung nicht bei einer anderen Hochschule als Dissertation eingereicht habe.

Ort, Datum   Unterschrift des Verfassers

Index

References

[1] Karl J Friston, "Functional and effective connectivity: a review", *Brain Connectivity*, vol. 1, no. 1, pp. 13–36, 2011. (Cited on pages 2, 3, and 6.)

[2] Mikail Rubinov and Olaf Sporns, "Complex network measures of brain connectivity: Uses and interpretations", *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010. (Cited on pages 2, 3, 22, 25, 27, and 29.)

[3] Olaf Sporns, "Contributions and challenges for network models in cognitive neuroscience", *Nature Neuroscience*, vol. 17, no. 5, pp. 652–660, 2014. (Cited on pages 2, 3, 5, 21, 103, 104, and 143.)

[4] R Cameron Craddock, Rosalia L Tungaraza, and Michael P Milham, "Connectomics and new approaches for analyzing human brain functional connectivity", *GigaScience*, vol. 4, no. 1, pp. 4734, 2015. (Cited on page 2.)

[5] Gagan S Wig, Bradley L Schlaggar, and Steven E Petersen, "Concepts and principles in the analysis of brain networks", *Annals of the New York . . .* , vol. 1224, no. 1, pp. 126–146, 2011. (Cited on page 2.)

[6] Olaf Sporns, Dante R Chialvo, Marcus Kaiser, and Claus C Hilgetag, "Organization, development and function of complex brain networks", *Trends in Cognitive Sciences*, vol. 8, no. 9, pp. 418–425, 2004. (Cited on page 2.)

[7] Karl J Friston, C D Frith, P F Liddle, and R S J Frackowiak, "Functional Connectivity: The Principal-Component Analysis of Large (PET) Data Sets", *Journal of Cerebral Blood Flow & Metabolism*, vol. 13, no. 1, pp. 5–14, 1993. (Cited on pages 3 and 6.)

[8] A Roebroeck, E Formisano, and R Goebel, "Mapping directed influence over the brain using Granger causality and fMRI", *NeuroImage*, vol. 25, no. 1, pp. 230–242, 2005. (Cited on pages 3 and 146.)

[9] Lutz Leistritz, Karin Schiecke, Laura Astolfi, and Herbert Witte, "Time-Variant Modeling of Brain Processes", *Proceedings of the IEEE*, vol. 104, no. 2, pp. 262–281, 2016. (Cited on pages 3, 5, and 6.)

[10] Arnaud Messé, Marc-Thorsten Hütt, Peter König, and Claus C Hilgetag, "A closer look at the apparent correlation of structural and functional connectivity in excitable neural networks", *Scientific Reports*, vol. 5, pp. 7870, 2015. (Cited on page 3.)

[11] Russell A Poldrack and Martha J Farah, "Progress and challenges in probing the human brain", *Nature*, vol. 526, no. 7573, pp. 371–379, 2015. (Cited on pages 3, 5, and 105.)

[12] Andrew Zalesky, Alex Fornito, and Edward T Bullmore, "Network-based statistic: Identifying differences in brain networks", *NeuroImage*, vol. 53, no. 4, pp. 1197–1207, 2010. (Cited on pages 3 and 16.)

[13] Edward T Bullmore and Danielle S Bassett, "Brain graphs: graphical models of the human brain connectome", *dx.doi.org*, vol. 7, no. 1, pp. 113–140, 2011. (Cited on page 4.)

[14] Olaf Sporns, *Networks of the Brain*, The MIT Press, 2010. (Cited on pages 4, 22, 28, and 30.)

[15] Danielle S Bassett, Nicholas F Wymbs, Mason A Porter, Peter J Mucha, Jean M Carlson, and Scott T Grafton, "Dynamic reconfiguration of human brain networks during learning", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 18, pp. 7641–7646, 2011. (Cited on pages 4, 16, 38, 44, 103, and 129.)

[16] Walter J Freeman and Rodrigo Quian Quiroga, *Imaging Brain Function With EEG*, Springer, 2013. (Cited on page 5.)

[17] Andrew S Blum and Seward B Rutkove, Eds., *The Clinical Neurophysiology Primer*, Humana Press, Totowa, NJ, 2007. (Cited on page 5.)

[18] Robert Oostenveld and Peter Praamstra, "The five percent electrode system for high-resolution EEG and ERP measurements", *Clinical Neurophysiology*, vol. 112, no. 4, pp. 713–719, 2001. (Cited on page 5.)

[19] S Baillet, J C Mosher, and R M Leahy, "Electromagnetic brain mapping", *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, 2001. (Cited on page 5.)

[20] L Gavit, S Baillet, J F Mangin, J Pescatore, and L Garnero, "A multiresolution framework to MEG/EEG source imaging.", *IEEE transactions on bio-medical engineering*, vol. 48, no. 10, pp. 1080–1087, 2001. (Cited on page 5.)

[21] S Ogawa, T M Lee, A R Kay, and D W Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 24, pp. 9868–9872, 1990. (Cited on page 5.)

[22] Juri D Kropotov, "Functional Magnetic Resonance Imaging", in *Functional Neuromarkers for Psychiatry*, pp. 17–25. Elsevier, 2016. (Cited on page 5.)

[23] Richard B Buxton, "The physics of functional magnetic resonance imaging (fMRI)", *Reports on Progress in Physics*, vol. 76, no. 9, pp. 096601, 2013. (Cited on page 5.)

[24] Arshi Khalid, Byung Sun Kim, Moo K Chung, Jong Chul Ye, and Daejong Jeon, "Tracing the evolution of multi-scale functional networks in a mouse model of depression using persistent brain network homology", *NeuroImage*, vol. 101, pp. 351–363, 2014. (Cited on pages 5, 17, and 19.)

[25] Karl J Friston, Rosalyn Moran, and Anil K Seth, "Analysing connectivity with Granger causality and dynamic causal modelling.", *Current Opinion in Neurobiology*, vol. 23, no. 2, pp. 172–178, 2013. (Cited on page 6.)

[26] Huifang E Wang, Christian Bernard, Pascale P Quilichini, Karl J Friston, Viktor K Jirsa, and Christophe Bernard, "A systematic framework for functional connectivity measures", *Frontiers in Neuroscience*, vol. 8, no. 57, pp. 31916, 2014. (Cited on page 6.)

[27] Duncan J Watts and Steven H Strogatz, "Collective dynamics of 'small-world' networks", *Nature*, vol. 393, no. 6684, pp. 440–442, 1998. (Cited on pages 10, 24, 28, and 48.)

[28] Albert-László Barabási and Réka Albert, "Emergence of scaling in random networks", *Science*, vol. 286, no. 5439, pp. 509–512, 1999. (Cited on pages 10 and 47.)

[29] W K Chen, *Graph Theory and Its Engineering Applications*, World Scientific, 1997. (Cited on pages 11 and 13.)

[30] Jørgen Bang-Jensen and Gregory Gutin, *Digraphs: Theory, Algorithms and Applications*, Springer Verlag, 1st edition edition, 2007. (Cited on page 11.)

[31] Mark E. J. Newman, *Networks*, An Introduction. Oxford University Press, 2010. (Cited on pages 11, 22, 24, 27, 28, 29, 30, 46, 47, and 48.)

[32] Frank Harary, *Graph Theory*, Addison-Wesley, 1969. (Cited on page 11.)

[33] Frank Harary, Robert Z Norman, and Dorwin Cartwright, *Structural Models*, An Introduction to the Theory of Directed Graphs. John Wiley & Sons, Inc., 1965. (Cited on page 11.)

[34] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein, *Introduction to Algorithms*, The MIT Press, 3rd edition edition, 2009. (Cited on pages 13 and 44.)

[35] Jon Kleinberg and Éva Tardos, *Algorithm Design*, Addison-Wesley, 1st edition edition, 2005. (Cited on page 13.)

[36] Steve Horvath, *Weighted Network Analysis*, Applications in Genomics and Systems Biology. Springer, 2011. (Cited on pages 16, 20, and 22.)

[37] Bernadette C M van Wijk, Cornelis J Stam, and Andreas Daffertshofer, "Comparing brain networks of different size and connectivity density using graph theory", *PLoS ONE*, vol. 5, no. 10, pp. e13701, 2010. (Cited on pages 16, 17, and 107.)

[38] Andrew C Thomas and Joseph K Blitzstein, "Valued Ties Tell Fewer Lies: Why Not To Dichotomize Network Edges With Thresholds", *arXiv*, 2011. (Cited on page 16.)

[39] Christopher Aicher, Abigail Z Jacobs, and Aaron Clauset, "Learning latent block structure in weighted networks", *Journal of Complex Networks*, vol. 3, no. 2, pp. 221–248, 2015. (Cited on pages 17, 20, 25, and 33.)

[40] Luca Ferrarini, Ilya M Veer, Evelinda Baerends, Marie-José van Tol, Remco J Renken, Nic J A van der Wee, Dirk J Veltman, André Aleman, Frans G Zitman, Brenda W J H Penninx, Mark A van Buchem, Johan H C Reiber, Serge A R B Rombouts, and Julien Milles, "Hierarchical functional modularity in the resting-state human brain", *Human Brain Mapping*, vol. 30, no. 7, pp. 2220–2231, 2009. (Cited on page 17.)

[41] Lutz Leistritz, Thomas Weiss, Karl-Jürgen Bär, Fabrizio De Vico Fallani, Fabio Babiloni, Herbert Witte, and Thomas Lehmann, "Network redundancy analysis of effective brain networks; a comparison of healthy controls and patients with major depression", *PLoS ONE*, vol. 8, no. 4, pp. e60956, 2013. (Cited on pages 17 and 18.)

[42] Hyekyoung Lee, Hyejin Kang, M K Chung, Bung-Nyun Kim, and Dong Soo Lee, "Persistent Brain Network Homology From the Perspective of Dendrogram", *Medical Imaging*, vol. 31, no. 12, pp. 2267–2277, 2012. (Cited on pages 17 and 19.)

[43] Regina Nuzzo, "Statistical errors", *Nature*, vol. 506, no. 7487, pp. 150–152, 2014. (Cited on page 17.)

[44] John P A Ioannidis, "Why Most Published Research Findings Are False", *PLOS Med*, vol. 2, no. 8, pp. e124, 2005. (Cited on page 17.)

[45] Monya Baker, "Statisticians issue warning over misuse of P values", *Nature*, vol. 531, no. 7593, pp. 151–151, 2016. (Cited on page 17.)

[46] Geoff Cumming, "The new statistics: Why and how", *Psychological Science*, vol. 25, no. 1, pp. 7–29, 2014. (Cited on page 17.)

[47] Steven N Goodman, "Of P-Values and Bayes: A Modest Proposal", *Epidemiology*, vol. 12, no. 3, pp. 295–297, 2001. (Cited on page 17.)

[48] Valen E Johnson, "Revised standards for statistical evidence", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 48, pp. 19313–19317, 2013. (Cited on page 17.)

[49] David Colquhoun, "An investigation of the false discovery rate and the misinterpretation of p-values", *Royal Society Open Science*, vol. 1, no. 3, pp. 140216–140216, 2014. (Cited on page 17.)

[50] Norm Matloff, *From Algorithms to Z-Scores: Probalistic and Statistical Modeling in Computer Science*, open-source textbook, University of California, Davis. (Cited on page 17.)

[51] Sophie Achard, Raymond Salvador, Brandon Whitcher, John Suckling, and Edward T Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs", *The Journal of Neuroscience*, vol. 26, no. 1, pp. 63–72, 2006. (Cited on page 17.)

[52] Mark S Granovetter, "The strength of weak ties", *American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973. (Cited on pages 17 and 129.)

[53] Mark S Granovetter, "The strength of weak ties: A network theory revisited", *Sociological theory*, vol. 1, pp. 201–233, 1983. (Cited on pages 17 and 129.)

[54] Sinan Aral, "The Future of Weak Ties", *American Journal of Sociology*, vol. 121, no. 6, pp. 1931–1939, 2016. (Cited on pages 17 and 129.)

[55] Ronald S Burt, "Bridge decay", *Social Networks*, vol. 24, no. 4, pp. 333–363, 2002. (Cited on pages 17, 24, and 129.)

[56] Marco Catani and Marsel Mesulam, "What is a disconnection syndrome?", *Cortex*, vol. 44, no. 8, pp. 911–913, 2008. (Cited on page 17.)

[57] Bradley Efron and Robert J Tibshirani, *An Introduction to the Bootstrap*, Monographs on statistics and applied probability. Chapman & Hall, 1993. (Cited on pages 18 and 138.)

[58] João R Sato, Daniel Y Takahashi, Silvia M Arcuri, Koichi Sameshima, Pedro A Morettin, and Luiz A Baccalá, "Frequency domain connectivity identification: An application of partial directed coherence in fMRI", *Human Brain Mapping*, vol. 30, no. 2, pp. 452–461, 2009. (Cited on pages 18, 146, and 148.)

[59] Sture Holm, "A simple sequentially rejective multiple test procedure", *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1979. (Cited on pages 18, 61, and 64.)

[60] Yoav Benjamini and Yosef Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing", *Journal of the Royal Statistical Society Series B-Methodological*, vol. 57, no. 1, pp. 289–300, 1995. (Cited on pages 19 and 61.)

[61] Christoph Schmidt, Britta Pester, Mahesh Nagarajan, Herbert Witte, Lutz Leistritz, and Axel Wismueller, "Impact of multivariate Granger Causality analyses with embedded dimension reduction on network modules", *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2797–2800, 2014. (Cited on pages 19, 104, 105, 106, 109, 120, 145, 148, and 150.)

[62] Bin Zhang and Steve Horvath, "A General Framework for Weighted Gene Co-Expression Network Analysis", *Statistical applications in genetics and . . .*, vol. 4, no. 1, 2005. (Cited on pages 19, 20, and 47.)

[63] Ulrik Brandes, "On variants of shortest-path betweenness centrality and their generic computation", *Social Networks*, vol. 30, no. 2, pp. 136–145, 2008. (Cited on pages 20, 21, 23, and 24.)

[64] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani, "Modeling the evolution of weighted networks", *Physical Review E*, vol. 70, no. 6, pp. 066149, 2004. (Cited on pages 20, 24, and 30.)

[65] Tore Opsahl and Pietro Panzarasa, "Clustering in weighted networks", *Social Networks*, vol. 31, no. 2, pp. 155–163, 2009. (Cited on pages 20 and 28.)

[66] Tore Opsahl, Filip Agneessens, and John Skvoretz, "Node centrality in weighted networks: Generalizing degree and shortest paths", *Social Networks*, vol. 32, no. 3, pp. 245–251, 2010. (Cited on pages 21, 23, and 24.)

[67] R N Mantegna, "Hierarchical structure in financial markets", *The European Physical Journal B*, vol. 11, no. 1, pp. 193–197, 1999. (Cited on pages 21 and 98.)

[68] Mark E. J. Newman, "The structure and function of complex networks", *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003. (Cited on pages 22 and 28.)

[69] S Boccaletti, Vito 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. (Cited on pages 22, 23, 27, and 28.)

[70] Santo Fortunato, "Community detection in graphs", *Physics Reports*, vol. 486, no. 3-5, pp. 75–174, 2010. (Cited on pages 22, 27, 31, 42, 45, 46, 47, 103, and 130.)

[71] Edward T Bullmore and Olaf Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems", *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009. (Cited on page 22.)

[72] Stanley Wasserman and Faust, Katherine, *Social Network Analysis*, Methods and Applications. Cambridge University Press, 1994. (Cited on pages 22, 28, and 30.)

[73] Ernesto Estrada, *The Structure of Complex Networks*, Theory and Applications. Oxford University Press, 2011. (Cited on pages 22 and 27.)

[74] Alex Bavelas, "A Mathematical Model for Group Structures", *Applied Anthropology*, vol. 7, no. 3, pp. 16–30, 1948. (Cited on pages 23 and 24.)

[75] Per Hage and Frank Harary, "Eccentricity and centrality in networks", *Social Networks*, vol. 17, no. 1, pp. 57–63, 1995. (Cited on pages 23, 24, and 29.)

[76] Linton C Freeman, "Centrality in social networks conceptual clarification", *Social Networks*, vol. 1, no. 3, pp. 215–239, 1978. (Cited on pages 23, 24, and 30.)

[77] Diego Garlaschelli and Maria I Loffredo, "Patterns of link reciprocity in directed networks.", *Physical Review Letters*, vol. 93, no. 26 Pt 1, pp. 268701, 2004. (Cited on pages 24 and 29.)

[78] Ulrik Brandes, "A faster algorithm for betweenness centrality", *Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163–177, 2001. (Cited on page 24.)

[79] Ulrik Brandes, Stephen P Borgatti, and Linton C Freeman, "Maintaining the duality of closeness and betweenness centrality", *Social Networks*, vol. 44, pp. 153–159, 2016. (Cited on page 24.)

[80] M R da Silva, Hongwu Ma, and An-Ping Zeng, "Centrality, Network Capacity, and Modularity as Parameters to Analyze the Core-Periphery Structure in Metabolic Networks", *Proceedings of the IEEE*, vol. 96, no. 8, pp. 1411–1420, 2008. (Cited on pages 24, 31, and 33.)

[81] Linton C Freeman, "A Set of Measures of Centrality Based on Betweenness", *Sociometry*, vol. 40, no. 1, pp. 35, 1977. (Cited on page 24.)

[82] Mark E. J. Newman, "A measure of betweenness centrality based on random walks", *Social Networks*, vol. 27, no. 1, pp. 39–54, 2005. (Cited on page 24.)

[83] O Green, R McColl, and D A Bader, *A Fast Algorithm for Streaming Betweenness Centrality*, IEEE, 2012. (Cited on page 24.)

[84] Miray Kas, Matthew Wachs, Kathleen M Carley, and L Richard Carley, *Incremental algorithm for updating betweenness centrality in dynamically growing networks*, ACM, New York, New York, USA, 2013. (Cited on page 24.)

[85] Meghana Nasre, Matteo Pontecorvi, and Vijaya Ramachandran, "Betweenness Centrality – Incremental and Faster", in *Mathematical Foundations of Computer Science 2014*, pp. 577–588. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. (Cited on page 24.)

[86] N Kourtellis, G D F Morales, and F Bonchi, "Scalable Online Betweenness Centrality in Evolving Graphs", *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 9, pp. 2494–2506, 2015. (Cited on page 24.)

[87] Ulrik Brandes and CHRISTIAN PICH, "Centrality estimation in large networks", *International Journal of Bifurcation and Chaos*, vol. 17, no. 07, pp. 2303–2318, 2011. (Cited on page 24.)

[88] Stephen P Borgatti, "Identifying sets of key players in a social network", *Computational and Mathematical Organization Theory*, vol. 12, no. 1, pp. 21–34, 2006. (Cited on pages 24 and 31.)

[89] Roger Guimerà and Luís A Nunes Amaral, "Functional cartography of complex metabolic networks", *Nature*, vol. 433, no. 7028, pp. 895–900, 2005. (Cited on page 24.)

[90] Giorgio Fagiolo, "Clustering in complex directed networks", *Physical Review E*, vol. 76, no. 2, pp. 026107, 2007. (Cited on pages 24 and 28.)

[91] Alexei Vázquez, Romualdo Pastor-Satorras, and Alessandro Vespignani, "Large-scale topological and dynamical properties of the Internet", *Physical Review E*, vol. 65, no. 6, pp. 066130, 2002. (Cited on page 24.)

[92] Erzsébet Ravasz and Albert-László Barabási, "Hierarchical organization in complex networks", *Physical Review E*, vol. 67, no. 2, pp. 026112, 2003. (Cited on pages 24 and 33.)

[93] Jukka-Pekka Onnela, Jari Saramäki, János Kertész, and Kimmo Kaski, "Intensity and coherence of motifs in weighted complex networks", *Physical Review E*, vol. 71, no. 6, pp. 065103, 2005. (Cited on pages 24 and 27.)

[94] A Barrat, Marc Barthélemy, R Pastor-Satorras, and A Vespignani, "The architecture of complex weighted networks", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 11, pp. 3747–3752, 2004. (Cited on pages 24 and 100.)

[95] Sergey Brin and Lawrence Page, "Reprint of: The anatomy of a large-scale hypertextual web search engine", *Computer Networks*, vol. 56, no. 18, pp. 3825–3833, 2012. (Cited on page 25.)

[96] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd, "The PageRank citation ranking: Bringing order to the web", Tech. Rep., 1999. (Cited on page 25.)

[97] Christoph Schmidt, Thomas Weiss, Christian Komusiewicz, Herbert Witte, and Lutz Leistritz, "An Analytical Approach to Network Motif Detection in Samples of Networks with Pairwise Different Vertex Labels", *Computational and Mathematical Methods in Medicine*, vol. 2012, no. 395, pp. 1–12, 2012. (Cited on pages 25, 60, 84, 85, 89, 93, and 145.)

[98] Robert J Prill, Pablo A Iglesias, and Andre Levchenko, "Dynamic properties of network motifs contribute to biological network organization", *PLoS Biology*, vol. 3, no. 11, pp. e343, 2005. (Cited on pages 25 and 26.)

[99] Ron Milo, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, Shai Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, and Uri Alon, "Superfamilies of evolved and designed networks", *Science*, vol. 303, no. 5663, pp. 1538–1542, 2004. (Cited on pages 25 and 26.)

[100] David Aparício, Pedro Ribeiro, and Fernando Silva, "Network comparison using directed graphlets", 2015. (Cited on pages 25 and 27.)

[101] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon, "Network motifs: simple building blocks of complex networks", *Science*, vol. 298, no. 5594, pp. 824–827, 2002. (Cited on pages 25, 26, 49, and 59.)

[102] Nadav Kashtan, Shalev Itzkovitz, Ron Milo, and Uri Alon, "Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs", *Bioinformatics*, vol. 20, no. 11, pp. 1746–1758, 2004. (Cited on pages 25 and 26.)

[103] Sebastian Wernicke, "A faster algorithm for detecting network motifs", in *Proceedings of the 5th Workshop on Algorithms in Bioinformatics (WABI'05)*. 2005, pp. 165–177, Springer. (Cited on page 25.)

[104] Giovanni Ciriello and Concettina Guerra, "A review on models and algorithms for motif discovery in protein-protein interaction networks", *Briefings in Functional Genomics and Proteomics*, vol. 7, no. 2, pp. 147–156, 2008. (Cited on pages 25 and 26.)

[105] Falk Schreiber and Henning Schwöbbermeyer, "Frequency concepts and pattern detection for the analysis of motifs in networks", in *Transactions on Computational Systems Biology III*, Corrado Priami, Emanuela Merelli, Pablo Gonzalez, and Andrea Omicini, Eds., pp. 89–104. Springer Berlin Heidelberg, 2005. (Cited on page 25.)

[106] P Foggia, C Sansone, and M Vento, "A performance comparison of five algorithms for graph isomorphism", in *Proceedings of the 3rd IAPR TC-15 Workshop on Graph-based Representations in Pattern Recognition*, 2001, pp. 188–199. (Cited on page 26.)

[107] Olaf Sporns and Rolf Kötter, "Motifs in brain networks", *PLoS Biology*, vol. 2, no. 11, pp. e369, 2004. (Cited on page 26.)

[108] Etay Ziv, Robin Koytcheff, Manuel Middendorf, and Chris Wiggins, "Systematic identification of statistically significant network measures", *Physical Review E*, vol. 71, no. 1, pp. 016110, 2005. (Cited on page 26.)

[109] Shai S Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon, "Network motifs in the transcriptional regulation network of Escherichia coli", *Nature Genetics*, vol. 31, pp. 64–68, 2002. (Cited on pages 26 and 59.)

[110] Tong Ihn Lee, Nicola J Rinaldi, Francois Robert, Duncan T Odom, Ziv Bar-Joseph, Georg K Gerber, Nancy M Hannett, Christopher T Harbison, Craig M Thompson, Itamar Simon, Julia Zeitlinger, Ezra G Jennings, Heather L Murray, D Benjamin Gordon, Bing Ren, John J Wyrick, Jean-Bosco Tagne, Thomas L Volkert, Ernest Fraenkel, David K Gifford, and Richard A Young, "Transcriptional regulatory networks in Saccharomyces cerevisiae", *Science*, vol. 298, no. 5594, pp. 799–804, 2002. (Cited on page 26.)

[111] S Mangan and Uri Alon, "Structure and function of the feed-forward loop network motif.", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 21, pp. 11980–11985, 2003. (Cited on page 26.)

[112] Sebastian Wernicke, *Combinatorial Algorithms to Cope with the Complexity of Biological Networks*, PhD thesis, Verl. Dr. Hut, 2006. (Cited on page 26.)

[113] Arun S Konagurthu and Arthur M Lesk, "On the origin of distribution patterns of motifs in biological networks", *BMC Systems Biology*, vol. 2, no. 1, pp. 73, 2008. (Cited on pages 26 and 95.)

[114] Nadav Kashtan, Shalev Itzkovitz, Ron Milo, and Uri Alon, "Topological generalizations of network motifs", *Physical Review E*, vol. 70, no. 3, pp. 031909, 2004. (Cited on page 26.)

[115] Nataša Pržulj, D G Corneil, and I Jurisica, "Modeling interactome: scale-free or geometric?", *Bioinformatics*, vol. 20, no. 18, pp. 3508–3515, 2004. (Cited on pages 27 and 47.)

[116] Nataša Pržulj, "Biological network comparison using graphlet degree distribution", *Bioinformatics*, vol. 23, no. 2, pp. e177–e183, 2007. (Cited on page 27.)

[117] Mark E. J. Newman and Tiago P Peixoto, "Generalized Communities in Networks", *Physical Review Letters*, vol. 115, no. 8, pp. 088701, 2015. (Cited on page 27.)

[118] Mason A Porter, Jukka-Pekka Onnela, and Peter J Mucha, "Communities in networks", *Notices of the American Mathematical Society*, vol. 56, no. 9, pp. 1082–1097, 2009. (Cited on pages 27, 31, 35, and 44.)

[119] Filippo Radicchi, Claudio Castellano, Federico Cecconi, Vittorio Loreto, and Domenico Parisi, "Defining and identifying communities in networks", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 9, pp. 2658–2663, 2004. (Cited on page 27.)

[120] Anil K Jain, "Data clustering: 50 years beyond K-means", *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010. (Cited on page 27.)

[121] Danielle Smith Bassett and Edward T Bullmore, "Small-world brain networks", *The Neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006. (Cited on page 27.)

[122] Peter Csermely, András London, Ling-Yun Wu, and Brian Uzzi, "Structure and dynamics of core/periphery networks", *Journal of Complex Networks*, vol. 1, no. 2, pp. 93–123, 2013. (Cited on pages 27, 30, 33, and 48.)

[123] Vito Latora and Massimo Marchiori, "Efficient behavior of small-world networks", *Physical Review Letters*, vol. 87, no. 19, pp. 198701–6, 2001. (Cited on page 29.)

[124] Nikola T Markov, Mária Ercsey-Ravasz, David C Van Essen, Kenneth Knoblauch, Zoltán Toroczkai, and Henry Kennedy, "Cortical High-Density Counterstream Architectures", *Science*, vol. 342, no. 6158, pp. 1238406–1238406, 2013. (Cited on pages 29 and 48.)

[125] Vito Latora and Massimo Marchiori, "Economic small-world behavior in weighted networks", *The European Physical Journal B*, vol. 32, no. 2, pp. 249–263, 2003. (Cited on page 29.)

[126] Fabrizio De Vico Fallani, Vito Latora, Laura Astolfi, Febo Cincotti, Donatella Mattia, M G Marciani, Serenella Salinari, A Colosimo, and Fabio Babiloni, "Persistent patterns of interconnection in time-varying cortical networks estimated from high-resolution EEG recordings in humans during a simple motor act", *Journal of Physics A: Mathematical and Theoretical*, vol. 41, no. 22, pp. 224014, 2008. (Cited on pages 29 and 129.)

[127] Mark E. J. Newman, "Assortative mixing in networks", *Physical Review Letters*, vol. 89, no. 20, pp. 208701, 2002. (Cited on page 30.)

[128] MEJ Newman, "Mixing patterns in networks", *Physical Review E*, vol. 67, no. 2, 2003. (Cited on pages 30 and 48.)

[129] Jacob G Foster, David V Foster, Peter Grassberger, and Maya Paczuski, "Edge direction and the structure of networks", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 24, pp. 10815–10820, 2010. (Cited on page 30.)

[130] Bowen Yan and Jianxi Luo, "Multicore-periphery structure in networks", *arXiv*, 2016. (Cited on pages 30 and 33.)

[131] Linton C Freeman, Stephen P Borgatti, and Douglas R White, "Centrality in valued graphs: A measure of betweenness based on network flow", *Social Networks*, vol. 13, no. 2, pp. 141–154, 1991. (Cited on page 30.)

[132] Mark D Humphries and Kevin Gurney, "Network 'small-world-ness': a quantitative method for determining canonical network equivalence", *PLoS ONE*, vol. 3, no. 4, pp. e0002051, 2008. (Cited on pages 30 and 50.)

[133] Claus C Hilgetag and Alexandros Goulas, "Is the brain really a small-world network?", *Brain Structure and Function*, vol. 221, no. 4, pp. 2361–2366, 2015. (Cited on page 30.)

[134] Uri Alon, "Network motifs: theory and experimental approaches", vol. 8, no. 6, pp. 450–461, 2007. (Cited on pages 31 and 95.)

[135] S Schuster, T Pfeiffer, F Moldenhauer, I Koch, and T Dandekar, "Exploring the pathway structure of metabolism: decomposition into subnetworks and application to Mycoplasma pneumoniae", *Bioinformatics*, vol. 18, no. 2, pp. 351–361, 2002. (Cited on page 31.)

[136] Petter Holme, Mikael Huss, and Hawoong Jeong, "Subnetwork hierarchies of biochemical pathways", *Bioinformatics*, vol. 19, no. 4, pp. 532–538, 2003. (Cited on page 31.)

[137] B W Kernighan and S Lin, "An efficient heuristic procedure for partitioning graphs", *Bell System Technical Journal, The*, vol. 49, no. 2, pp. 291–307, 1970. (Cited on pages 31 and 34.)

[138] Andrea Lancichinetti and Santo Fortunato, "Consensus clustering in complex networks", *Scientific Reports*, vol. 2, pp. 336, 2012. (Cited on pages 31, 34, and 37.)

[139] Mark E. J. Newman and Aaron Clauset, "Structure and inference in annotated networks.", *Nature Communications*, vol. 7, pp. 11863, 2016. (Cited on pages 31 and 38.)

[140] Thomas M J Fruchterman and Edward M Reingold, "Graph drawing by force-directed placement", *Software: Practice and Experience*, vol. 21, no. 11, pp. 1129–1164, 1991. (Cited on page 32.)

[141] Tomihisa Kamada and Satoru Kawai, "An algorithm for drawing general undirected graphs", *Information Processing Letters*, vol. 31, no. 1, pp. 7–15, 1989. (Cited on pages 32 and 114.)

[142] Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J Honey, Van J Wedeen, and Olaf Sporns, "Mapping the structural core of human cerebral cortex", *PLoS Biology*, vol. 6, no. 7, pp. e159, 2008. (Cited on page 32.)

180

[143] E Ravasz, A L Somera, D A Mongru, Z N Oltvai, and Albert-László Barabási, "Hierarchical organization of modularity in metabolic networks", *Science*, vol. 297, no. 5586, pp. 1551–1555, 2002. (Cited on pages 33 and 47.)

[144] Gergely Palla, Illés J Farkas, Péter Pollner, Imre Derényi, and Támás Vicsek, "Directed network modules", *New Journal of Physics*, vol. 9, pp. 186, 2007. (Cited on page 33.)

[145] Illés J Farkas, Dániel Ábel, Gergely Palla, and Támás Vicsek, "Weighted network modules", *New Journal of Physics*, vol. 9, pp. 180, 2007. (Cited on page 33.)

[146] Gergely Palla, Imre Derényi, Illés J Farkas, and Támás Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society", *Nature*, vol. 435, pp. 814–818, 2005. (Cited on page 33.)

[147] Mark E. J. Newman, "Spectral methods for community detection and graph partitioning", *Physical Review E*, vol. 88, no. 4, pp. 042822, 2013. (Cited on page 33.)

[148] Mark E. J. Newman, "Communities, modules and large-scale structure in networks", *Nature Physics*, vol. 8, no. 1, pp. 25–31, 2012. (Cited on pages 33, 42, and 103.)

[149] Paul W Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt, "Stochastic blockmodels: First steps", *Social Networks*, vol. 5, no. 2, pp. 109–137, 1983. (Cited on page 33.)

[150] Brian Karrer and Mark E. J. Newman, "Stochastic blockmodels and community structure in networks", *Physical Review E*, vol. 83, no. 1, pp. 016107, 2011. (Cited on page 33.)

[151] Daniel B Larremore, Aaron Clauset, and Abigail Z Jacobs, "Efficiently inferring community structure in bipartite networks", *Physical Review E*, vol. 90, no. 1, pp. 012805, 2014. (Cited on page 33.)

[152] Mark E. J. Newman, "Modularity and community structure in networks", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 23, pp. 8577–8582, 2006. (Cited on pages 33, 117, and 127.)

[153] Mark E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices", *Physical Review E*, vol. 74, no. 3, pp. 036104, 2006. (Cited on pages 33, 117, and 127.)

[154] E A Leicht and Mark E. J. Newman, "Community structure in directed networks", *Physical Review Letters*, vol. 100, no. 11, pp. 118703, 2008. (Cited on pages 33, 42, 117, and 127.)

[155] Michael Molloy and Bruce Reed, "A critical point for random graphs with a given degree sequence", *Random Structures & Algorithms*, vol. 6, no. 2-3, pp. 161–180, 1995. (Cited on pages 34 and 49.)

[156] Christoph Schmidt, Thomas Weiss, Thomas Lehmann, Herbert Witte, and Lutz Leistritz, "Extracting labeled topological patterns from samples of networks", *PLoS ONE*, vol. 8, no. 8, pp. e70497, 2013. (Cited on pages 34, 60, 62, 64, 84, 89, 93, and 145.)

[157] Ron Milo, Nadav Kashtan, Shalev Itzkovitz, Mark E. J. Newman, and Uri Alon, "On the uniform generation of random graphs with prescribed degree sequences", *arXiv:cond-mat/0312028*, 2003. (Cited on pages 34, 49, and 50.)

[158] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre, "Fast unfolding of communities in large networks", *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, pp. P10008, 2008. (Cited on pages 34, 117, and 127.)

[159] Martin Rosvall and Carl T Bergstrom, "Mapping change in large networks", *PLoS ONE*, vol. 5, no. 1, pp. e8694, 2010. (Cited on pages 34, 80, and 138.)

[160] Aaron Clauset, Mark E. J. Newman, and Cristopher Moore, "Finding community structure in very large networks", *Physical Review E*, vol. 70, no. 6, pp. 066111, 2004. (Cited on pages 34 and 117.)

[161] Mark E. J. Newman, "Fast algorithm for detecting community structure in networks", *Physical Review E*, vol. 69, no. 6, pp. 066133, 2004. (Cited on page 34.)

[162] Pascal Pons and Matthieu Latapy, "Computing communities in large networks using random walks", in *Computer and Information Sciences - Iscis 2005, Proceedings*, Berlin, Heidelberg, 2005, pp. 284–293, Springer Berlin Heidelberg. (Cited on pages 35 and 117.)

[163] Pascal Pons and Matthieu Latapy, "Computing communities in large networks using random walks", *Journal of Graph Algorithms and Applications*, vol. 10, no. 2, pp. 191–218, 2006. (Cited on pages 35 and 117.)

[164] K H Fischer and J A Hertz, *Spin Glasses*, Cambridge University Press, 1993. (Cited on page 35.)

[165] M Mezard, G Parisi, and M Virasoro, *Spin Glass Theory and Beyond*, vol. 9 of *An Introduction to the Replica Method and Its Applications*, World Scientific, 1986. (Cited on page 35.)

[166] F Y WU, "The Potts model", *Reviews of Modern Physics*, vol. 54, no. 1, pp. 235–268, 1982. (Cited on page 35.)

[167] Jörg Reichardt and Stefan Bornholdt, "Statistical mechanics of community detection", *Physical Review E*, vol. 74, no. 1, pp. 016110, 2006. (Cited on pages 35, 36, 44, 117, 121, 122, and 127.)

[168] Martin Rosvall and Carl T Bergstrom, "Maps of random walks on complex networks reveal community structure", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 4, pp. 1118–1123, 2008. (Cited on pages 36 and 117.)

[169] Peter D Grünwald, In Jae Myung, and Mark A Pitt, *Advances in Minimum Description Length*, Theory and Applications. The MIT Press, 2005. (Cited on page 36.)

[170] M Rosvall, D Axelsson, and C T Bergstrom, "The map equation", *The European Physical Journal Special Topics*, vol. 178, no. 1, pp. 13–23, 2009. (Cited on pages 36 and 37.)

[171] David Huffman, "A Method for the construction of minimum-redundancy codes", *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, 1952. (Cited on page 36.)

[172] Derek Greene, Dónal Doyle, and Pádraig Cunningham, "Tracking the Evolution of Communities in Dynamic Social Networks", in *2010 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2010)*. pp. 176–183, IEEE. (Cited on pages 37, 68, 129, 130, and 162.)

[173] William M Rand, "Objective Criteria for the Evaluation of Clustering Methods", *Journal of the American Statistical Association*, vol. 66, no. 336, pp. 846–850, 1971. (Cited on pages 39 and 40.)

[174] Matthijs J Warrens, "On the Equivalence of Cohen's Kappa and the Hubert-Arabie Adjusted Rand Index", *Journal of Classification*, vol. 25, no. 2, pp. 177–183, 2008. (Cited on pages 39 and 40.)

[175] Lawrence Hubert and Phipps Arabie, "Comparing partitions", *Journal of Classification*, vol. 2, no. 1, pp. 193–218, 1985. (Cited on page 40.)

[176] Marina Meilă, "Comparing clusterings — an information based distance", *Journal of Multivariate Analysis*, vol. 98, no. 5, pp. 873–895, 2007. (Cited on pages 40 and 42.)

[177] Michael Levandowsky and David Winter, "Distance between Sets", *Nature*, vol. 234, no. 5323, pp. 34–35, 1971. (Cited on page 40.)

[178] Trevor F Cox and Michael A A Cox, *Multidimensional Scaling*, Chapman and Hall/CRC, second edition edition, 2000. (Cited on pages 40 and 126.)

[179] Brian Karrer, Elizaveta Levina, and Mark E. J. Newman, "Robustness of community structure in networks", *Physical Review E*, vol. 77, no. 4, pp. 046119, 2008. (Cited on pages 41, 42, 43, 82, and 135.)

[180] Stijn Dongen, "Performance criteria for graph clustering and Markov cluster experiments", Tech. Rep., 2000. (Cited on page 42.)

[181] Vladimir I Levenshtein, "*Binary codes capable of correcting deletions, insertions, and reversals*", *Soviet Physics-Doklady*, vol. 10, no. 8, pp. 707–710, 1966. (Cited on page 42.)

[182] Sanjoy Dasgupta, Christos H Papadimitriou, and Umesh Vazirani, *Algorithms*, McGrawHill, 1st edition edition, 2006. (Cited on page 42.)

[183] Mark E. J. Newman and M Girvan, "Finding and evaluating community structure in networks", *Physical Review E*, vol. 69, no. 2, pp. 026113, 2004. (Cited on page 42.)

[184] Roger Guimerà, Marta Sales-Pardo, and Luís A Nunes Amaral, "Module identification in bipartite and directed networks", *Physical Review E*, vol. 76, pp. 036102, 2007. (Cited on pages 42 and 43.)

[185] Alex Arenas, J Duch, A Fernández, and S Gómez, "Size reduction of complex networks preserving modularity", *New Journal of Physics*, vol. 9, no. 6, pp. 176–176, 2007. (Cited on pages 43 and 44.)

[186] Santo Fortunato and Marc Barthélemy, "Resolution limit in community detection", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 1, pp. 36–41, 2007. (Cited on page 44.)

[187] Ulrik Brandes, Daniel Delling, Marco Gaertler, R Goerke, Martin Hoefer, Zoran Nikoloski, and Dorothea Wagner, "Maximizing Modularity is hard", *arXiv*, 2006. (Cited on page 44.)

[188] Benjamin H Good, Yves-Alexandre de Montjoye, and Aaron Clauset, "Performance of modularity maximization in practical contexts", *Physical Review E*, vol. 81, no. 4, pp. 046106, 2010. (Cited on page 44.)

[189] Aaron Clauset, "Finding local community structure in networks", *Physical Review E*, vol. 72, no. 2, pp. 026132, 2005. (Cited on page 44.)

[190] Peter J Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis", *Journal of Computational and Applied Mathematics*, vol. 20, no. 0, pp. 53–65, 1987. (Cited on page 45.)

[191] E A Leicht, Petter Holme, and Mark E. J. Newman, "Vertex similarity in networks", *Physical Review E*, vol. 73, no. 2 Pt 2, pp. 026120, 2006. (Cited on pages 46 and 47.)

[192] Deepayan Chakrabarti, Ravi Kumar, and Andrew Tomkins, "Evolutionary clustering", in *the 12th ACM SIGKDD international conference*, New York, New York, USA, 2006, p. 554, ACM Press. (Cited on pages 46 and 47.)

[193] Lada A Adamic and Eytan Adar, "Friends and neighbors on the web", *Social Networks*, vol. 25, no. 3, pp. 211–230, 2003. (Cited on pages 46 and 47.)

[194] Andy M Yip and Steve Horvath, "Gene network interconnectedness and the generalized topological overlap measure", *BMC Bioinformatics*, vol. 8, no. 1, pp. 22, 2007. (Cited on page 47.)

[195] Michael B Eisen, Paul T Spellman, Patrick O Brown, and David Botstein, "Cluster analysis and display of genome-wide expression patterns", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 25, pp. 14863–14868, 1998. (Cited on page 47.)

[196] Jing Zhao, Lili Miao, Jian Yang, Haiyang Fang, Qian-Ming Zhang, Min Nie, Petter Holme, and Tao Zhou, "Prediction of Links and Weights in Networks by Reliable Routes", *Scientific Reports*, vol. 5, pp. 12261, 2015. (Cited on page 47.)

[197] Jacob Cohen, "A Coefficient of Agreement for Nominal Scales", vol. 20, no. 1, pp. 37–46, 1960. (Cited on pages 47 and 108.)

[198] Ernesto Estrada, "Spectral scaling and good expansion properties in complex networks", *Epl*, vol. 73, no. 4, pp. 649–655, 2007. (Cited on page 47.)

[199] Ernesto Estrada, "Topological structural classes of complex networks", *Physical Review E*, vol. 75, no. 1, pp. 016103, 2007. (Cited on page 47.)

[200] Paul Erdős and Alfréd Rényi, "On random graphs", *Publicationes Mathematicae (Debrecen)*, vol. 6, pp. 290–297, 1959. (Cited on page 47.)

[201] Réka Albert, Hawoong Jeong, and Albert-László Barabási, "Error and attack tolerance of complex networks", *Nature*, vol. 406, no. 6794, pp. 378–382, 2000. (Cited on page 47.)

[202] Mathew Penrose, *Random geometric graphs*, Oxford University Press, 2003. (Cited on page 47.)

[203] Shalev Itzkovitz and Uri Alon, "Subgraphs and network motifs in geometric networks", *Physical Review E*, vol. 71, no. 2, pp. 026117, 2005. (Cited on page 47.)

[204] Duncan J Watts, "Networks, dynamics, and the small-world phenomenon", *American Journal of Sociology*, vol. 105, no. 2, pp. 493–527, 1999. (Cited on page 48.)

[205] Alessandro Vespignani, "Complex networks: The fragility of interdependency", *Nature*, vol. 464, no. 7291, pp. 984–985, 2010. (Cited on page 48.)

[206] Christian M Schneider, André A Moreira, José S Andrade, Shlomo Havlin, and Hans J Herrmann, "Mitigation of malicious attacks on networks.", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 10, pp. 3838–3841, 2011. (Cited on page 48.)

[207] Nicholas J Gotelli and Werner Ulrich, "Statistical challenges in null model analysis", *Oikos*, vol. 121, no. 2, pp. 171–180, 2012. (Cited on pages 48, 49, and 51.)

[208] Nicholas J Gotelli and Gary R Graves, *Null models in ecology*, Smithsonian Institution Press, Washington, DC, 1996. (Cited on pages 48 and 95.)

[209] Chiara Orsini, Marija M Dankulov, Pol Colomer-de Simón, Almerima Jamakovic, Priya Mahadevan, Amin Vahdat, Kevin E Bassler, Zoltán Toroczkai, Marian Boguna, Guido Caldarelli, Santo Fortunato, and Dmitri Krioukov, "Quantifying randomness in real networks", *Nature Communications*, vol. 6, pp. 8627, 2015. (Cited on pages 48, 50, and 95.)

[210] Sergei Maslov, Kim Sneppen, and Alexei Zaliznyak, "Detection of topological patterns in complex networks: correlation profile of the internet", *Physica A: Statistical Mechanics and its Applications*, vol. 333, pp. 529–540, 2004. (Cited on pages 48 and 49.)

[211] Sami Hanhijärvi, Gemma Garriga, and Kai Puolamäki, "Randomization techniques for graphs", Tech. Rep., 2009. (Cited on page 48.)

[212] Joseph Blitzstein and Persi Diaconis, "A sequential importance sampling algorithm for generating random graphs with prescribed degrees", *Internet Mathematics*, vol. 6, no. 4, pp. 489–522, 2011. (Cited on pages 48 and 62.)

[213] Sergei Maslov and Kim Sneppen, "Specificity and stability in topology of protein networks", *Science*, vol. 296, no. 5569, pp. 910–913, 2002. (Cited on pages 48 and 49.)

[214] Shalev Itzkovitz, Ron Milo, Nadav Kashtan, Mark E. J. Newman, and Uri Alon, "Reply to "Comment on 'Subgraphs in random networks' "", *Physical Review E*, vol. 70, no. 5, pp. 058102, 2004. (Cited on pages 48 and 49.)

[215] Fabien Viger and Matthieu Latapy, "Efficient and simple generation of random simple connected graphs with prescribed degree sequence", *Journal of Complex Networks*, vol. 4, no. 1, pp. 15–37, 2015. (Cited on page 48.)

[216] Yael Artzy-Randrup, Sarel J Fleishman, Nir Ben-Tal, and Lewi Stone, "Comment on "Network motifs: simple building blocks of complex networks" and "Superfamilies of

evolved and designed networks"", *Science*, vol. 305, no. 5687, pp. 1107–1107c, 2004. (Cited on pages 49 and 95.)

[217] Moritz Emanuel Beber, Christoph Fretter, Shubham Jain, Nikolaus Sonnenschein, Matthias Müller-Hannemann, and Marc-Thorsten Hütt, "Artefacts in statistical analyses of network motifs: general framework and application to metabolic networks", *Journal of the Royal Society, Interface*, vol. 9, no. 77, pp. 3426–3435, 2012. (Cited on page 49.)

[218] Mark E. J. Newman, S H Strogatz, and Duncan J Watts, "Random graphs with arbitrary degree distributions and their applications", *Physical Review E*, vol. 64, no. 2, pp. 026118, 2001. (Cited on page 49.)

[219] Oliver D King, "Comment on "Subgraphs in random networks"", *Physical Review E*, vol. 70, no. 5, pp. 058101, 2004. (Cited on pages 49 and 50.)

[220] Mark E. J. Newman, "Random graphs as models of networks", in *Handbook of Graphs and Networks: From the Genome to the Internet*, Stefan Bornholdt and Hans Georg Schuster, Eds., pp. 35–68. Wiley-VCH Verlag GmbH & Co. KGaA, 2005. (Cited on page 49.)

[221] Ravi Kannan, Prasad Tetali, and Santosh Vempala, "Simple Markov-chain algorithms for generating bipartite graphs and tournaments", *Random Structures & Algorithms*, vol. 14, no. 4, pp. 293–308, 1999. (Cited on page 49.)

[222] D Aldous and U Vazirani, ""Go with the winners" algorithms", in *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on.* 1994, pp. 492–501, IEEE Comput. Soc. Press. (Cited on page 50.)

[223] Gerrit Ansmann and Klaus Lehnertz, "Surrogate-assisted analysis of weighted functional brain networks", *Journal of Neuroscience Methods*, vol. 208, no. 2, pp. 165–172, 2012. (Cited on page 50.)

[224] Martijn P van den Heuvel and Olaf Sporns, "Rich-Club Organization of the Human Connectome", *The Journal of Neuroscience*, vol. 31, no. 44, pp. 15775–15786, 2011. (Cited on pages 50 and 51.)

[225] Olaf Sporns and Jonathan D Zwi, "The small world of the cerebral cortex", *Neuroinformatics*, vol. 2, no. 2, pp. 145–162, 2004. (Cited on page 51.)

[226] Martijn P van den Heuvel and Hilleke E Hulshoff Pol, "Exploring the brain network: A review on resting-state fMRI functional connectivity", *European Neuropsychopharmacology*, vol. 20, no. 8, pp. 519–534, 2010. (Cited on page 57.)

[227] Helen H Shen, "Core Concept: Resting-state connectivity", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 46, pp. 14115–14116, 2015. (Cited on pages 57 and 105.)

[228] H W Kuhn, "The Hungarian method for the assignment problem", *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955. (Cited on page 65.)

[229] Kenneth C Gilbert and Ruth B Hofstra, "Multidimensional assignment problems", *Decision Sciences*, vol. 19, no. 2, pp. 306–321, 1988. (Cited on page 65.)

[230] David W Pentico, "Assignment problems: A golden anniversary survey", *European Journal of Operational Research*, vol. 176, no. 2, pp. 774–793, 2007. (Cited on page 65.)

[231] Karin Schiecke, Christoph Schmidt, Diana Piper, Peter Putsche, Martha Feucht, Herbert Witte, and Lutz Leistritz, "Assignment of Empirical Mode Decomposition Components and Its Application to Biomedical Signals", *Methods of Information in Medicine*, vol. 54, no. 5, pp. 461–473, 2015. (Cited on pages 65 and 145.)

[232] "lpSolve", *lpsolve.sourceforge.net*. (Cited on page 67.)

[233] Gergely Palla, Albert-László Barabási, and Támás Vicsek, "Quantifying social group evolution", *Nature*, vol. 446, no. 7136, pp. 664–667, 2007. (Cited on pages 70 and 162.)

[234] Andrea Lancichinetti, Santo Fortunato, and Filippo Radicchi, "Benchmark graphs for testing community detection algorithms", *Physical Review E*, vol. 78, no. 4, pp. 046110, 2008. (Cited on page 74.)

[235] Anne Condon and Richard M Karp, "Algorithms for graph partitioning on the planted partition model", in *Randomization, Approximation, and Combinatorial Optimization*, pp. 221–232. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999. (Cited on page 74.)

[236] R J Hyndman and Y N Fan, "Sample quantiles in statistical packages", *American Statistician*, vol. 50, no. 4, pp. 361–365, 1996. (Cited on page 75.)

[237] Kaisa Miettinen, *Nonlinear Multiobjective Optimization*, vol. 12 of *International Series in Operations Research & Management Science*, Springer US, Boston, MA, 1998. (Cited on pages 75 and 76.)

[238] Matthias Ehrgott, *Multicriteria Optimization*, Springer-Verlag, Berlin/Heidelberg, second edition edition, 2005. (Cited on page 76.)

[239] Petri Eskelinen and Kaisa Miettinen, "Trade-off analysis approach for interactive nonlinear multiobjective optimization", *Or Spectrum*, vol. 34, no. 4, pp. 803–816, 2012. (Cited on page 76.)

[240] Yi Cao, "Pareto Front". (Cited on page 77.)

[241] Andreas Mierau, Thorben Hülsdünker, and Heiko K Strüder, "Changes in cortical activity associated with adaptive behavior during repeated balance perturbation of unpredictable timing", *Frontiers in Behavioral Neuroscience*, vol. 9, pp. 233, 2015. (Cited on pages 78, 130, and 131.)

[242] Dirk Eddelbuettel and Romain François, "Rcpp: Seamless Rand C++Integration", *Journal of Statistical Software*, vol. 40, no. 8, 2011. (Cited on page 79.)

[243] David Gfeller, Jean-Cédric Chappelier, and Paolo De Los Rios, "Finding instabilities in the community structure of complex networks", *Physical Review E*, vol. 72, no. 5, pp. 056135, 2005. (Cited on pages 82 and 135.)

[244] Lutz Leistritz, Thomas Weiss, Jaroslav Ionov, Karl-Jürgen Bär, Wolfgang H R Miltner, and Herbert Witte, "Connectivity analysis of somatosensory evoked potentials in patients with major depression", *Methods of Information in Medicine*, vol. 49, no. 5, pp. 484–491, 2010. (Cited on pages 84 and 85.)

[245] S Lautenbacher, J Spernal, W Schreiber, and J C Krieg, "Relationship between clinical pain complaints and pain sensitivity in patients with depression and panic disorder.", *Psychosomatic Medicine*, vol. 61, no. 6, pp. 822–827, 1999. (Cited on pages 84 and 93.)

[246] Lisa Renee Miller and Annmarie Cano, "Comorbid chronic pain and depression: who is at risk?", *The Journal of Pain*, vol. 10, no. 6, pp. 619–627, 2009. (Cited on page 84.)

[247] Sabrina Klauenberg, Christoph Maier, Hans-Jörg Assion, Axel Hoffmann, Elena K Krumova, Walter Magerl, Andrea Scherens, Rolf-Detlef Treede, and Georg Juckel, "Depression and changed pain perception: Hints for a central disinhibition mechanism", *Pain*, vol. 140, no. 2, pp. 332–343, 2008. (Cited on page 84.)

[248] Irina A Strigo, Alan N Simmons, Scott C Matthews, Arthur D Craig, and Martin P Paulus, "Association of major depressive disorder with altered functional brain response during anticipation and processing of heat pain", *Archives of General Psychiatry*, vol. 65, no. 11, pp. 1275–1284, 2008. (Cited on pages 84 and 92.)

[249] Karl-Jürgen Bär, Stanislaw Brehm, Michael Karl Boettger, Silke Boettger, Gerd Wagner, and Heinrich Sauer, "Pain perception in major depression depends on pain modality", *Pain*, vol. 117, no. 1-2, pp. 97–103, 2005. (Cited on pages 84 and 93.)

[250] Karl-Jürgen Bär, Gerd Wagner, Mandy Koschke, Silke Boettger, Michael Karl Boettger, Ralf Schlösser, and Heinrich Sauer, "Increased prefrontal activation during pain perception in major depression", *Biological Psychiatry*, vol. 62, no. 11, pp. 1281–1287, 2007. (Cited on pages 84 and 92.)

[251] Stefan Lautenbacher and Jürgen-Christian Krieg, "Pain perception in psychiatric disorders: a review of the literature", *Journal of Psychiatric Research*, vol. 28, no. 2, pp. 109–122, 1994. (Cited on page 84.)

[252] Janneke Terhaar, Filipa Campos Viola, Marcel Franz, Sandy Berger, Karl-Jürgen Bär, and Thomas Weiss, "Differential processing of laser stimuli by Aδ and C fibres in major depression.", *Pain*, vol. 152, no. 8, pp. 1796–1802, 2011. (Cited on pages 84 and 93.)

[253] G D Iannetti and A Mouraux, "From the neuromatrix to the pain matrix (and back)", *Experimental Brain Research*, vol. 205, no. 1, pp. 1–12, 2010. (Cited on page 84.)

[254] Winfried Meissner, Thomas Weiss, Ralf H Trippe, Holger Hecht, Clemens Krapp, and Wolfgang H R Miltner, "Acupuncture decreases somatosensory evoked potential amplitudes to noxious stimuli in anesthetized volunteers", *Anesthesia & Analgesia*, vol. 98, no. 1, pp. 141–147, 2004. (Cited on page 85.)

[255] Thomas Weiss, Kai Kumpf, Jens Ehrhardt, Ingmar Gutberlet, and Wolfgang H R Miltner, "A bioadaptive approach for experimental pain research in humans using laser-evoked brain potentials", *Neuroscience Letters*, vol. 227, no. 2, pp. 95–98, 1997. (Cited on page 85.)

[256] Richard J Davidson, Diego Pizzagalli, Jack B Nitschke, and Katherine Putnam, "Depression: perspectives from affective neuroscience", *Annual Review of Psychology*, vol. 53, pp. 545–574, 2002. (Cited on page 92.)

[257] Carsten Diener, Christine Kuehner, Wencke Brusniak, Bettina Ubl, Michèle Wessa, and Herta Flor, "A meta-analysis of neurofunctional imaging studies of emotion and cognition in major depression", *NeuroImage*, vol. 61, no. 3, pp. 677–685, 2012. (Cited on page 92.)

[258] Isabella Mutschler, Tonio Ball, Johanna Wankerl, and Irina A Strigo, "Pain and emotion in the insular cortex: evidence for functional reorganization in major depression", *Neuroscience Letters*, vol. 520, no. 2, pp. 204–209, 2012. (Cited on page 92.)

[259] Richard J Davidson and William Irwin, "The functional neuroanatomy of emotion and affective style", *Trends in Cognitive Sciences*, vol. 3, no. 1, pp. 11–21, 1999. (Cited on page 92.)

[260] Richard J Davidson, Daren C Jackson, and Ned H Kalin, "Emotion, plasticity, context, and regulation: perspectives from affective neuroscience", *Psychological Bulletin*, vol. 126, no. 6, pp. 890–909, 2000. (Cited on page 92.)

[261] Tom Johnstone, Carien M van Reekum, Heather L Urry, Ned H Kalin, and Richard J Davidson, "Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression", *The Journal of Neuroscience*, vol. 27, no. 33, pp. 8877–8884, 2007. (Cited on page 92.)

[262] Karl-Jürgen Bär, Stanislaw Brehm, Michael Karl Boettger, Gerd Wagner, Silke Boettger, and Heinrich Sauer, "Decreased sensitivity to experimental pain in adjustment disorder", *European Journal of Pain*, vol. 10, no. 5, pp. 467–471, 2006. (Cited on page 93.)

[263] Johannes F Knabe, Chrystopher L Nehaniv, and Maria J Schilstra, "Do motifs reflect evolved function? - No convergent evolution of genetic regulatory network subgraph topologies", *Biosystems*, vol. 94, no. 1-2, pp. 68–74, 2008. (Cited on page 95.)

[264] Piers J Ingram, Michael P H Stumpf, and Jaroslav Stark, "Network motifs: structure does not determine function", *BMC Genomics*, vol. 7, no. 1, pp. 108, 2006. (Cited on page 95.)

[265] Oren Shoval and Uri Alon, "SnapShot: network motifs", *CELL*, vol. 143, no. 2, pp. 326–326.e1, 2010. (Cited on page 95.)

[266] Madalina E Tivarus, Britta Pester, Christoph Schmidt, Thomas Lehmann, Tong Zhu, Jianhui Zhong, Lutz Leistritz, and Giovanni Schifitto, "Are Structural Changes Induced by Lithium in the HIV Brain Accompanied by Changes in Functional Connectivity?", *PLoS ONE*, vol. 10, no. 10, pp. e0139118, 2015. (Cited on pages 96, 97, 99, 100, and 145.)

[267] Giovanni Schifitto, Jianhui Zhong, David Gill, Derick R Peterson, Michelle D Gaugh, Tong Zhu, Madalina Tivarus, Kim Cruttenden, Sanjay B Maggirwar, Howard E Gendelman, Stephen Dewhurst, and Harris A Gelbard, "Lithium therapy for human immunodeficiency virus type 1-associated neurocognitive impairment", *Journal of NeuroVirology*, vol. 15, no. 2, pp. 176–186, 2009. (Cited on page 97.)

[268] H Garavan, T J Ross, S J Li, and E A Stein, "A parametric manipulation of central executive functioning.", *Cerebral Cortex*, vol. 10, no. 6, pp. 585–592, 2000. (Cited on page 97.)

[269] Frank Wilcoxon, "Individual Comparisons by Ranking Methods", *Biometrics bulletin*, vol. 1, no. 6, pp. 80, 1945. (Cited on page 99.)

[270] A Cnaan, N M Laird, and P Slasor, "Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data.", *Statistics in medicine*, vol. 16, no. 20, pp. 2349–2380, 1997. (Cited on page 100.)

[271] Wise Young, "Review of Lithium Effects on Brain and Blood", *Cell Transplantation*, vol. 18, no. 9, pp. 951–975, 2009. (Cited on page 102.)

[272] Danielle S Bassett, Daniel L Greenfield, Andreas Meyer-Lindenberg, Daniel R Weinberger, Simon W Moore, and Edward T Bullmore, "Efficient physical embedding of topologically complex information processing networks in brains and computer circuits", *PLoS Computational Biology*, vol. 6, no. 4, pp. e1000748, 2010. (Cited on page 103.)

[273] Christoph Schmidt, Britta Pester, Nicole Schmid-Hertel, Herbert Witte, Axel Wismüller, and Lutz Leistritz, "A Multivariate Granger Causality Concept towards Full Brain Functional Connectivity", *PLoS ONE*, vol. 11, no. 4, pp. e0153105, 2016. (Cited on pages 104, 107, 117, 120, 127, 145, 147, 148, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, and 161.)

[274] Boris Gourévitch, Régine Le Bouquin-Jeannès, and Gérard Faucon, "Linear and nonlinear causality between signals: methods, examples and neurophysiological applications", *Biological Cybernetics*, vol. 95, no. 4, pp. 349–369, 2006. (Cited on page 106.)

[275] S M Smith, M Jenkinson, M W Woolrich, C F Beckmann, TEJ Behrens, H Johansen-Berg, P R Bannister, M De Luca, I Drobnjak, D E Flitney, R K Niazy, J Saunders, J Vickers, Y Y Zhang, N De Stefano, J M Brady, and P M Matthews, "Advances in functional and structural MR image analysis and implementation as FSL", *NeuroImage*, vol. 23, pp. S208–S219, 2004. (Cited on page 107.)

[276] Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik, *Statistical Methods for Rates and Proportions*, Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, third edition edition, 2003. (Cited on page 108.)

[277] Michael Schmuhl, "graphopt". (Cited on page 122.)

[278] Britta Pester, Lutz Leistritz, Herbert Witte, and Axel Wismueller, "Exploring effective connectivity by a Granger Causality approach with embedded dimension reduction", *Biomedizinische Technik / Biomedical engineering*, 2013. (Cited on pages 123 and 148.)

[279] Gábor Csárdi and Tamás Nepusz, "The igraph software package for complex network research", *InterJournal*, vol. Complex Systems, pp. 1695, 2006. (Cited on page 126.)

[280] Albert-László Barabási, *Network Science*, barabasi.com. (Cited on page 129.)

[281] Sitaram Asur, Srinivasan Parthasarathy, and Duygu Ucar, "An event-based framework for characterizing the evolutionary behavior of interaction graphs", *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 3, no. 4, pp. 16–36, 2009. (Cited on pages 129, 130, and 162.)

[282] Daniel J Fenn, Mason A Porter, Mark McDonald, Stacy Williams, Neil F Johnson, and Nick S Jones, "Dynamic communities in multichannel data: An application to the

foreign exchange market during the 2007–2008 credit crisis", *Chaos*, vol. 19, no. 3, pp. 033119, 2009. (Cited on page 130.)

[283] J V Jacobs and F B Horak, "Cortical control of postural responses", *Journal of Neural Transmission*, vol. 114, no. 10, pp. 1339–1348, 2007. (Cited on page 130.)

[284] B E Maki and W E McIlroy, "Cognitive demands and cortical control of human balance-recovery reactions", *Journal of Neural Transmission*, vol. 114, no. 10, pp. 1279–1296, 2007. (Cited on page 130.)

[285] Selma Papegaaij, Wolfgang Taube, Stéphane Baudry, Egbert Otten, and Tibor Hortobágyi, "Aging causes a reorganization of cortical and spinal control of posture", *Frontiers in Aging Neuroscience*, vol. 6, no. MAR, 2014. (Cited on page 130.)

[286] T Hülsdünker, A Mierau, C Neeb, H Kleinöder, and H K Strüder, "Cortical processes associated with continuous balance control as revealed by EEG spectral power.", *Neuroscience Letters*, vol. 592, pp. 1–5, 2015. (Cited on page 134.)

[287] Parashkev Nachev, Christopher Kennard, and Masud Husain, "Functional role of the supplementary and pre-supplementary motor areas", *Nature Reviews Neuroscience*, vol. 9, no. 11, pp. 856–869, 2008. (Cited on page 134.)

[288] François-Benoît Vialatte, Monique Maurice, Justin Dauwels, and Andrzej Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives", *Progress in Neurobiology*, vol. 90, no. 4, pp. 418–438, 2010. (Cited on page 135.)

[289] Keiji Tanaka, "Mechanisms of visual object recognition: monkey and human studies", *Current Opinion in Neurobiology*, vol. 7, no. 4, pp. 523–529, 1997. (Cited on page 135.)

[290] Giacomo Rizzolatti and Massimo Matelli, "Two different streams form the dorsal visual system: anatomy and functions", *Experimental Brain Research*, vol. 153, no. 2, pp. 146–157, 2003. (Cited on page 135.)

[291] M Guerraz and A M Bronstein, "Ocular versus extraocular control of posture and equilibrium", *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 38, no. 6, pp. 391–398, 2008. (Cited on page 135.)

[292] John Hopcroft, Omar Khan, Brian Kulis, and Bart Selman, "Tracking evolving communities in large linked networks", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101 Suppl 1, no. Supplement 1, pp. 5249–5253, 2004. (Cited on page 135.)

[293] Mikko Kivelä, Alex Arenas, Marc Barthélemy, James P Gleeson, Yamir Moreno, and Mason A Porter, "Multilayer networks", *Journal of Complex Networks*, vol. 2, no. 3, pp. 203–271, 2014. (Cited on pages 138 and 143.)

[294] Peter J Mucha, Thomas Richardson, Kevin Macon, Mason A Porter, and Jukka-Pekka Onnela, "Community structure in time-dependent, multiscale, and multiplex networks", *Science*, vol. 328, no. 5980, pp. 876–878, 2010. (Cited on page 138.)

[295] Peter J Mucha and Mason A Porter, "Communities in multislice voting networks", *Chaos*, vol. 20, no. 4, pp. 041108, 2010. (Cited on page 138.)

[296] Eric Jonas and Konrad Kording, "Could a neuroscientist understand a microprocessor?", *bioRxiv*, p. 055624, 2016. (Cited on page 142.)

[297] Stefan Haufe, Vadim V Nikulin, Klaus-Robert Müller, and Guido Nolte, "A critical assessment of connectivity measures for EEG data: a simulation study.", *NeuroImage*, vol. 64, no. 1, pp. 120–133, 2013. (Cited on page 143.)

[298] A Varma and B O Palsson, "Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110.", *Applied and Environmental Microbiology*, vol. 60, no. 10, pp. 3724–3731, 1994. (Cited on page 143.)

[299] Stefan Schuster, Luís F de Figueiredo, and Christoph Kaleta, "Predicting novel pathways in genome-scale metabolic networks", *Biochemical Society Transactions*, vol. 38, no. 5, pp. 1202–1205, 2010. (Cited on page 143.)

[300] C W J Granger, "Investigating causal relations by econometric models and cross-spectral methods", *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969. (Cited on page 146.)

[301] Helmut Lütkepohl, *New Introduction to Multiple Time Series Analysis*, Springer, 1st edition edition, 2005. (Cited on pages 146, 147, and 150.)

[302] D Hemmelmann, M Ungureanu, W Hesse, T Wüstenberg, J R Reichenbach, O W Witte, Herbert Witte, and Lutz Leistritz, "Modelling and analysis of time-variant directed interrelations between brain regions based on BOLD-signals", *NeuroImage*, vol. 45, no. 3, pp. 722–737, 2009. (Cited on page 147.)

[303] H Akaike, "A new look at the statistical model identification", *IEEE transactions on automatic control*, vol. 19, no. 6, pp. 716–723, 1974. (Cited on page 147.)

[304] Thomas Milde, Lutz Leistritz, Laura Astolfi, Wolfgang H R Miltner, Thomas Weiss, Fabio Babiloni, and Herbert Witte, "A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis

of laser-evoked brain potentials", *NeuroImage*, vol. 50, no. 3, pp. 960–969, 2010. (Cited on page 147.)

[305] Zhenyu Zhou, Mingzhou Ding, Yonghong Chen, Paul Wright, Zuhong Lu, and Yijun Liu, "Detecting directional influence in fMRI connectivity analysis using PCA based Granger causality", *Brain Research*, vol. 1289, pp. 22–29, 2009. (Cited on page 148.)

[306] Alessandro Londei, Alessandro D Ausilio, Demis Basso, and Marta Olivetti Belardinelli, "A new method for detecting causality in fMRI data of cognitive processing", *Cognitive Processing*, vol. 7, no. 1, pp. 42–52, 2005. (Cited on page 148.)

[307] Luiz A Baccalá and Koichi Sameshima, "Partial directed coherence: a new concept in neural structure determination", *Biological Cybernetics*, vol. 84, no. 6, pp. 463–474, 2001. (Cited on pages 148 and 149.)

[308] Luiz A Baccalá, K Sameshima, and D Y Takahashi, "Generalized partial directed coherence", in *Digital Signal Processing, 2007 15th International Conference on.* 2007, pp. 163–166, IEEE. (Cited on pages 148 and 149.)

[309] Anna Korzeniewska, Małgorzata Mańczak, Maciej Kamiński, Katarzyna J Blinowska, and Stefan Kasicki, "Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method.", *Journal of Neuroscience Methods*, vol. 125, no. 1-2, pp. 195–207, 2003. (Cited on page 149.)