Functional modelling of cortical macro-networks

A dissertation submitted for the degree of Doctor of Philosophy

Mihály Bányai

Department of Measurement and Information Systems,
Budapest University of Technology and Economics,
Budapest, Hungary

Advisor:

Dr. Fülöp Bazsó, Ph.D.,
Wigner Research Centre for Physics,
Hungarian Academy of Sciences, Budapest, Hungary

Consultant:

Dr. György Strausz, Ph.D.,
Department of Measurement and Information Systems,
Budapest University of Technology and Economics, Budapest, Hungary

Ph.D. School: Computer Science
Head: Dr. János Levendovszky
Ph.D. Program: Intelligent systems

Budapest, 2013

Abstract

This dissertation is about interaction networks of cortical areas, macroscopic ensembles of neurons connected by synapses, implementing behavioural functionalities such as perception, action and cognition. Different mathematical modelling approaches can be used to test hypotheses about such networks' structure, function, and the relation of the two.

The first part of the dissertation studies the structure of cortical macronetworks: by introducing a graph theoretical measure applicable for directed networks, I identify roles of vertices in information processing and introduce a network comparison and classification method. To create a random graph model of the cortical network, I propose a graph generation algorithm. I use the introduced methodology to investigate the information processing role of the prefrontal areas.

The second part of the dissertation studies functional subnetwoks of the cortex: a model consisting five areas, implementing associative learning, is proposed, and tested against measurement data using statistical model inversion and comparison. I investigate alterations of the connectivity model in schizophrenia to test possible deficits in cognitive control. I study the intrinsic connectivity changes in the prefrontal neural networks, possibly contributing to memory dysfunction.

Kivonat

Disszertációm az agykérgi területek hálózatáról szól. Ezek szinapszisok által összekötött makroszkopikus neuronegyüttesek, amelyek megvalósítják a viselkedéshez szükséges érzékelési, cselekvési és kognitív funkcionalitást. Különböző matematikai madellek segítségével tesztelhetünk hipotéziseket ezek struktúrájáról, funkciójáról és a kettő kapcsolatáról.

A disszertáció első fele a kérgi makrohálózatok struktúráját vizsgálja: egy irányított hálózatokra alkalmazható gráfelméleti mérték segítségével meghatározom az egyes csúcsok hálózati információfeldolgozásban játszott szerepét és bemutatok egy módszert hálózatok összehasonlítására és osztályozására. A kortikális hálózat randomgráf-modelljének létrehozása céljából új gráfgeneráló algoritmust javaslok. A bemutatott mdszerek segítségével megvizsgálom a prefrontális területek kérgi infromációfeldolgozásban játszott szerepét.

A disszertáció második része a kéreg funkcionális alhálózatait vizsgálja: egy öt területből álló modellt javaslok, amely az asszociatív tanulás megvalósításában vesz részt, majd ezt mérési adatokon ellenőrzöm statisztikai modellinverzió és -összehasonlítás segítségével. Megvizsgálom az összeköttetések skizofrénia esetén tapasztalható elváltozását a kognitív kontroll lehetséges hiányosságainak tesztelése céljából. Emellett tanulmányozom a prefrontális neurális hálózatok megváltozott belső kapcsolódását, amely szintén hozzájárulhat memóriaproblémák kialakulásához.

Nyilatkozat

Alulírott Bányai Mihály kijelentem, hogy ezt a doktori értekezést magam készítettem és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalomban, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Budapest, 2013. április 25.

.....
Bányai Mihály

Acknowledgments

I would like to thank my advisor, Fülöp Bazsó for supervising my work, introducing me the methodology of the scientific disciplines necessary for my research, especially the mathematical and physical bases of neural modelling, and also for helping to choose the directions of my study and summarising and publishing the results.

I would like to thank Péter Érdi for creating and leading the group I was able to work in, and also providing me the opportunity to gain overseas experience and for collaborations, and for giving advice in all aspects of scientific research.

My thanks go to my co-authors, László Négyessy and Vaibhav Diwadkar, who made my theoretical studies possible by their experimental work, and shared their biological knowledge about the studied phenomena with me.

I recieved invaluable help from the Computational Neuroscience and Complex Systems Group of the Wigner Research Centre for Physics, Tamás Nepusz, Zoltán Somogyvári, László Zalányi, Balázs Ujfalussy, Tamás Kiss and all past and present members of the group, who created a uniqe athmosphere for research, where I always could find someone to talk to about neuroscience, mathematics or anything else.

I thank the Department of Theory of the Wigner RCP and the Intelligent Systems Group at the Department of Measurement and Information Systems of BME, especially my co-advisor, György Strausz, who accomodated my research topic that lies further from their area of research.

Finally I would like to thank my family and Abigél for supporting me throughout my graduate studies and the writing of this dissertation.

Köszönetnyilvánítás

Szeretném megköszönni témavezetőmnek, Bazsó Fülöpnek, hogy felügyelte munkámat, bevezetett a kutatásaimhoz szükséges tudományterületek módszertanába, különös tekintettel az idegrendszeri modellezés matematikai és fizikai alapjaira, valamint segített a kutatási irányok megválasztásában, az eredmények összegzésében és publikálásában.

Köszönöm Érdi Péternek, hogy létrehozta és vezette a csoportot, amiben dolgozhattam, emellett lehetőséget biztosított számomra a külföldi tapasztalatszerzésre és együttműködések létrehozására, és ellátott tanácsaival a kutatás minden aspektusára vonatkozóan.

Köszönet illeti szerzőtársaimat, Négyessy Lászlót és Vaibhav Diwadkart, akik kísérleti munkájukkal lehetővé tették elméleti vizsgálataim lefolytatását, és megosztották biológiai tudásukat a vizsgált jelenségek kapcsán.

Felbecsületetlen segítséget jelentett a Wigner Fizikai Kutatóközpont Elméleti Idegtudomány és Komplex Rendszerek Csoportja, Nepusz Tamás, Somogyvári Zoltán, Zalányi László, Ujfalussy Balázs, Kiss Tamas és a csoport minden volt és jelenlegi tagja azzal, hogy egyedülálló kutatási atmoszférát hoztak létre, ahol mindig akadt valaki, akivel konzultálni tudtam idegtudományi, matematikai és egyéb kérdésekről.

Köszönöm a Wigner FK Elméleti Osztályának és a BME Méréstechnika és Információs Rendszerek Tanszék Intelligens Rendszerek Csoportjának, különösen Strausz György társtémavezetőmnek, hogy befogadták az irányvonalukba kevésbé illeszkedő kutatási területemet.

Végül köszönöm családomnak és Abigélnek, hogy mindenben támogattak a kutatómunka és a disszertáció elkészítése során.

Contents

List of Figures

List of Tables

List of Algorithms

"...Esz az ember a cék reklámja. Régen festő volt, de a nagy háborúban sújosan megsebesült, amikor megszököt a fetyházból. Úty ment bele a fejibe neki a golyó, hoty az alvási agyizmát, ami centrumot képez, megsértette. [...] Ezek a centrumosok soha nem tudnak aludni, még a legnemesebb regények olvasásától se. [...] A festő is ilyen: sem aludni, sem tiszteséges lenni nem tud. Mostan fontos szereplő vigyázásból, mivel a tisztelt bentlévők a rendőrség mellőzését kéretik. És esz nem alszik, ha kint áll, őrölés köszben. Mivel ojan golyó érte centrumban asz agyát, hoty sem aludni, sem józan lenni, sem mosakodni többé nem tud. Viszont verekedési centrumot nem sérte a golyó. Esztet látam egy jobbkezeséből, de balkészel megálapítottam, hoty elájulni mék tud. Esz a centruma 1.-ső rangú... "

Rejtő Jenő (1941): *Piszkos Fred közbelép Fülig Jimmy őszinte sajnálatára*,
Albatrosz könyvek, Budapest, 1966.

"...This man is the sign of the company. He used to be a painter, but he had been gravely wounded during the Great War, when he broke out of the gaol. The bullet entered his head in such a way, that his sleeping brain muscle, that constitutes a center, was hurt. [...] These center people cannot ever sleep, even from reading the most grand novels. [...] The painter is such: he cannot sleep, nor be honest. Now he is an important man looking out, as the revered customers insist not to be contacted by the police. And this guy doesn't sleep while guarding. Since such a bullet reached the center in his brain that he cannot sleep, be sober or bath. However, no brawling center has been hurt by the bullet. This I could see from his right hook, but I determined with my left hand, that he can still pass out. This center of his is 1st class... "

Jenő Rejtő (1941): *Dirty Fred intervenes for the sincere regrets of Jimmy All-Ears*,
Albatrosz books, Budapest, 1966.

1

Introduction

This dissertation is about networks in the brain. From the multitude of phenomena that can be addressed as such, I will mostly cover networks of cortical areas. What do we know about their structure, what kind of implications can we draw from it regarding their function, and how can we explain behavioural traits based on their connectivity?

Olaf Sporns coined the term "connectome" [126], in reflection to terms like "genome", to describe all kinds of knowledge about the brain that provides answer to a question on what is connected to what and sometimes when. The maps resulting from the different approaches organise into a hierarchical system: the behaviour of the living organism, human or animal, is a resultant of the interplay between functional units, which correspond to brain areas or their still-macroscopic subsystems according to our current level of knowledge, however, keeping in mind the more detailed picture is essential when interpreting measurements, or trying to explain dysfunctions. Also, the macro-network picture is necessarily phenomenological, and for creating mechanistic, procedural ideas about brain function, the appropriate level of abstraction is quite probably lower. Thus, I begin the introduction to this work by an overview of this elaborate machinery. This will be necessary to create a framework for formulating hypotheses about the ultimate topic of interest, the emergence (or presence, nonetheless) of functionality. As the essence of scientific understanding is formal reasoning, after looking at the sources of information about structure, function and their possible relationships, I will review modelling techniques for different types of phenomena. This will finally enable the formulation of the scientific questions this dissertation wishes to answer.

1.1 Levels of organisation in the brain

As the neuron doctrine states, the morphological, developmental and functional unit of the nervous system, thus of the cortex in the brain, is the neuron (for a somewhat dissenting opinion see Freeman [54]). The membrane potential dynamics of these spatially extended cells is able to produce quick spikes of quite invariable height and width, called the action potentials. These spikes travel along the elongated axons of the cell, and are transmitted to other cells trough the synapses, chemical connections of variable efficacy.

The neuron and the synapse themselves operate as sophisticated interaction networks of proteins and other molecules. This can be regarded as the lowest level of the cortical hierarchy of networks. However, this level will only be discussed in this dissertation in an example for the usability of a general graph analysis method in Section 2.4.1, otherwise, I consider spike transmission as the basis of all information processing in the brain, sticking to the neuron doctrine on the functional level.

1.1.1 Neuronal networks

The fundamental role of the connections between individual neurons in the implementation of any brain function was proposed by Santiago Ramón y Cajal (see [156] for the original, reviewed in [36]). Current theories of cortical activity agree that neural state can be derived from neuronal firing or a statistical measure of that, and learned information is stored in the synaptic efficacies, the connection strengths between neurons. This structure was one of the central concepts of cybernetics [94], and thus exerted a great influence on machine learning and pattern recognition techniques. Non-synaptic types of communication between cells exist, but most of the communication, especially for longer distances, is realised by spike transmission, according to the current scientific opinion. Regarding other signalling pathways, the reader is referred to [142] and [143].

The most precisely known biological neural network is that of the simple worm, Caenorhabditis elegans, that has 302 neurons, and all connections between those are mapped [152], as there is no inter-specimen variability in this nervous system. Although synaptic plasticity is already present in this network, the examination of such a system does not provide substantial information about more complicated functionality, as such, especially cognitive behaviour, can only be associated with higher animals. The complete mapping of a vertebrate nervous system is not available to this day, and measurement techniques only allow the gathering of fragmentary information about the structural, and even more so about the functional architecture of

a cerebral cortex.

In the human cortex there are about 10^{11} neurons connected by about 10^{15} synapses. These numbers would call for a higher level approach themselves, but there is clearly observable structure in the neural tissue that helps to define a meaningful clustering of neurons.

1.1.2 Area networks

Cortical areas are ensembles of neurons constituting the grey matter of the cortex, and can be separated from each other based on anatomical measures, such as the distribution cell types and connectivity, but often also based on the simultaneous change in neuronal activity related to a certain function, as detailed in Section 1.2. Distant areas can be connected through axon bundles through the white matter.

The first systematic area map was provided by Brodmann in 1909 (see [26] for the original, reviewed in [135], for listing of areas mentioned in this text see Table A.4), defining 52 regions for the human and monkey cortex. The area numbering introduced by him is still used in some cases. A more detailed division of the cortex was given by Economo and Koskinas in 1925 (see [40] for original, reviewed in [137]), which consisted of 103 areas. This map is depicted in Figure 1.1.

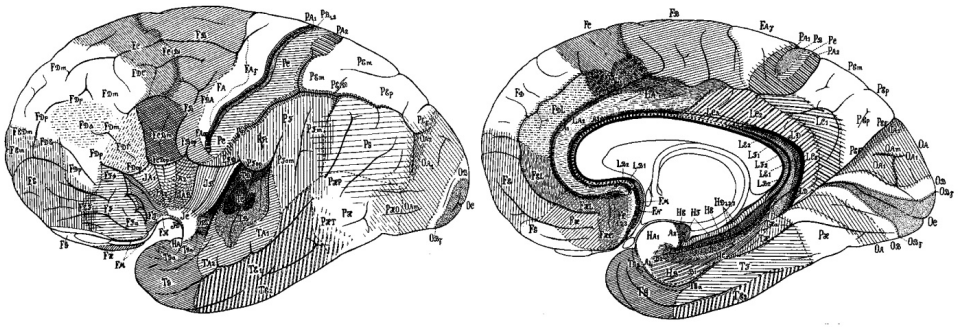


Figure 1.1: Division of the human cerebral cortex into 103 areas by Economo and Koskinas, from [40].

Nowadays scholars use a cortical area system which is more refined by improved data collection techniques and the integration of functional measurements, but there is no standard division of the cortex to a fixed number of areas.

The microscopic method of neural connectivity mapping is axonal tracing. Certain chemicals are injected into the cells of the neural tissue, and as they

are able to move along the axon in an anterograde or retrograde fashion, the location of the cells connected to the marked neuron can be determined using microscopic measurements. For a review of axonal tracing techniques, see [83]. The drawbacks of such methods are that they are highly invasive, and require many lengthy measurements to determine connections.

As for some mapping techniques non-human brains are much more accessible, we often have more detailed information about the neural circuitry of certain animal species. There has been multiple initiatives to facilitate the efficient integration and sharing of such anatomical data using online databases. One of such sites is the Brain Architecture Management System (BAMS)[1] database [25].

The species with the closest cortical architecture to humans, thus most relevant to human neuroscience are primates. Their cortical organisation is not identical to that of the homo sapiens, but several functions and structures are homologous, allowing us to draw implications from one species to the other. A connectivity map of the macaque visuo-tactile cortex was created in 1991 [49], and refined continuously since then. The data accumulated so far is accessible from the CoCoMac database[2], which also accepts submissions [130]. A new version of this network is depicted in Figure 1.2.

A non-invasive way to gather information about the brain is through magnetic resonance (MR) imaging. This technique detects the spatial distribution of certain materials using strong magnetic fields. When tuned to measure the concentration of water, with an appropriate decoding method, it can be used to trace axonal fibres in the white matter. This analysis pipeline is called diffusion tensor imaging (DTI) [77], or with a slightly different methodology, diffusion spectrum imaging (DSI) [69, 148], and can provide invaluable data about structural connectivity between brain regions. However, DTI and DSI cannot reveal the direction of synaptic connections in the axonal fibres, and can only reveal bundles of myelinated axons, fine details of connectivity, including local or close connections cannot be seen through them. Thus, the microscopic axonal tracing methods remain necessary to discover the cortical connectome.

1.1.3 Intermediate levels of hierarchy

In order to fully understand cortical computation, it is necessary to define the intermediate organisational levels between the elementary neuron and the cortical area responsible for general aspects of behaviour. The main

[1] http://brancusi.usc.edu/bkms/
[2] http://cocomac.org

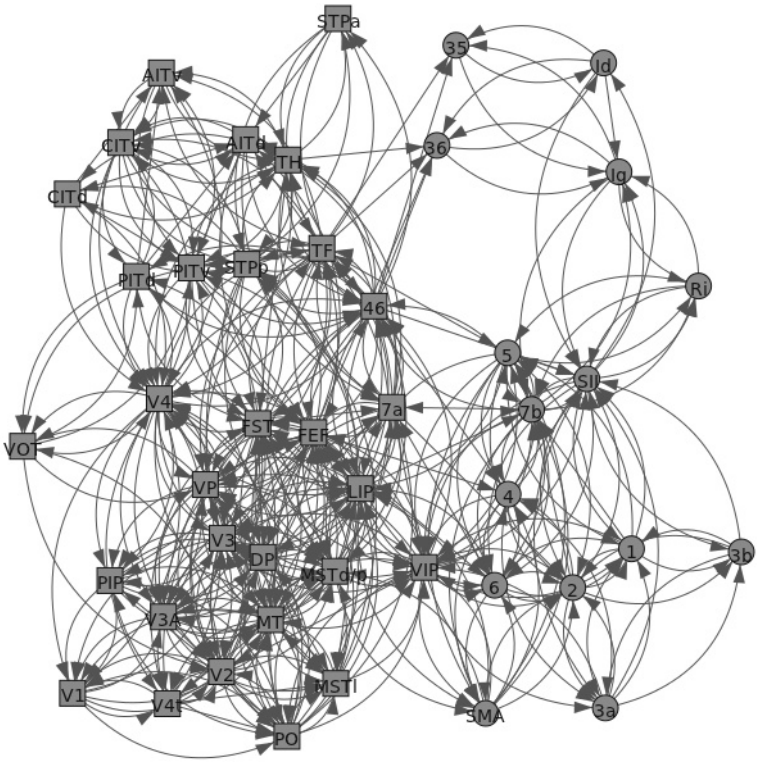


Figure 1.2: Visuo-tactile area subnetwork of the macaque cortex. Areas belonging to the tactile subsystem are denoted by circles, areas of the visual system and higher-level integration with squares. For abbreviated and numbered areas see Tables A.3 and A.4 respectively.

candidates for such a role are the cortical columns, that can be defined in several ways.

Micro-columns (or minicolumns) originate from the same progenitor cell during ontogenetic development, and are thought to encode the same feature according to the modality of the area they are located in. Such columns consist of around 80 pyramidal neurons. Cortical macro-columns (or hypercolumns) are thought to encode all features related to a certain receptive field, consisting 4-8000 pyramidal neurons. These are expressed more in certain areas than others, as anatomically in the barrel cortex of the rat, or functionally in the ocular dominance columns of the mammalian visual cortex. For a review of the topic from Mountcastle, see [100], for a more recent discussion, see [117].

A significant field of interest is the possible definition of a computational unit of the cortex, named the canonical microcircuit. This is supposed to be

replicated in most of the cortical areas carrying out computations using the same basic rules. Although consensus has not been reached on this matter, there have been attempts to give a joint anatomical-mathematical candidate for such units [18].

1.1.4 Bridging the levels

The natural goal of neuroscience is to integrate all levels of the hierarchy mentioned above into a model that can give a mechanistic explanation of brain functions. Although a widely accepted framework is still lacking this day, many attempts were made to incorporate certain levels into a model on different grounds. One direction is to define networks of neuron models (see Section 1.3.4) large enough to test hypotheses on the cortical area level [78].

Another direction is to use graph theoretical tools to analyse connectivity(see Section 1.3.1). The attempt to treat the brain as a large hierarchical network appears in recent reviews about multi-level graph theoretical analysis [29, 87], which also tries to integrate structural networks with information about functional connectivity. Such integration is inevitable, as the multiscale organisation of the neural tissue is ultimately subsumed, and shaped by the functionality it implements. Thus, I continue this introduction with the examination of possible data sources and concepts binding structure and function together.

1.2 Structure and function

We can often associate the implementation of one or more types of tasks, functionalities to cortical areas. These include primary and associative processing of sensory inputs, motor control, working and associative memory, processing and production of language and cognitive functions, such as control and decision making.

The analysis of the cortical circuitry from the triple viewpoint of structure, function and dynamics was given by János Szentágothai and co-authors in [8].

Quantitative information about functional connectivity in the brain can be recorded using various techniques. Electroencephalography (EEG) and magnetoencephalography (MEG) provide good temporal but poorer spatial resolution, while positron emission tomography (PET) and more recently functional magnetic resonance imaging (fMRI) record with a good spatial resolution from the whole brain at the expense of temporal precision.

The magnetic resonance imaging technique can be used to detect temporal changes in the cortex by tuning the apparatus to measure the distribution of deoxyhemoglobin. At the current technical level, fMRI measurements can record a three dimensional image of the whole brain in every 1-2 seconds with the spatial resolution of 2-3 millimetres, associating one intensity value of the blood oxygen level dependent (BOLD) signal to each such cube, called voxel. Oxygen concentration in blood vessels can be linked to the spiking activity of neighbouring neural tissue, as the energy consumption of neurons is dominated by the process of restoring equilibrium membrane potential after an action potential. However, the exact mapping between neuronal dynamics and BOLD remains to be determined. Comparative parallel measurements and mathematical models allow the definition of mapping functions of different flexibility that can be used to infer neural activity [88].

Using fMRI measurements, we can determine which cortical areas participate in the implementation of a certain function [30]. A comprehensive review about functional subsystems in the brain and the analysis of magnetic resonance recordings can be found in [51].

1.2.1 Functional connectivity

When we talk about functional connectivity in the cortex, the first two questions are the following: what are the physical correlates of the vertices between which we infer the edges, and under what external, sensory circumstances do we conduct the measurement. For the first question, the two typical answers are voxels and cortical areas, and this dissertation focuses on the latter. We can assign a single time series to each area we are interested in by merging the intensities of the voxels related to the given area by averaging, taking maximum variance or the first principal component. To determine which voxels belong to an anatomical area, probabilistic maps of standardised brain coordinates can be used (the two major coordinate systems used are MNI[3] and Talairach[4]), and inter-subject variance in anatomical layouts can be accounted for by considering simultaneous, sensory input-related activation of voxels.

For the second question, we have to answer by specifying an experimental paradigm, which is selective to the function(s) of which we want to determine the activated subnetworks. As the co-activation of cortical areas varies heavily based on the type of information processing going on in the brain, we can only talk about task-dependent functional subnetworks, instead of

[3]http://www.loni.ucla.edu/ICBM/Downloads/Downloads_Atlases.shtml
[4]http://www.talairach.org/

a functional network in general. Such subnetworks can be organised into a hierarchy as well based on the generality or specificity of the tasks they are associated with. Detailed characterisation of such subnetworks related to many different tasks is given in [114].

A special, heavily investigated case of the functional subnetworks is when we do not apply any sensory input in the experimental paradigm. The brain remains quite active in this situation also, and the set of activated areas is called the resting state network. For details of functional resting state connectivity, see [138].

Psychiatric diseases can often be related to dysfunctions of the cortical connectome [28]. Alterations of functional connectivity can be either disconnection of two or more areas, or over-connectivity, over-synchronisation of the network elements. Although the investigation of such quantities does not reveal the "cause" of illnesses, it can provide substantial information about the mechanism of the disorder, allowing implications about healthy cortical functions, and possibly directions for new therapies.

We can consider function on many different levels from elementary sensory processing to the highest levels of cognition involving decision making and strategic planning. As the latter type of functionality is the closest to human (and animal) behaviour and poses the greatest challenge to achieve any understanding about its mechanisms, I dedicate a more detailed discussion to the part of the cortex such functions are thought to be located in: the prefrontal cortex.

1.2.2 Role of the prefrontal cortex

As Elemér Lábos stated about fifteen years ago, "the prefrontal cortex is the hippocampus of the future". This refers to the fact that the hippocampus, with its well known and specific structural connectivity and its involvement in cognitive processes like spatial navigation or associative memory formation, has been in the focus of interest of neuroscientists both from the experimental and theoretical points of view. Lately, the prefrontal cortex proved to be a similar contender for the interest of the field [104]. This region is the main centre implementing cognitive control, as it integrates information from most parts of the cortex that are involved in higher-level, associative processing and might as well prove be the top of the cortical hierarchy.

The prefrontal cortex is a large part of the frontal lobe, consisting of multiple regions and several cortical areas. Using the Brodmann notation, areas 9, 10, 11, 45, 46 and sometimes also 8, 12, 13, 14 and 47 can be considered as prefrontal. As area 46 in the dorsolateral region is highly connected to visual and multimodal associative areas, I will accentuate its

role in the modelling approaches of this work.

Miller and Cohen presented a comprehensive treatment of the structure and functionality implemented in the prefrontal cortex [95]. They argue that representations of situation-independent knowledge is located there, thus control of intentional, cognitive behaviour in the sense of providing top-down signals towards primary and associative sensory areas to filter information and drive motion to complete a task. Such behaviour is expressed in the models of Section 3.2.3. In order to make informed decisions, filtered sensory information must converge in the prefrontal cortex, noted by Miller and Cohen. The structural basis of this phenomena is studied in Section 2.4.1.

The prefrontal cortex plays a significant role in psychiatric disorders too. Several decades ago even its complete removal was practised as treatment [53], but contemporary theories of cognitive disorders, like schizophrenia (as discussed in Section 3.1.1), also stress the significance of prefrontal alterations in the mechanism of deficits in patients [67]. These alterations usually affect the flow of control signals from higher level, prefrontal areas towards lower levels, where association of sensory data to other modalities and existing knowledge is implemented, such generating symptoms suggesting an altered sensation of reality, context and relationships.

To gain understanding about the interplay between cortical structure and function, we need a mathematical toolset to formulate and test hypotheses with measurement data. Such frameworks will be discussed in the next section.

1.3 Modelling approaches

Mathematical modelling is the core activity of scientific theory formulation. Its goal is always to predict the behaviour of a physical system, and it proved to be generally quite successful in it, Eugene Wigner even contemplated on whether the success rate could be explained [153]. In the physical sciences, the models are often descriptive, and on a level as low as possible.

As we know extremely little about the human brain given its physical complexity (and without a radical improvement in measurement techniques, this will remain the case in the near future), the role of models at high levels of abstraction is greater that in many other disciplines. Additionally, the evolutionary pressure on living organisms to efficiently maintain specimen and species, results in a functional specialisation unobserved in physical sciences. These circumstances sometimes necessitate the application of a normative approach, that is, to begin the modelling with the information about what the system is for [17].

1.3.1 Graph theoretical analysis

The most straightforward mathematical framework to handle network models is graph theory, in which the system at hand is modelled as a set of vertices and directed edges connecting pairs of them: $G(V, E)$ (undirected graph models are also often useful, but will not be covered in this dissertation). The in- and out-degrees of the vertex i are denoted by d_i^{in} and d_i^{out} respectively. Several measures have been developed to characterise the possible roles of vertices and edges in implementing function, and also to classify, characterise and compare whole networks.

Vertex-based measures try to capture the relative importance of certain vertices in the information (or other quantitative) flow, such as vertex centrality [52], or the identification of hubs in the network [127].

Edge-based measures try to capture the role individual edges play in the information processing, such as the generalisation of betweenness centrality to edges [66].

To describe a whole network in a concise manner, the most often used property is the degree distribution, which proved to be a power law for some phenomena of interest [15]. For a comparison of other graph metrics, see [21, 145].

Clusterings, grouping of the vertices of the graph according to some criteria, most commonly connectivity, can reveal interesting macro-structural information about the network. For an overview of clustering algorithms see [82].

1.3.2 Random graph models

Random graphs greatly contribute to the understanding of real-world networks by describing the formation mechanisms of networks as sampling statistical generative models. They reproduce statistical properties of the real networks, such as degree distribution [108].

The simplest random graph model is the Erdős-Rényi graph[46], which is only defined by the number of vertices and the probability of having an edge between any two of them. The procedure that generates such graphs is given in Algorithm 1. Such networks can serve as bases of comparisons, aiming to detect meaningful structure in real-world or other model networks, as Erdős-Rényi graphs are meant to be structureless.

Watts and Strogatz introduced the notion of small-world networks [147], which can be characterised by short average path length and high transitivity. Networks with such properties are present in many systems, as high transitivity ensures the presence of densely connected clusters, and short path

Algorithm 1 The Erdős-Rényi graph generation algorithm

Require: $n \geq 0$, $0 \leq p \leq 1$
1: $G := G(V, E)$, $|V| = n$, $|E| = 0$
2: **for** $i = 1$ to n **do**
3: **for** $j = 1$ to n **do**
4: $r :=$ random real number between 0 and p
5: **if** $r \leq p$ **then**
6: $E := E \cup \{i, j\}$
7: **end if**
8: **end for**
9: **end for**
10: **return** G

length ensures that there are edges connecting such clusters. A procedure that generates graphs with small-world property is given in Algorithm 2 for the special case when the starting state is a directed circle, as in case we can start the algorithm from a k-connected regular circle. However, the notion of small-worldness is quite general, very many networks fit in it, and such models may serve better as a more structured null hypothesis than an Erdős-Rényi graph in the case of a network with apparent community structure.

Algorithm 2 The Watts-Strogatz graph generation algorithm

Require: $n \geq 0$, $0 \leq r \leq n$
1: $G := G(V, E)$, $|V| = |E| = n$, $d_i^{in} = d_i^{out} = 1 \qquad \forall i \in V$
2: **for** $i = 1$ to r **do**
3: s,$q :=$ random integers between 1 and n, inclusive
4: $t :=$ target index of edge with source index s
5: $E := (E \cup \{s, t\}) \setminus \{s, q\}$
6: **end for**
7: **return** G

As a model for growing networks, scale-free graphs were introduced by Barabási and Albert [15]. These give a simple mathematical formulation of the "rich gets richer" principle, resulting in power-law distributions over vertex degrees. The procedure that generates such networks is called preferential attachment and is given in Algorithm 3.

Algorithm 3 The preferential attachment graph generation algorithm

Require: $n \geq 3$
1: $G := G(V, E)$, $|V| = 2$, $|E| = 1$
2: **for** $i = 3$ to n **do**
3: $\quad p_i := \frac{d_i^{in}}{\sum_{j=1}^{i-1} d_j^{in}}$
4: $\quad r :=$ random real number between 0 and 1
5: $\quad t : \sum_{j=1}^{t-1} p_j \leq r \leq \sum_{j=1}^{t} p_j$
6: $\quad v := V \cup i$
7: $\quad E := E \cup \{i, t\}$
8: **end for**
9: **return** G

1.3.3 Statistical generative models

As functional activity in vertices of the cortical area network is measured in an indirect manner by fMRI, EEG, MEG, etc., network discovery in such settings is about inferring unobserved quantities of neural activity, and dealing with a considerable amount of uncertainty about those and causal links between them. For such problems, statistical generative models offer a consistent mathematical framework by the definition of a complete model of the data generation mechanism and the source and form of uncertainty on the quantities in it.

The inference of probability distributions of hidden variables (also called latent or unobserved) of a generative model by considering a set of measurements as values of an observed variable, is called the *inversion* of the model. Hidden variables are usually divided to state variables and parameters, differing in interpretation and characteristic time scales. One can be interested in any of the two or both. Inference methods are based on analytical or numerical methods of calculating conditional probability distributions, and are discussed in Section 3.2.4. The main advantage of statistical generative models is the possibility to make a consistent probabilistic comparison between different model structures, as outlined in Section 3.2.5.

A subset of statistical models include temporal dynamics, consisting of deterministic and stochastic components. The structure of such a model is outlined in Figure 1.3. The observed variables are split into inputs u and outputs y. Latent variables include state variables x, state noise ε^x, observation noise ε^y, parameters governing the deterministic temporal evolution of the state θ^{xd} and the output θ^{yd}, and probability distributions of the noise variables θ^{xs}, θ^{ys}. The parameters are random variables themselves, and their distributions are defined by hyperparameters h. We can gener-

ate synthetic samples of the observed output, or infer the hidden state or (hyper)parameters, which is called learning. Inference and learning involves the estimation of conditional distributions during model inversion, which can be a hard problem in a general case, but can be greatly simplified by making linear assumptions (or applying local linearisations) about the temporal dynamics, and using Gaussians for the probability distributions.

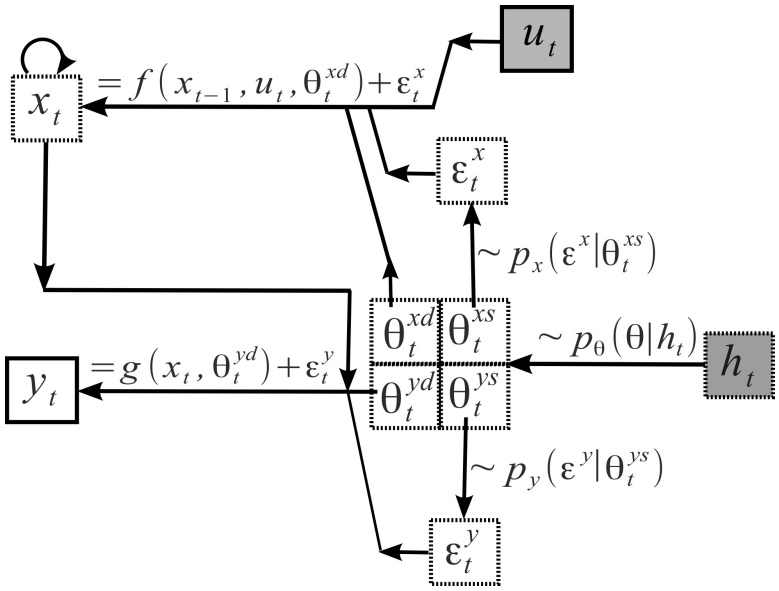


Figure 1.3: General architecture and sampling procedure of a dynamical statistical model. Arrows denote causal interactions between variables, with formulae given ($\sim$ meaning a sample from a distribution, $\theta = \{\theta^{xd}, \theta^{xs}, \theta^{yd}, \theta^{ys}\}$). Observed variables are depicted in solid line boxes, latent ones in dashed line boxes. Shaded boxes show deterministic variables (marginals are always deltas at the actual value). Assumptions of the model are the formulae for f, g, p_x, p_y, p_θ and initial values for x_0 and h_0.

Network connectivity can be inferred in statistical models as interaction structure between hidden state variables. In the literature, cortical networks discovered by data-driven statistical methods, such as correlation maps, are sometimes called "functional connectivity", as opposed to ones resulting from model-based methods, named "effective connectivity" [57, 58]. However, this nomenclature will not be adopted in this dissertation, where by "functional" I always mean something related to function, as opposed to structure.

1.3.4 Simplified neural network models

After Hodgkin and Huxley published their groundbreaking model of the electrical dynamics of neural membranes based on measurements in the giant axon of the squid [74], several simplified models came into existence relying on abstractions of different features of the original model's spiking behaviour. The type of models that emphasise the importance of the exact time when spikes are observed (thus promoting the use of a temporal code in the nervous system [140], instead of, for example, a rate or phase code [73]) are called integrate-and-fire neurons. The idea goes back as far as to Lapicque (see [84] for original, reviewed in [1]), but developed into a mathematically practical tool much later [75]. The time dynamics of sub-threshold membrane potential is described by 1.1, where $I(t)$ incorporates all appropriately scaled synaptic and other inputs to the cell.

$$\tau \frac{dV}{dt} = -V(t) + RI(t) \tag{1.1}$$

When V exceeds a certain threshold, a spike is recorded, and V is set back to a refractory value, and after the refractory period, a base value. These models capture the most characteristic property of neurons, namely the spiking behaviour, while remaining simple enough for the dynamic trajectories to be computed by simple integration schemes. When building networks of such neurons applying a synapse model, the computational costs are low enough to handle a larger circuitry with regularly available computers, and simulate longer time intervals of spiking activity.

1.4 Further reading

Additional details of the topics mentioned here can be found in a recent book about networks in the brain [125]. About random graph models, the reader is referred to a paper from Barabási and Albert [2]. A review of several network analysis methods for fMRI data, mostly ones complementary to the approach used here, can be found in [122]. The general problem of statistical modelling together with the most widely used techniques is covered in great detail by Bishop [22], and applied to model brain function in [116]. One of the standard guides to neural network modelling is the handbook of Koch and Segev [80].

The most important source of data about brain networks in the near future will be the Human Connectome Project of the National Institute of Health of the USA. Within the project, the MGH/Harvard-UCLA consor-

tium [5] works on the refinement of structural data collection methods, and the WU-Minn consortium[6] collects structural and functional datasets from a large number of subjects providing an invaluable freely accessible database of multimodal data [139].

For an overview of the topics of this dissertation, targeted to the general audience in Hungarian, see [10].

1.5 Questions to be answered

The main goal of this dissertation is to provide answers to the following scientific questions: (i) What implications can be drawn from the sole structure of the cortical macro-network regarding information processing functionality, what kind of procedure can be applied to generate a similar structure, and how can the role of prefrontal areas in the cortical signal flow be assessed by structural measures? (ii) What functional subnetwork can be identified based on a physiological measurement during task completion and how is this network altered by a psychiatric disease, and what is the possible role of prefrontal areas is this system, and what kind of dysfunction of those may lie in the background of the illness? According to the two question groups, the majority of the dissertation is organised into two chapters: Chapter 2 investigates structure and function based on anatomical data and graph theoretical models, while Chapter 3 uses statistical models to analyse functional physiological data.

[5] http://www.humanconnectomeproject.org/
[6] http://www.humanconnectome.org/

2

Signal flow in directed networks

The network of cortical areas can be represented as a directed graph. The edges are sets of synapses, and information is transmitted between the vertices as arrays of spike trains. In this chapter, I will present a graph theoretical approach to the question of interaction between network structure and function. Especially, the convergence or divergence of information on certain vertices and edges will be investigated.

As the methods presented here do not only apply to cortical are network, but have a broader significance regarding directed networks in general, I formulate all statements in the most general sense, and apply them to many different networks as well.

2.1 Overview

My goal is to identify functional properties of nodes based on the network structure. Many attempts were made to find functional signatures in the network structure, such as [76, 128], for a review see [107]. As tagging network nodes and edges with functional attributes depends on external information

---

Related publications:

Bányai M., Nepusz T., Négyessy L., Bazsó F.: *Convergence properties of some random networks.* 7-th International Symposium on Intelligent Systems and Informatics (SISY), 2009, Subotica, Serbia. 241-245, 2009

Bányai M., Négyessy L., Bazsó F.: *Organisation of signal flow in directed networks.* Journal of Statistical Mechanics: theory and experiment P06001, 2011.

Négyessy L., **Bányai M.**, Bazsó F.: *What makes the prefrontal cortex so appealing in the era of brain imaging? A network analytical perspective.* Acta Biologica Hungarica, **63**(Suppl1):38-53, 2011.

and is not a completely unique procedure, the original problem needs reformulation which is tractable with graph-theoretical tools.

The function real-world networks perform constrains their structure. Yet, one often has more detailed information about the network structure than about the functions it may perform. I focus on systems, either natural or artificial, which process signals and are comprised of many interconnected elements. From a signal processing point of view, global information about network structure is encoded in the shortest paths, i.e. if signal processing is assumed to be fast, most of network communication is propagated along the shortest paths. Therefore global and local properties of shortest paths are relevant for understanding organisation of the signal processing in the system represented with a suitable network. During signal transmission, signals are being spread and condensed in the nodes, as well as along network edges. It was previously shown [102, 103] that in case of cerebral cortex, using a simplified version of the convergence degree (CD), it was possible to connect structural and functional features of the network. In complex networks, signal processing characteristics are also determined by the level of network circularity (which in biology and especially neural science is known as reverberation, for obvious reasons). Possibility to go around *chordless* circles necessitates simultaneous quantification of signal condensing, spreading along network edges and edge circularity. Here I generalise edge convergence and divergence [103], and take into account the existence of circles in the network, treating their effects separately from the effect of branching. For that reason I refine the definition of edge convergence and introduce the overlapping set of an edge, both notions are to be defined in a precise manner later in the text. My approach may be viewed as generalisation of in-, out and strongly connected components of a graph to the level of network edges. Notions introduced have an extra gain, they help clarifying the otherwise murky notion of network causality. The functional role of a node in a network is defined by the amount of information it injects to or absorbs from the system, or passes on to other nodes. In case of real-world networks I test my findings using external validation, given the existing body of knowledge about each specific network. I illustrate the advantage of edge-based approach with the case of strongly connected graphs, where edge-based measures offer deeper understanding of signal processing and transmitting roles of nodes than an analysis which concentrates solely on nodes and their properties.

Measures I work with are applicable to networks of all sizes, there is no assumption about "sufficient" network size. More precisely, networks I work with can be small, and applicability to large networks is limited only by the computational capacity needed to find all shortest paths in the network. The semantics of my approach is tailored to explain signal flow, though

my methodology is applicable to directed networks in general. In cases of information processing, regulatory, transportation or any other network the appropriate semantics of the approach has to be given.

In Section 2.2 I introduce the notions of convergence degree and overlapping set, in Section 2.3 I define the flow representation, in Section 2.4 I analyse four real-world networks and discuss signal transmission, processing and control properties of the small-world networks. I compute CD-s and (nontrivial) overlap probability distributions for three model networks. In the last section I discuss my results and draw conclusions.

2.2 In-, out and overlapping-sets and the convergence degree

Convergence degree was introduced in [103] for the analysis of cortical networks. For application to some random networks see Section 2.4.2. I modify the measure introduced therein, in order to capture the structure of shortest paths in a more detailed way. I will discuss both global and local properties of the shortest paths, relevant notions will be distinguished with self explanatory indices G and L respectively.

Let $SP(G)$ be the set of all the shortest paths in the graph G. For any edge $e_{i,j} \in E(G)$ we can choose a subset $SP(G, e_{i,j})$ comprised of all the shortest paths which contain the chosen edge $e_{i,j}$. $SP(G, e_{i,j})$ uniquely determine two further sets: $In_G(i, j)$ the set of all the nodes from which the shortest paths in $SP(G, e_{i,j})$ originate, and $Out_G(i, j)$ the set of all the nodes in which the shortest paths in $SP(G, e_{i,j})$ terminate. By using the G index I indicate that these sets are constructed from the global structure of the graph through the shortest paths, in contrast with the local definition introduced later. By definition I assume that node i is in $In_G(i, j)$ and node j is in $Out_G(i, j)$. A third set can be defined as $Int(i, j) = In(i, j) \cap Out(i, j)$, the intersection of In- and Out sets and call it the overlapping set. Note that $In_G(i, j)$ ($Out_G(i, j)$, respectively $Int_G(i, j)$) is the edge-level equivalent of the in-component (out-component, respectively strongly connected component) of the directed network, introduced in [108] and later refined by [39]. Notions relevant for understanding the convergence degree and overlapping set are shown in Figure 2.1.

From the perspective of the chosen edge, the whole network splits to two, possibly overlapping sets, both of which have rich structure. Shortest paths induce natural stratification on the set $In_G(i, j)$, nodes at distance 1, 2 and so on from the node i are uniquely determined. Points at distance m from

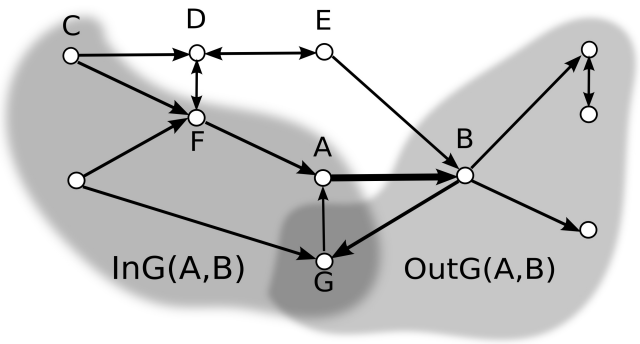


Figure 2.1: Global sets are displayed as shaded regions, local sets are comprised of first in-neighbours of node A and first out-neighbours of node B inside the shaded regions, with the exception of node G, which is contained in the local and global overlap of $In(A, B)$ and $Out(A, B)$. Note the omission of points D and E from the global input and output sets.

the tail form the m-th stratum of $In_G(i, j)$. Each point in the m-th stratum is a tail of an edge with a head in the $m-1$-th stratum. Edges connecting m-th stratum with any stratum $n < m-1$ are prohibited. Edges from the In strata to the Out strata are prohibited, since those would alter the shortest paths between the sets. The set $Out_G(i, j)$ is stratified in a similar fashion. Points in the intersection of $In_G(i, j)$ with $Out_G(i, j)$ inherit both stratifications. Stratification of In_G and Out_G sets is illustrated in Figure 2.2.

Local versions of these sets are defined as follows: $In_L(i, j)$ is the set of all the first predecessors of the node i, while $Out_L(i, j)$ is the set of first successors of the node j. When indices G or L are omitted, either is used. If the graph has circles, In and Out sets may overlap, thus it makes sense to introduce strict SIn and $SOut$ sets, which are defined as follows:

$$SIn(i, j) = In(i, j) \setminus Int(i, j) \quad (2.1)$$

$$SOut(i, j) = Out(i, j) \setminus Int(i, j) \quad (2.2)$$

In, Out, SIn and $SOut$ are generalisations of the notion of first predecessors and successors of a node, and accordingly, cardinalities of these sets are generalisations of the in- and out-degrees of nodes. Note that global and local versions of the In, Out and overlapping sets are two extremes of two set families defined as follows. Let $In(i, j, r_1)$ be the set of points from which paths at distance less or equal to r_1 from the point i begin, analogously let $Out(i, j, r_2)$ be the set of points at which paths at distance less or equal to

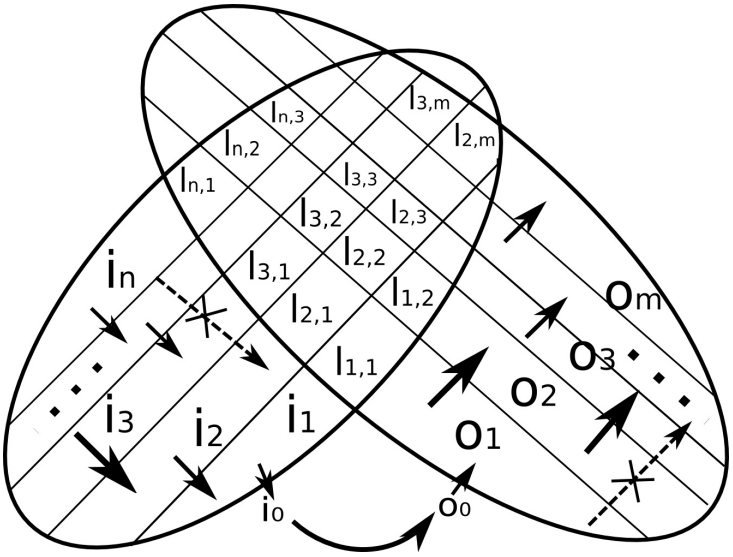


Figure 2.2: Input strata are labelled with indices i, output strata are labelled with indices o and overlap strata have double indices l. Examples of prohibited edges are shown with dashed lines, necessary edges are shown with full line. Strata i_0 and o_0 are connected with the edge itself and they do not overlap.

r_2 from the point j terminate. The two sets are balls centred at i and j with radii r_1 and r_2. Instead of balls, one may consider the surfaces of the balls, in which case points at distances r_1 and r_2 are considered. The global In-set is thus $In_G(i,j) = In(i,j,\infty,\infty)$, whilst the local In-set corresponds to points at surfaces with radii 1, $In_L(i,j) = In(i,j,1,1)$.

The notion of strict in-, out- and overlapping sets is important for understanding causality relations in network systems. Global signal flow through an edge $e_{i,j}$ induces separation of network nodes into four classes:

1. $SIn_G(i,j)$, in which are the causes of the flow.

2. $SOut_G(i,j)$, in which the effects of flow are manifested.

3. The overlap, whose elements represent neither cause nor effect. Relation between elements in the overlap is often described as circular- or network causality.

4. Points which are not members of $In_G(i,j) \cup Out_G(i,j)$ form the remaining, fourth category which has no causal relationship with the signal flowing through the given edge.

I stress that for a generic graph no such partition is possible based on node properties. E.g. if we tried to define analogous notions based on node properties, all analogue node classes would coincide for the case of strongly connected graphs. The In and Out sets would coincide, and all distinction between different node classes would have been lost.

For each edge, I define three additional measures, namely the relative size of the strict in-set ($RIn(i,j)$), the relative size of the strict out-set ($ROut(i,j)$), and the relative size of the overlap between in-set and out-set $ROvl(i,j)$, as follows:

$$RIn(i,j) = \frac{|SIn(i,j)|}{|In(i,j) \cup Out(i,j)|} \quad (2.3)$$

$$ROut(i,j) = \frac{|SOut(i,j)|}{|In(i,j) \cup Out(i,j)|} \quad (2.4)$$

$$ROvl(i,j) = \frac{|In(i,j) \cap Out(i,j)|}{|In(i,j) \cup Out(i,j)|} \quad (2.5)$$

where $|S|$ denotes the cardinality of the set S.

Note that Equation 2.5 is the Jaccard coefficient [79] of the $In(i,j)$ and $Out(i,j)$ sets. It is possible to generate networks which have edges with large global overlaps, one simply adds randomly a small number of edges to an initial oriented circle. This example helps understanding the meaning of (possibly large) global overlaps: they are characteristic of edges in chordless circles. More precisely, for and edge to have a nonempty overlapping set it is necessary, but not sufficient, to be on a chordless circle of length at least three. I illustrate this by an example. In the graph shown in Figure 2.3, the only edge with nonempty overlapping set is $e_{1,2}$, with $Int(1,2) = \{3\}$. $e_{1,2}$ is on the chordless circle (3,1,2,3), whilst the edges $e_{3,1}$ and $e_{2,3}$ on the same chordless circle have zero overlapping sets.

Local overlaps are related to the clustering coefficient of the graph, since they define the probability that the vertices in the neighbourhood of a given vertex are connected to each other.

Overlap represents global mutual relationship and a measure of dependence (in terms of chordless circles) between In- and Out sets. This dependence is inherent in the network

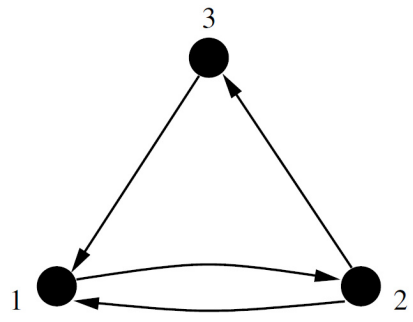


Figure 2.3: containing edges with empty and nonempty overlapping sets.

structure. Large Jaccard coefficient of the $In(i,j)$ and $Out(i,j)$ sets is not detectable with edge betweenness, as it may obtain large values for edges with non-overlapping sets.

The edge convergence degree $CD(i,j)$ of the edge $e_{i,j}$ is defined as follows:

$$CD(i,j) = RIn(i,j) - ROut(i,j) = \frac{|SIn(i,j)| - |SOut(i,j)|}{|In(i,j) \cup Out(i,j)|} \tag{2.6}$$

Note that the definition of CD uses the normalised sizes of the strict In- and Out-sets to make the measure independent of the network size. Furthermore, this formula is related to the complement of the Jaccard coefficient (denoted as $Jacc(\ ,\ )$) of the In- and Out-sets, or equivalently to their normalised set-theoretic difference, thus connecting the CD to information theoretical quantities. The following inequality is obvious:

$$|CD(i,j)| \leq 1 - Jacc(In(i,j), Out(i,j)) = 1 - ROvl(i,j) \tag{2.7}$$

Directionality of the edge gives meaning to cardinality subtraction, as In and Out sets can be distinguished. If the CD value is close to one, the signal flow through the edge is originating from many sources and terminating in very few sinks, while CD values close to -1 indicate flow formed of few sources and many sinks. This property justifies rough division of edges according to their CD properties to convergent (condensing), balanced and divergent (spreading). An oriented circle with at least three nodes has the maximum possible global overlap for each edge, while the absolute value of the global CD is the smallest possible, in accordance with the inequality (2.7). I note that CD in an oriented chain monotonously decreases along the chain, whilst the overlap is zero along the chain. This simple example again illustrates how CD and overlap are sensitive to the network topology.

Applicability of the convergence degree is limited by the following facts. Definition of convergence degree makes sense only if not all connections are reciprocal, stated otherwise if there is a definite directionality in the network. If every connection is reciprocal, the network may be considered unoriented. For fully reciprocal networks, the In and Out sets would coincide. Second, convergence degree makes sense for a network which is at least weakly connected.

2.3 Flow representation of the network

Since the number of edges exceeds the number of nodes in a typical connected network, and in many cases we are interested in the role of individual nodes, it is desirable to condense primarily edge-based measures to a node-centric view. The condensed view should reveal several features of interest: local vs global signal processing properties of network nodes, directionality of the information, i.e. whether we are interested in the properties of the incoming or outgoing edges, the third aspect is the statistics, i.e. total or average property of the edges, and finally we may choose edges according to the sign of their CD. Condensing the information about overlapping sets follows the same lines, with the exception of the sign.

I proceed by an example and introduce the following six quantities defined for each node i. Let $\sigma^{-,av}_{in,L}(i)$ denote the sum of all incoming negative local convergence degrees divided by the node's in-degree, and let $\sigma^{+,av}_{in,L}(i)$ denote the sum of all incoming positive convergence degrees divided by the node's in-degree, i.e. $\sigma^{-,av}_{in,L}(i)$ is the average negative inwards pointing local CD of the node i.

In a similar way we can also define $\sigma^{-,av}_{out,L}(i)$ and $\sigma^{+,av}_{out,L}(i)$ for outgoing convergence degrees. For clarity I give formulae for $\sigma^{-,av}_{in,L}(i)$ and $\sigma^{-,av}_{out,L}(i)$. $d_{in}(i)$ and $d_{out}(i)$ denote in-degree and out-degree of the node i, θ is the unit step function continuous from the left. $\Gamma_{in}(i)$ denotes the first in-neighbours of the node i, the analogous notation $\Gamma_{out}(i)$ is self-explanatory.

$$\sigma^{-,av}_{in,L}(i) = \frac{1}{d_{in}(i)} \sum_{j \in \Gamma_{in}(i)} \theta(-CD_L(j,i)) CD_L(j,i) \tag{2.8}$$

$$\sigma^{-,av}_{out,L}(i) = \frac{1}{d_{out}(i)} \sum_{j \in \Gamma_{out}(i)} \theta(-CD_L(i,j)) CD_L(i,j) \tag{2.9}$$

I also define $\sigma^{ovl,av}_{in,L}(i)$, the sum of all incoming local overlaps and $\sigma^{ovl,av}_{out,L}(i)$, the sum of all outgoing local overlaps each being normalised with the corresponding node degree.

$$\sigma^{ovl,av}_{in,L}(i) = \frac{1}{d_{in}(i)} \sum_{j \in \Gamma_{in}(i)} ROvl_L(j,i) \tag{2.10}$$

$$\sigma^{ovl,av}_{out,L}(i) = \frac{1}{d_{out}(i)} \sum_{j \in \Gamma_{out}(i)} ROvl_L(i,j) \tag{2.11}$$

Factors before the sums serve normalisation purposes, each σ should have a value within the $[-1, 1]$ interval. These quantities are average local CD-s and relative overlaps corresponding to each node. One is also interested in the total of the in- and out pointing edges of a given CD sign, and define the corresponding version of the node-reduced convergence degree. For normalisation purposes the sums in σ^{tot}'s are divided by $n-1$, the maximal possible number of the outgoing (incoming) connections a node can have, where n denotes the number nodes in the network.

Thus, using the quantities $\sigma^{\{+,-\},\{tot,av\}}_{\{in,out\},\{G,L\}}$ and $\sigma^{ovl,\{tot,av\}}_{\{in,out\},\{G,L\}}$ one can construct four different CD flow representations of a network, namely CD^{tot}_G, CD^{av}_G, CD^{tot}_L and CD^{av}_L.

The incoming node-reduced CD values are understood as coordinates of the x axis, while the outgoing CD values are interpreted as the coordinates of the y axis. In order to display overlaps together with the convergence degrees in a single figure, overlaps are treated as the coordinates of the z axis, the incoming overlaps being positive and the outgoing understood negative. Each point is represented in each octant of the flow representation. The points in the xy plane are not independent, given the values in the diagonal quadrants, the other two quadrants can be reconstructed with reflections.

Representation of graph nodes in the xy plane is related to the CD flow through the nodes in the following way. The CD flow ϕ through the node i is defined as follows:

$$\phi(i) = \sum_{j=1}^{d_{out}(i)} CD(i,j) - \sum_{j=1}^{d_{in}(i)} CD(j,i) \tag{2.12}$$

The first sum is equal to $\rho_{out}(i)\left(\sigma^{+}_{out} + \sigma^{-}_{out}\right)$, where $\rho(i)$ is the appropriate weight, whilst the second sum equals $\rho_{in}(i)\left(\sigma^{+}_{in} + \sigma^{-}_{in}\right)$. The flow can be rewritten as

$$\phi(i) = \rho_{out}(i)\sigma^{+}_{out}(i) - \rho_{in}(i)\sigma^{-}_{in}(i) + \rho_{out}(i)\sigma^{-}_{out}(i) - \rho_{in}(i)\sigma^{+}_{in}(i) \tag{2.13}$$

If the first difference on the right hand side of Equation (2.13) is large (small), i.e. the representative point is close to the diagonal $y = -x$ and is far from the origin in the top left (bottom right) quadrant, and the second difference is small (large), i.e. the representative point is close to the diagonal $y = -x$ and is far from the origin in the bottom right (top left) quadrant, the node i is *source* (*sink*) of the CD flow. Analogously, the CD flow can be written as:

$$\phi(i) = \rho_{out}(i)\sigma^{-}_{out}(i) - \rho_{in}(i)\sigma^{-}_{in}(i) + \rho_{out}(i)\sigma^{+}_{out}(i) - \rho_{in}(i)\sigma^{+}_{in}(i) \qquad (2.14)$$

where the two differences determine the router characteristics of the node i. In this sense flow representation is a means to independently study different components of the CD flow. Different circles may have common nodes, thus the overlap flow defines whether different circles passing through the given node have more common parts after of before the given node, i.e. whether a node is a source or sink of circularity. Precise meaning of large and small depends on the criteria used to classify the representative points of the flow representation.

Nodes can be classified based on the CD (relative overlap) flow, besides distinction based on the sign, the scale is continuous, there is no a-priori grouping of nodes. Further classification can be made based on the structure of the CD (relative overlap) flow, i.e. based on properties of different terms defining the CD (relative overlap) flow. Components of the flow representation for two toy graphs are shown in Figure 2.4. We can observe that same nodes may be global, but not local CD flow sinks or sources.

Each octant represents different aspect of convergence-divergence relations in the network. These quantities bring us to the actual interpretation of edge convergence and divergence as a characterisation of signal flow on the nodes of a network. To make statements about the signal flow derived from the CD flow, we have to make an inversion of properties, as nodes which behave as a sink of convergence, actually inject information to the network, thus they are sources of signal. Respectively, CD sources are sinks of signal. Assuming this interpretation we can extract useful information from the flow representation regarding the signal processing roles of nodes in the network. Nodes which have incoming edges with cardinalities of the Insets (Outsets) being larger than cardinalities of the Outsets (Insets), and outgoing edges with cardinalities of the Outsets (Insets) being larger than cardinalities of the Insets (Outsets) are, from the signal processing perspective, identified as sources of signals. The combination of divergent input (negative incoming CD sum) and convergent output (positive outgoing CD sum) is, considering the signal flow, equivalent to absorption of signals in the network. This is represented in the top left quadrant of the xy plane. On the opposite, the combination of convergent input and divergent output corresponds to the source characteristics of the nodes (bottom right quadrant of the xy plane). The top right and bottom left quadrants can be interpreted as a display of *signed* relay characteristics of the nodes. Nodes which have incoming edges with cardinalities of the Outsets (Insets) being larger than cardinalities of

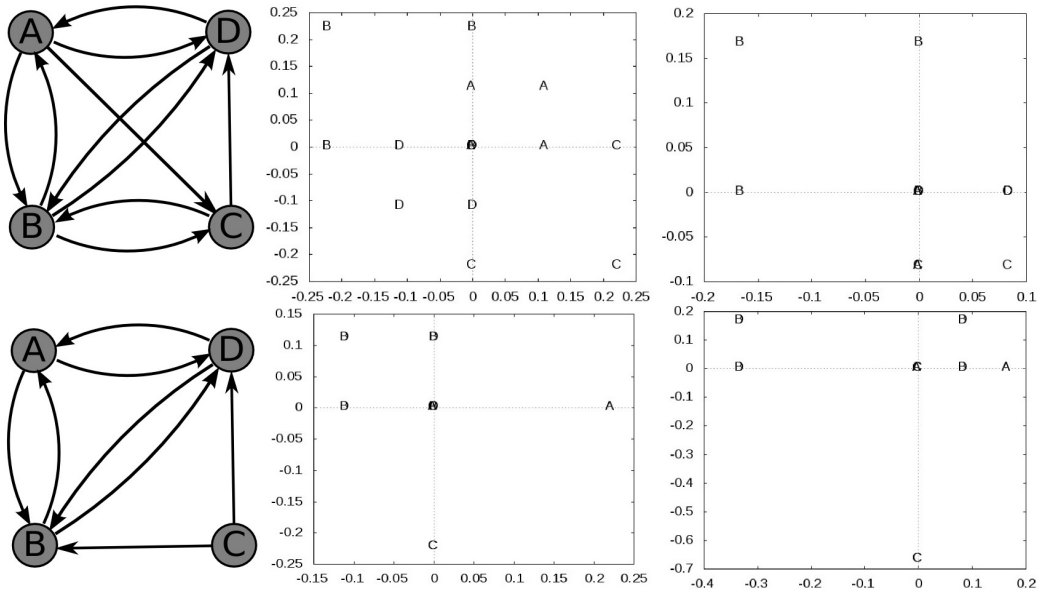


Figure 2.4: The lower graph differs in two edges from the top graph. The middle column represents graph nodes with σ_G^{tot}, the right column represents graph nodes with σ_L^{tot}. Every overlapping set is empty for the lower graph, because all chordless circles are of length two. Some points have the same coordinates in the flow representation. E.g., point D is is global, but not local CD flux sink.

the Insets (Outsets), and outgoing edges with cardinalities of the Outsets (Insets)being larger than cardinalities of the Insets (Outsets), are called negative (positive) router nodes. At the same time routing characteristics can be read from the top right and bottom left quadrants. Routers *redistribute* incoming CD of a given sign to outgoing CD of the *same* sign. Additional information is obtained from the z coordinate, which gives the average overlap of incoming and respectively, outgoing edges. This quantity identifies the degree of a node's participation in signal circulation in the network, a property typically associated with control circuits.

Graphical presentation of a network is not unique, e.g. isomorphic graphs may look totally different, the Petersen graph being a typical example. Community structure is not unique, grouping of points, thus presenting a network can be achieved in a multitude of ways. Yet, the flow representation of a network is *unique*, though due to possible symmetries it may have a significant amount of redundancy. This 3D plot of the network is unique in the sense that there is no arbitrariness in the position of the points in the three dimensional space. The flow representation can be considered as a network fingerprint since isomorphic graphs are mapped to the same plot, and differences between flow representations can be attributed to structural and functional properties of the network. If all edges are reciprocal or the graph is undirected, the flow representation of the network shrinks to a single point. The same argument applies to all graphs in which some nodes can not be distinguished due to symmetries. More precisely, nodes in the orbit of an element generated by the automorphism group of the graph are represented with the same point on the flow representation, as all the value of σ-s are constants on the orbits generated by the automorphism group of the graph.

Usefulness and application of the flow representation will be illustrated in the analysis of the real-world networks in Section 2.4.1.

2.4 Results

2.4.1 Signal flow characteristics of real-world networks

In this section I analyse functional clusters in real-world networks and the statistical properties of their interconnection. I analysed two biological and two artificial networks: macaque visuo-tactile cortex [102, 103], signal-transduction network of a CA1 neuron [92], the call graph of the Linux kernel version 2.6.12-rc2[1], and for comparison purposes the street network of Rome[2]. Nodes

[1] http://kernel.org/pub/linux/kernel/people/akpm/patches/2.6/2.6.12-rc2/

[2] http://www.dis.uniroma1.it/∼challenge9/data/rome/rome99.gr

and edges are defined as follows: in the macaque cortex nodes are cortical areas and edges are cortical fibres, in the signal-transduction network nodes are reactants and edges are chemical reactions, in the call graph nodes are functions and edges are function calls, in the street network the nodes are intersections between roads and edges correspond to roads or road segments. The first three networks perform computational tasks, Linux kernel manages the possibly scarce computational resources, signal-transduction network can be considered as the operating system of a cell, while cortex is an ubiquitous example of a system which simultaneously performs many computationally complex tasks. The street network is an oriented transportation network, which has a rich structure, as its elements have traffic regulating roles.

The call graph of the Linux kernel was constructed in the following way. I created the call graph of the kernel source which included the smallest number of components necessary to ensure functionality. The call graph was constructed using the CodeViz software[3], but it was not identical to the actual network of the functions calling each other, because the software detects only calls that are coded in the source and not the calls only realised during runtime. The resulting call graph had more than 10^4 vertices. As I wanted to perform clustering and statistical tests, the original data was prohibitively large, therefore I applied a community clustering algorithm [113] to create vertex groups. I generated a new graph in which the vertices represented the communities of the original call graph and have added edges between vertices representing communities whenever the original nodes in the communities were connected by any number of edges. Definition of the call graph nodes and their connections is analogous to the nodes and connections of the cortical network, as millions of neurons form a cortical area, and two areas are considered to be connected if a relatively small number of neurons in one area is connected to a small number of neurons in another area. The call graph of the Linux kernel will be discussed in Section 2.4.1.

The flow representations of two real-world networks are shown in Figure 2.5 and for comparison, in part A, the Erdős-Rényi network. We can identify the most important nodes and some general features of the networks as follows. Part B refers to the macaque visuo-tactile cortex. It is characterised by the alignment of the nodes along a straight line along the main diagonal, and hyperbolic-like pattern in the first and third quadrants, showing reverse ordering in the opposite quadrants, and absence of routers, which refers to a hierarchical organisation. In part C one can see the signal-transduction network of a hippocampal neuron. In the signal-transduction network of the hippocampal neurons, the molecules with the most negative CD flow are in-

[3]http://freshmeat.net/projects/codeviz/

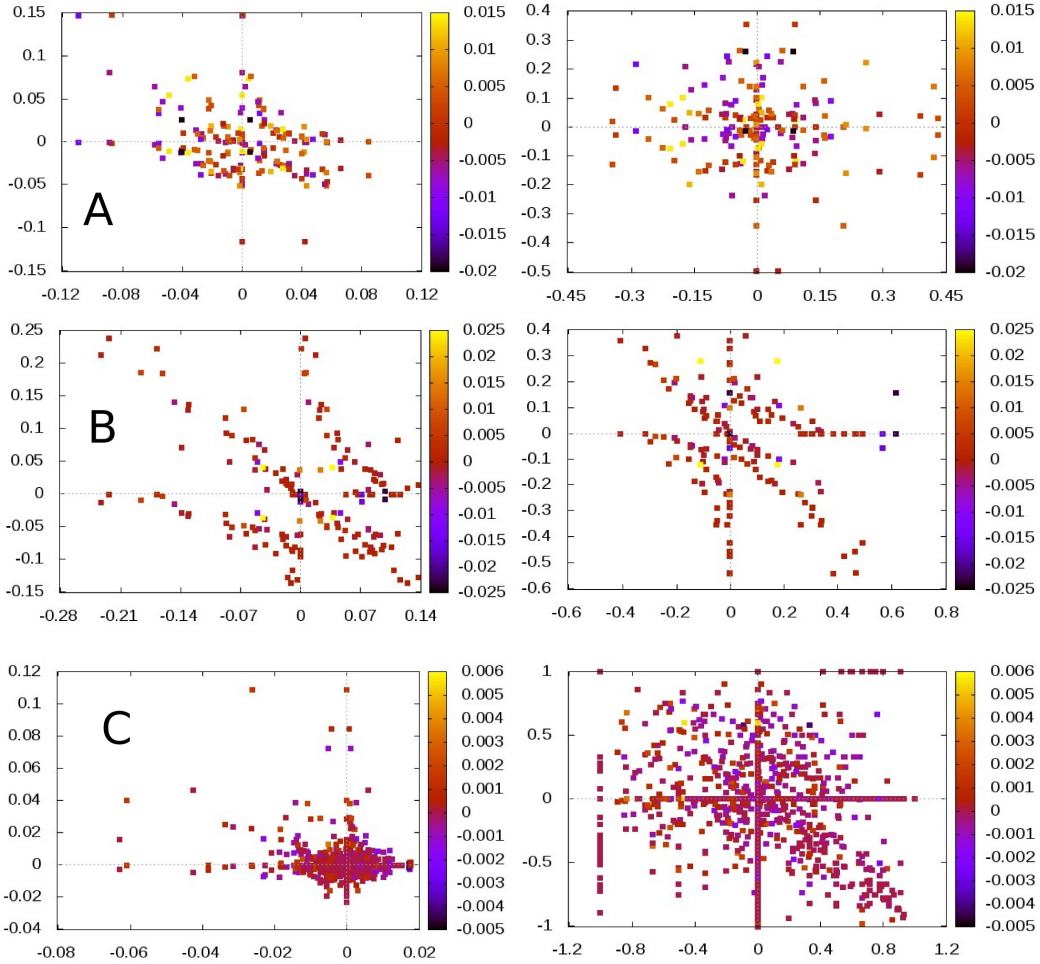


Figure 2.5: Components of the total CD_G flow are shown in the left column, components of the average CD_L are shown in the right column. Displayed are: Erdős-Rényi graph (row A), macaque visuo-tactile cortex (row B) and signal-transduction (row C). Relative overlap flow is indicated by colour intensity. Quantities on the axes are the same as in Figure 2.4.

volved, among other functions, in the regulation of key participants of the signal transduction cascade such as the cAMP second messengers. Molecules with large positive CD flow play function in cell survival and differentiation, as well as apoptosis. Router-like proteins are involved in diverse functions, notably the regulation of synaptic transmission in addition to those mentioned above. However, it should be noted that partly because of the paucity of my knowledge about many of the components of this network, as well as because of redundancy, i.e. overlapping functionality, we could give here only a very superficial classification. All edges of the signal transduction network fall in one of the three classes: excitatory, inhibitory and neutral, [92]. CD and overlap data were unrelated to the inhibitory, excitatory or neutral nature of network edges. Empirical distributions of CD-s and overlaps were alike for each edge class, see Figure A.1 in the Appendix.

Comparison of local and global structural organisation

I have analysed the flow representations in order to identify different features of signal processing. Network nodes are points represented in a 6D space of the flow representation, and in order to identify different signal processing, transmitting and controlling groups of nodes I performed clustering using Gaussian mixture and Bayesian information criterion implemented in R[4]. I wish to stress that the clustering I performed is not a form of community detection, but grouping of nodes with respect to their functional signal processing properties. Community detection can identify dense substructures, but it provides no information about the nature of signal processing, transmission or control. In each network I determined local and global, total and average signal processing clusters, have determined their properties, and have analysed the nature of CD-s and relative overlaps within and between clusters.

Clustering of nodes with respect to their functional properties resulted in contingency tables, with clusters being labels of the contingency table, and entries in the contingency able being numbers of edges within and between respective clusters. To estimate the randomness of the contingency tables I used the Monte Carlo implementation of the two sided Fisher's exact test. Number of replicates used in the Monte Carlo test was 10^4 in each case. The exact Fisher's test characterises the result of the clustering procedure, it quantifies how much the distribution of edges within and between clusters differ. I summarise the results in Table 2.1. For comparison purposes benchmark graphs were generated using algorithms described in [82].

[4]http://r-project.org

This algorithm is able to generate random modular networks with a better defined community structure than the ER-model. These graphs consisted of 45 vertices, and average and maximum in-degrees were set to 10 and 20 respectively.

Table 2.1: Number of functional clusters (n) and the corresponding p-values calculated using Fisher's exact test of the contingency tables. Q denotes the modularity of the community structure. Two numbers in a single cell denote the first two moments derived from sample size of 100 graph instances. Networks are denoted as follows: VTc - macaque visuo-tactile cortex, stn - signal-transduction network of the hippocampal CA1 neuron, kernel - call-graph of the Linux kernel, Rome - Rome street network, ER - Erdős-Rényi graphs and bench - benchmark graphs. Numbers were rounded to minimise the table size. Definitions of aggregated networks are given in Section 2.4.1.

network	n_{comm}	Q	$n_{G,tot}$	$p_{G,tot}$	$n_{L,av}$	$p_{L,av}$
VTc	4	0.332	6	0.48	9	10^{-4}
stn	58	0.530	3	0.75	19	10^{-4}
Rome	39	0.907	18	10^{-4}	19	10^{-4}
ER - μ	3.68	0.114	3.94	0.59	5.39	0.66
σ	1.55	0.020	2.34	0.30	3.34	0.29
benchm. - μ	3.19	0.449	3.83	0.19	5.21	0.10
σ	0.50	0.042	2.07	0.24	3.36	0.20
kernel aggr.	18	0.426	12	0.41	19	0.40
stn aggr.	9	0.34	18	0.38	7	0.05
Rome aggr.	6	0.46	8	0.24	5	0.86

For additional details on the CD distributions of real networks, see Appendix A.1

Based on Table 2.1, classification of nodes according to their functional properties does not match the network community structure. Classifying nodes according to their local and global functional properties differ substantially, further details are given in Table A.1. The p-values of the global and local groupings differ in the same way for all the networks analysed, though the difference is much smaller or absent for call graph of the Linux kernel. Distribution of edges between different node clusters measured by total CD_G flow in the signal transduction network was highly irregular, whilst very regular according to other flow measures. Note that the sizes of overlapping sets, and also the circularities were largest in the signal transduction

network, which was a consequence of edge sparseness. Measured by all the p-values, the street network had very regular structure, and was distinctively different from all other networks. In the case of Erdős-Rényi graphs there was practically no difference in randomness between local and global functional clusters, as presence of any community or structure in these networks was a matter of pure chance. Erdős-Rényi and benchmark networks were parametrised to match the macaque visuo-tactile network. The number of communities was comparable, but the number of functional clusters and the way in which edges connected functional clusters was different. The Erdős-Rényi and benchmark graphs were both structureless, but in different way. As one would expect, Erdős-Rényi graphs had much more randomness in the connection pattern between functional clusters than the benchmark graphs. In the macaque visuo-tactile network the connection according to the total CD_G was highly irregular, and resembled the Erdős-Rényi graph, according to other measures the connection pattern between functional clusters was regular, and differed from the either Erdős-Rényi or benchmark graphs. Summarising, the CD_G^{tot} flow representation is well suited to distinguish properties of signal and information processing networks and captures the characteristic features of signal transmission, processing and control.

Analysis of a prefrontal area network

As the dorsolateral prefrontal cortex is the area that plays a central role in the cognitive control of memory formations [23], a focused analysis of the areas related to this anatomical region is presented. Figure 2.6 shows the flow representation of an area subnetwork representing the prefrontal system. Results show that the most important sink of information in this network, where most of the information converges, is the dorsolateral prefrontal cortex, denoted by 46 in Brodmann numbering of cortical areas [26].

The other area that exhibits an accentuated sink property is the frontal eye field (FEF), which implements the control of visual attention, and the anterior cingulate area (Brodmann Area 24), which plays an important role in cognitive functions like decision making.

The dysfunction of the dorsolateral prefrontal cortex plays an important role in psychiatric disorders like schizophrenia [149]. An investigation of such dysfunctions based on fMRI measurements and statistical modelling is given in Chapter 3.

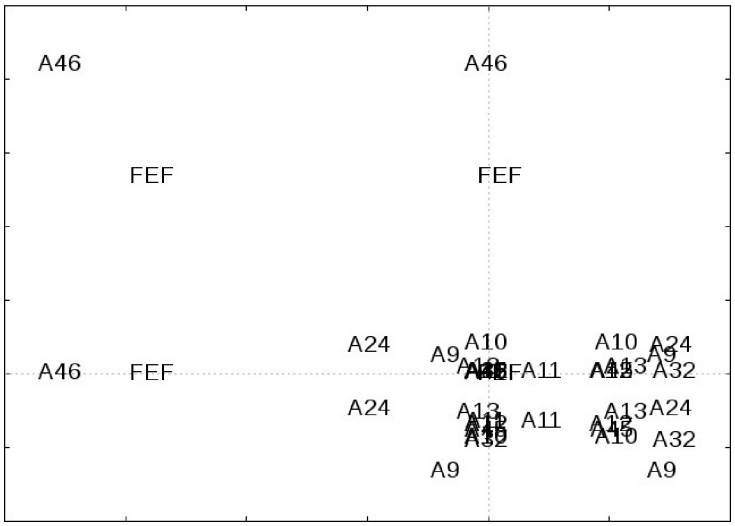


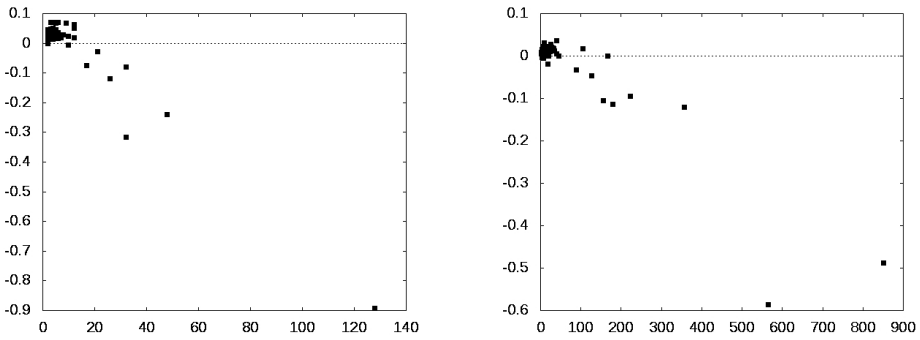
Figure 2.6: Depiction of prefrontal cortical areas in flow representation. From the position of the dorsolateral prefrontal cortex (Brodmann Area 46) assumed in quarter plane II, a role as an information sink can be implied. For abbreviated and numbered areas see Tables A.3 and A.4 respectively. Quantities on the axes are the same as in Figure 2.4.

Analysis of aggregated networks

The amount of data comprised in large networks necessitates community level understanding of signal flow. Communities themselves perform signal transmission, processing and control tasks, therefore determination of community level functional properties based on structural information poses a relevant problem. Number of communities in the street network and the hippocampal signal transduction network was large enough to define a non-trivial aggregated network which was subject of analysis. Each community in the original network was represented by a node in the aggregated network. Nodes of aggregated networks had additional structure, namely members of communities they represented, therefore allowing analysis relating CD and overlap flow with nodal structure.

The CD_G flow of the aggregated networks showed a regular pattern, nodes with positive CD_G flow were numerous and corresponded to small sized clusters in the original network, whilst nodes with negative CD_G flow were few and corresponded to large clusters in the original network, see Figure 2.7.

With some precaution (because of small network size and many unknown edges) analogous analysis of the whole macaque cortical network [124] can be performed. The aggregated network had four nodes, see Figure 2.8. Node



(a) Cellular signal transduction network (b) Linux call graph

Figure 2.7: Relation between CD_G flow (vertical axis) of the node in the aggregated network and the cluster size (horizontal axis) in the original network.

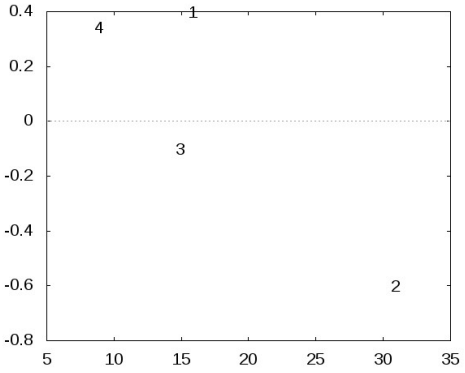


Figure 2.8: CD_G flow (vertical axis) of the node in the aggregated network versus cluster size (horizontal axis) in the macaque cortex. The communities are: 1- visual related, 2- higher cognitive functions, temporal, parietal prefrontal and hippocampal formation, 3- sensory-motor related, 4- auditory related.

with the largest negative CD_G flow corresponded to areas related to higher cognitive functions, the visual and auditory communities were smaller and had positive CD flows. Sensory-motor community had small negative CD flow, and was of intermediate size.

Similar analysis of the circularity flow revealed that nodes which corresponded to largest clusters in the original network had circularities close to zero. Because in- and out circularities of nodes corresponding to large clusters were nonzero, these nodes were well nested within chordless circles in the network. This nesting enables efficient performance of control-related tasks. CD flows of the original networks were mainly positive in the nodes corresponding to small, positive CD flow clusters. At the same time, only in nodes representing large clusters which had negative CD flow were numerous nodes with negative CD flows.Given the different nature of networks analysed, I conclude that organising principles in large-scale networks manifest dependence of functional roles on sizes of the network communities.

In case of the Linux call graph the most outlying nodes in the CD flow representation are the memory initialisation and buffer operators as CD flow sources, some of the CD flow sink nodes are connected to file system operations and the task scheduler. Flow properties of the aggregated street- and hippocampal signal transduction networks differ from the original networks, and resemble the properties of the macaque visuo-tactile cortex, as shown by aggregation of points along the $y = -x$ line in the diagonal quadrants, and grouping of points in the other two quadrants, see Figure 2.9. This is a signature of different organisation principles of signal transmission, processing and control properties at the community level, the net CD on the incoming side of a node is roughly redistributed on the outgoing side with a change of sign.

Statistical results of the analysis of functional properties were summarised in the lower part of Table 2.1. Randomness of connections between functional clusters in the aggregated street network strikingly differs from the original street network. Functional properties of the aggregated signal transduction network are similar to the functional properties of the cortical network, measured by the p-values. A possible explanation is that communities, i.e. functional cellular compartments of the signal transduction network have much better defined functional roles than single units, thus from the functional point of view, the role of nodes in the aggregated network is comparable to the cortex, when cortex is represented as a network of cortical areas.

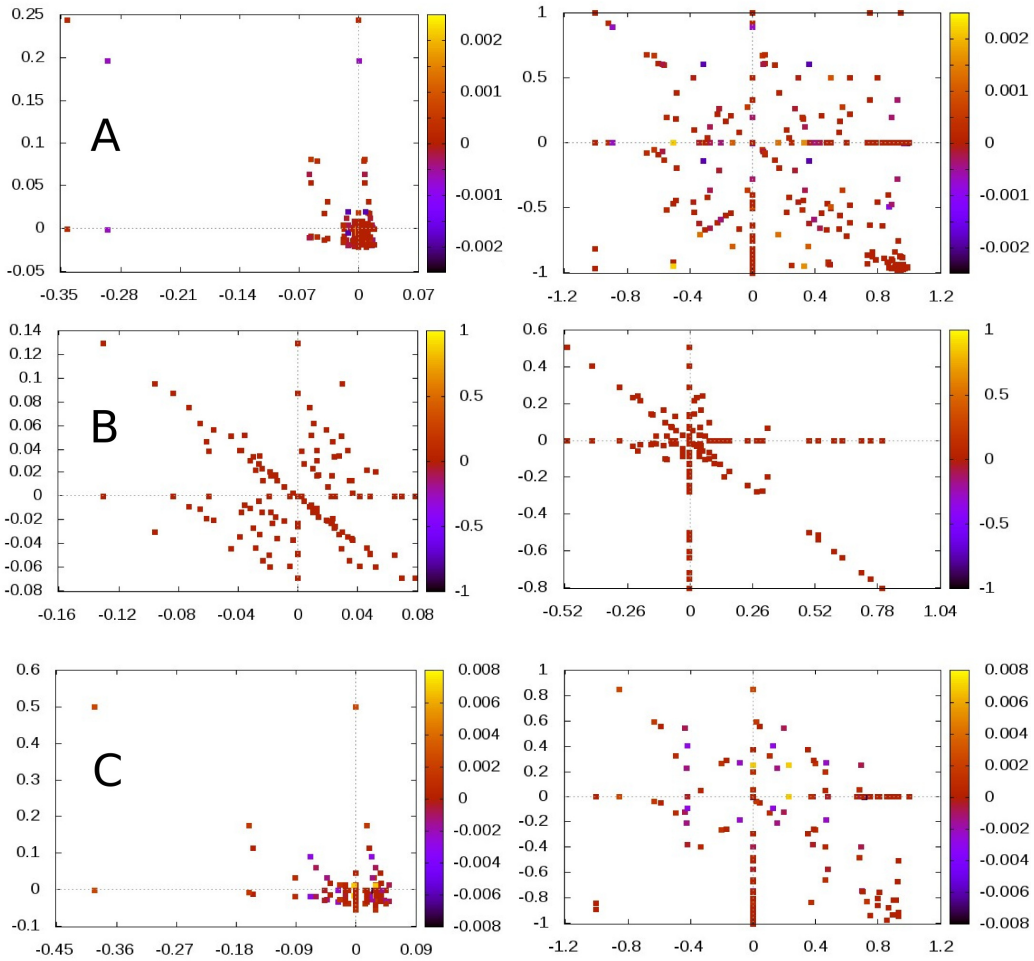


Figure 2.9: Components of the total CD_G flow are shown in the left column, components of the average CD_L flow are shown in the right column. Displayed are: Linux call graph (row A), street network (row B) and hippocampal signal transduction network (row C). Relative overlap flow is indicated by colour intensity. Quantities on the axes are the same as in Figure 2.4.

2.4.2 Signal flow in small-world-like networks

Random graphs [24] are useful models of real-world networks, thus one may ask which mechanism generates a cortical-like graph. Small-world property is often mentioned in relation to cortical (and other) networks. As CD- and overlap-related properties describe important features of signal transmission, processing and control, I studied whether signal flow properties can be obtained by the small-world generating algorithms. Macaque visuo-tactile cortex is strongly connected, even more, it contains numerous Hamilton circles. I constructed and analysed random graphs which matched prescribed properties of the cortical network.

The Watts-Strogatz graphs were generated as described in Algorithm 2, as given in [106]. If the reciprocity was preset, after each new edge with the probability defined by the reciprocity, I added an edge from the target to the source vertex as well. If there are several Hamiltonian circles in the network to be modelled, deletion of the edges in the original circle is unnecessary, as it only adds one Hamiltonian circle to the model graph. The resulting topology corresponds to a small-world network with a binomial degree distribution, indicating that all vertices are roughly equal as they have approximately the same number of connections. In this sense, small-world networks constructed this way are closely related to Erdős-Rényi random networks, the only difference being the Hamiltonian cycle that always exists. However, if some structure is to result from random addition of edges, the resulting graph has to differ from the structureless Erdős-Rényi networks. Therefore, I introduced preferentiality in the edge addition process to obtain a non-homogeneous graph.

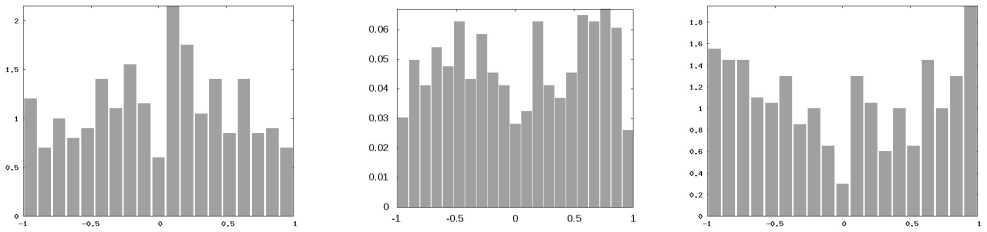
The new algorithm is called preferential small-world (PSW). It combines the ideas behind Watts-Strogatz and preferential attachment networks, as when adding random edges to the initial regular circle, the distribution of the source and target vertices were sampled as defined by the out- and in-degrees of the vertices respectively. This meant that a higher degree induced a proportionally higher probability for the vertex to be chosen as source or target. This procedure is detailed in Algorithm 4.

For statistical comparison I generated 100 graph instances of each network. Some numbers were rounded, in order to optimise the table size. I used Kolmogorov-Smirnov test to check whether CD-s and relative overlaps of the cortical and generated graphs originated from the same (statistically indistinguishable) probability density function. For each instance of generated graph the answer was negative, but the PSW algorithm generates a bimodal CD-distribution qualitatively resembling that of the VTc (see Figure 2.10). Statistical results are shown in Table 2.2.

Algorithm 4 The preferential small-world graph generation algorithm

Require: $n \geq 0$, $0 \leq r \leq n$

1: $G := G(V, E)$, $|V| = |E| = n$, $d_i^{in} = d_i^{out} = 1 \qquad \forall i \in V$
2: **for** $i = 1$ to r **do**
3: s,q := random integers between 1 and n, inclusive
4: $p_i := \frac{d_i^{in}}{\sum_{j=1}^{i-1} d_j^{in}}$
5: r := random real number between 0 and 1
6: $t : \sum_{j=1}^{t-1} p_j \leq r \leq \sum_{j=1}^{t} p_j$
7: $E := (E \cup \{s, t\}) \setminus \{s, q\}$
8: **end for**
9: **return** G



(a) Watts-Strogatz (b) Visuo-tactile cortex (c) PSW

Figure 2.10: CD histograms of the visuo-tactile cortex and its models

Table 2.2: ER denotes Erdős-Rényi graph, WS denotes Watts-Strogatz graphs, PSW denotes preferentail small-world, VTc denotes macaque visuo-tactile cortex. All networks were of the same size, $|V(G)| = 45$, $|E(G)| = 463$, and the proportion of the reciprocal edges was 0.8. Two numbers in a cell are the values of the first two empirical central moments, with the exception of Kolmogorov-Smirnov test results, where they denote D and p values respectively.

netw.	clust. coeff.	diam.	average SP	CD	KS-test D, p	Ovl	KS-test D, p
ER	0.550 $2 \cdot 10^{-3}$	3.1 $3 \cdot 10^{-2}$	1.88 $3 \cdot 10^{-3}$	$2 \cdot 10^{-3}$ 0.26	0.28 0	0.11 0.075	0.81 0
WS	0.600 $1 \cdot 10^{-3}$	3.06 $2 \cdot 10^{-2}$	1.89 $3 \cdot 10^{-3}$	$2 \cdot 10^{-3}$ 0.54	0.089 0.11	$5 \cdot 10^{-3}$ 0.03	0.047 0.66
PSW	0.623 $3 \cdot 10^{-3}$	4.32 $8 \cdot 10^{-2}$	1.93 $7 \cdot 10^{-3}$	$1.6 \cdot 10^{-2}$ 0.64	0.096 0.10	$5 \cdot 10^{-3}$ 0.03	0.046 0.299
VTc	0.517	5	2.15	$2 \cdot 10^{-2}$ 0.57		$8 \cdot 10^{-3}$ 0.45	

I conclude that description of cortical networks as small-world networks can be only a qualitative statement, as the small-world model fails to capture features relevant from the signal processing, transmission and control perspective.

In a second experiment investigating the role of other statistical properties of the networks, all quantities were obtained from averaging over 1000 randomly generated graphs. The measures in the focus of my interest were the averaged local clustering (or transitivity) coefficient and the average shortest path length [145, 147]. Clustering coefficient measures the probability of finding a triangle between all the possible triangles in the graph. It is a local characteristics. Average shortest path length is a global graph characteristics. Results describing the properties of generated networks are summarised in Table 2.3.

Comparison of graph features shown in Table 2.3 with the exception of clustering coefficient indicates similarity of macaque visuo-tactile cortex with the small-world graph with edge preference without preset edges connecting nodes at distance 2 along the initial circle. Therefore similarity at the large scales stems in the graph generating algorithm.

The flow representation of all the graph instances resemble each other.

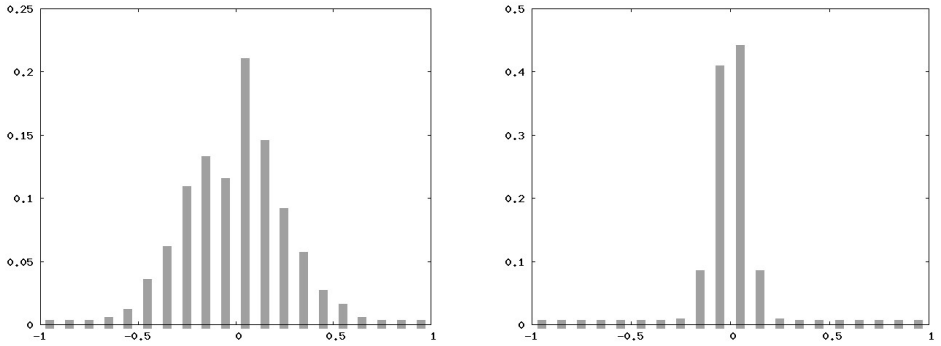
Table 2.3: Properties of networks generated using different procedures.

netw.	%	recipr.	cl. coeff.	diam.	avg. SP	CD
ER		.13	.55	3.10	1.88	$-4 \cdot 10^{-3}$, .54
ER		.83	.23	3.03	1.83	$-1 \cdot 10^{-3}$, .23
WS	0	.83	.6	3.09	1.88	$2 \cdot 10^{-3}$, .54
WS	40	.82	.6	3.08	1.87	$1 \cdot 10^{-3}$, .54
WS	60	.81	.6	3.08	1.86	$2 \cdot 10^{-4}$, .54
WS	100	.82	.6	3.06	1.85	$-3 \cdot 10^{-5}$, .53
PSW	0	.82	.62	4.63	1.94	$2 \cdot 10^{-2}$, .64
PSW	40	.8	.62	4.44	1.92	$2 \cdot 10^{-2}$, .63
PSW	60	.78	.61	4.21	1.90	$1 \cdot 10^{-2}$, .62
PSW	100	.8	.6	3.99	1.88	$1 \cdot 10^{-2}$, .62
VTc		0.816	0.517	5	2.15	$2 \cdot 10^{-2}$, .57

Numbers in the second column denote the percentage of edges which randomly connect nodes at distance 2 along the initial circle. Erdős-Rényi graph with pre-defined reciprocity has a given number of reciprocal edges, otherwise the addition of edges is completely random. For all random graph models, for all quantities, except for CD, $\sigma < 0.1$.

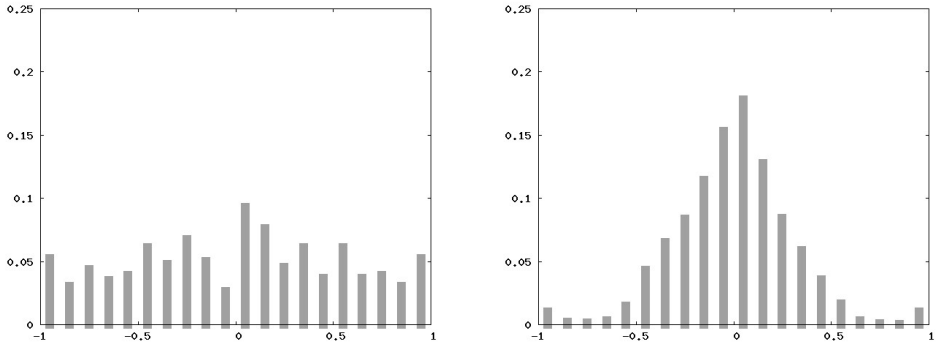
Yet, analysis of the flow representation reveals the underlying differences. Histogram of the CD values in all the graphs analysed sensitively distinguished different graph types. Erdős-Renyi graphs have unimodal, zero mean CD probability density function, with fast diminishing values on both ends (see Fig. 2.11(a)), small-world graphs have a unimodal probability density function, but non-vanishing tails (see Fig. 2.12(a)), while CD histograms of small-world graphs with preference are bimodal (shown in Fig. 2.13(a)). In case of small-world networks, the places of minima and maxima of the empirical CD probability density functions are exchanged, depending on graph generating algorithm.

To illustrate the dependence of CD histogram on network size, I show CD histograms generated from 10 sample networks with 450 nodes 46300 edges. The larger graphs were generated using the Watts-Strogatz algorithm (see Fig. 2.11(b)), without- and with edge preference (shown in Fig. 2.12(b) and Fig. 2.13(b) respectively). The probability of initial edges connecting nodes at distance two along the initial circle was set to zero. As we increase networks size, the limit of the resulting graph series is uniform for the ER graphs, indicating that the CD distribution would tend to a Dirac-delta. For the PSW graphs, the limit would not be structureless due to the preferential generation, thus their CD distribution tends to nonzero variance. For Watts-



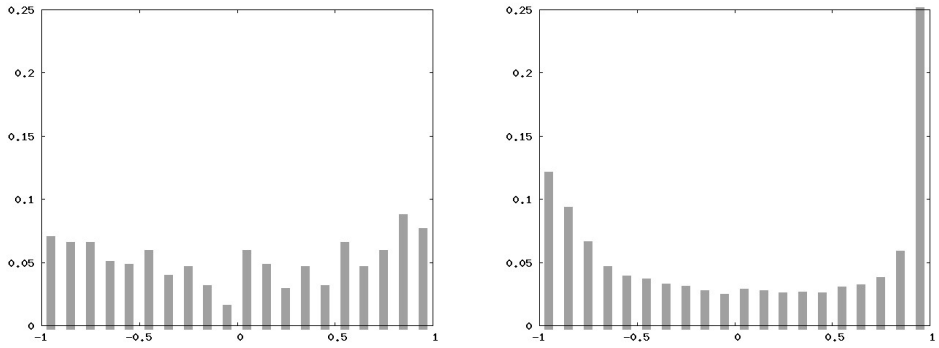
(a) $|V| = 45$, $|E| = 463$ (b) $|V| = 450$, $|E| = 4630$

Figure 2.11: Averaged CD histogram of 1000 Erdős-Rényi graphs



(a) $|V| = 45$, $|E| = 463$ (b) $|V| = 450$, $|E| = 4630$

Figure 2.12: Averaged CD histogram of 1000 small-world graphs generated by the Watts-Strogatz algorithm.



(a) $|V| = 45$, $|E| = 463$ (b) $|V| = 450$, $|E| = 4630$

Figure 2.13: Averaged CD histogram of 1000 small-world graphs generated by the preferential algorithm.

Strogatz graphs empirical evidence points towards having a Dirac-delta limit in the CD distribution as well.

The flow representation on the network is sensitive to the graph generating algorithm, i.e. graphs generated using different algorithms have different flow representations. I illustrate my statement on Figs. 2.14(a) and 2.14(b), where differences in histogram counts for graphs generated without- and with edge preference are shown. Differences in histogram counts are not results of chance as can be seen from the colour code. Red colour denotes small positive difference, yellow colour big positive difference, while blue colour denotes small negative difference and green colour big negative difference.

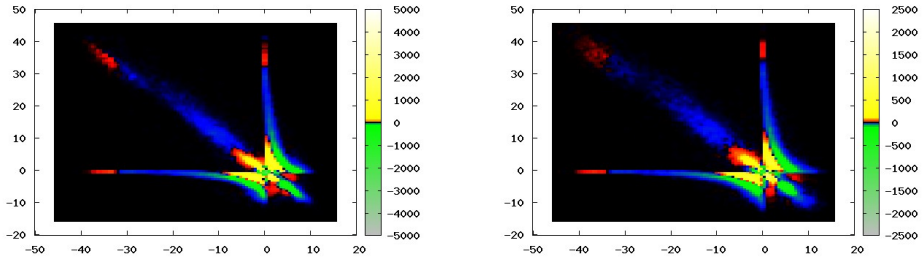
My simulation results also suggest that many global features of cortical-like structures can emerge as a result of random growth mechanism.

2.4.3 Model networks

It is possible to calculate the CD-s and overlaps or their probability density functions for some networks.

Arborescences

The purpose of calculating CD for arborescences is the comparison with networks grown with preferential attachment mechanism, see Section 2.4.3. I calculate global convergence degree of a complete directed tree – sometimes called arborescence. I assume that the root is at level 0, the number of levels is n, the branching ratio is constant and equals d and that all the edges are directed outwards from the root. For clarity, with the exception of the root,



(a) $p_2 = 0$ (b) $p_2 = 1$

Figure 2.14: Difference in histogram counts of the flow CD representation for 1000 graphs generated without- and with edge preference mechanism. p_2 is the probability of edges which connect points at distance 2 along the initial circle. $|V| = 45$, $|E| = 463$

all in-degrees are equal to 1, and with the exception of the leaves, all out-degrees are equal to d. If all assumptions are true, between any pair of nodes there is either no shortest path or there is only one. At level k ($0 \leq k \leq n$) the cardinality of any In set is k, while at level $k+1$ the size of any Out set is the sum of a geometric progression: $\frac{d^{n-k}-1}{d-1}$. Thus with some abuse of notation CD_G of any edge connecting nodes at levels k and $k+1$ equals:

$$CD_G(k, k+1) = 1 - \frac{2}{1 + \frac{k(d-1)}{d^{n-k}-1}} \tag{2.15}$$

We can observe that edges originating from the root have negative convergence degrees, but as the level index increases soon there are two possibly distinct levels k_1 and k_2, such that for $k \leq k_1$ CD_G is negative, whilst for $k \geq k_2$ CD_G is positive. k_1 and k_2 may coincide, or $k_2 = k_1 + 1$. k_1, and k_2 are determined by the solution of the equation $d^{n-k} + k(d-1) = 1$. Thus almost all edges have positive convergence degrees. One would naïvely expect that all the edges in such a tree are divergent, yet most of them are not. There is a level at which the number of the nodes in the In and Out sets results in the exchanged order of their (relative) sizes. The overall convergence in the whole network gives:

$$N(n, d) = \sum_{k=0}^{n-1} d^k CD_G(k, k+1) > 0 \tag{2.16}$$

Calculation of the local convergence degree is trivial:

$$CD_L(k, k+1) = \frac{1-d}{1+d}, \quad CD_L(n-1, n) = 1 \tag{2.17}$$

Contrary to the global CD there is only a trivial change in sign of the local CD.

Preferential attachment networks

Based on [19] I calculated the CD probability density function for the network grown with preferential attachment mechanism. This network has the structure of a random tree, therefore all overlapping sets are empty.

In growing networks it is natural to orient all the edges towards the root. For stratified networks, based on [19] one can derive local and global CD probability density function of nodes at distance n from the root, i.e. nodes at n-th level of the network. According to [19] the degree distribution at the level n is given as

$$f^{(n)}(k) = (1+y)\frac{\Gamma(2+y)\Gamma(k)}{\Gamma(2+k+y)} \tag{2.18}$$

where y is the depth measured in units of average depth:

$$y = \frac{n-1}{\langle n-1 \rangle} \tag{2.19}$$

Let x denote the CD_L of an edge connecting levels $n+1$ and n.

$$x = \frac{k_{n+1}-1}{k_{n+1}+1} \tag{2.20}$$

where k_{n+1} denotes the in-degree of the node at level $n+1$. Probability density of the local CD is calculated by changing the variable in Equation (2.18) according to Equation (2.20). The probability density of local CD having value x for an edge between levels $n+1$ and n is:

$$P_L(x, n) = \frac{2}{(1-x)^2} f^{(n+1)}\left(\frac{1+x}{1-x}\right) \tag{2.21}$$

Let $g^{(n)}(s)$ denote the probability of finding a tree rooted in the n-th layer of size s. $g^{(n)}(s)$ can be written as follows, [19]:

$$g^{(n)}(s) = \frac{1+y}{2+y}\frac{\Gamma\left(2+\frac{y}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\frac{\Gamma\left(s-\frac{1}{2}\right)}{\Gamma\left(s+1+\frac{y}{2}\right)} \tag{2.22}$$

Let x denote the random value of the global CD for an edge connecting levels $n+1$ and n.

$$x = \frac{s_{n+1} - n}{s_{n+1} + n} \tag{2.23}$$

where s_{n+1} denotes the fact that it is described with $g^{(n+1)}$. After changing the variable in (2.22), according to Equation (2.23), the probability density of the global CD for an edge connecting layers $n+1$ and n is:

$$P_G(x, n) = \frac{2n}{(1-x)^2}\, g^{(n+1)}\left(n\,\frac{1+x}{1-x}\right) \tag{2.24}$$

From the last term in the numerator of the Equation (2.22) one concludes that the domain of P_G is the open interval $\left(\frac{1-2n}{1+2n}, 1\right)$, which is the probabilistic equivalent of the global CD sign change observable in arborescences.

Erdős-Rényi graphs

Calculation of the CD and relative overlap probability density is based on the fact that all relevant probabilities are related to binomial distribution or a distribution derivable from a binomial one. Closed formulae for the local CD and overlap probability density function can be given, though they are lengthy, see Equations (2.25, 2.26). In the global case, the exact PDF are given by a recursive formula of considerable depths.

$$p_{LCD}(x) = \sum_{y,z=1}^{n-1} p_{LCD}\left(\frac{x(z-y)-y}{x-1}, y, z\right)\frac{|z-y|}{(x+1)^2} \tag{2.25}$$

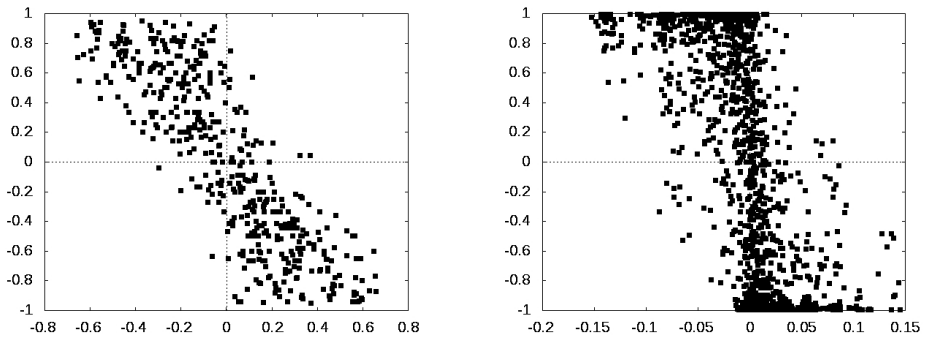
$$p_O(z) = \sum_{x,y=1}^{n-1} p_O\left(x, y, \frac{(x+y)z}{1+z}\right)\frac{x+y}{(z+1)^2} \tag{2.26}$$

For the complete derivation of the recursive formula see Appendix A.1.2

2.5 Discussion

2.5.1 Hierarchical organisation

Octants in the flow representation allow study of hierarchical organisation in the network, as flow sink nodes are assumed to be at lower hierarchical positions than the flow source nodes, [90, 103]. Flow sink nodes are connected with flow source nodes via edges with negative CD values, usually identified as feed-forward connections, while flow source nodes are connected to flow sink nodes via edges with positive CD, usually identified as feed-back



(a) Cortical network (b) Cellular signal transduction network

Figure 2.15: Relation between the CD flow through the nodes at the ends of an edge and CD of the same edge, points displayed have $(\phi_j - \phi_i, CD(i,j))$ coordinates.

connections, see Section 2.4.1. More precisely, based on graph structure it is possible to define a partial order relation on the set of nodes $V(G)$. Node i precedes node j according to the CD (ROvl) flow relation $\geq_{CD\,(ROvl)}$ if and only if $\phi_i > \phi_j$, where ϕ denotes the CD (ROvl) flow. In terms of hierarchical flow (HF) [90], $\geq_{HF} \equiv \leq_{CD}$. The consistency of classification edges as feed-forward or feed-back based on structural information is formulated as a relation between the CD flow through a node and the CD of edges attached to a node, and is shown in Figure 2.15, where the values of CD plotted against the difference of CD flows of the nodes at the two ends of and edge. The feed-forward or feed-back nature of edges could be verified using background information on the networks under study.

2.5.2 Functional implications of convergence degree

As my analysis of the real-world networks have shown, notions of convergence degree and overlapping sets may serve as initial steps in the task of relating a network's structure and functional properties it may have.

From the functional perspective, properties of the convergence degree and overlap can be understood as follows. Signals propagating through a given edge originate from the In-set, and are received in the Out-set. At the same time, signals are not simply transmitted or processed, as many real-world networks perform control tasks (about the controllability of network vertices, see **??**). Traditionally, in case of biological networks edges were classified as feed-forward and feed-backward and parts of control architecture were understood in such terms. Such an approach can be complemented with the

introduction of simplest control loops. The basic building blocks of control systems are comprised of chordless circles. Overlapping set and circularity grasp some properties of the control systems inherent in the network structure. The methodology introduced relies on the notion of shortest paths. Many real-world networks have large number of non shortest paths, for example to ensure fault tolerance. It is possible that not all the signals are transmitted along the shortest paths. The effect of non shortest paths can be grasped without introducing dynamics. My methodology can be extended in principle to answer how the functionality of network elements is altered. One may work with paths exceeding the length of shortest paths by one, and from the set of all such paths for each edge define the In and Out multisets, and proceed as I did. The procedure can be iterated if necessary.

My analysis of CD and overlap flows can be interpreted in terms of information flow and circulation. Identification of routers, sinks, sources and circulating nodes in the real-world networks was in accordance with the known functional roles of the nodes, for related previous work see [103]. Control and other loops were already investigated, [92] and classified as positive or negative depending on the nature of edges (excitatory or inhibitory) they contained. My methodology allows identification of an edge being feed-forward or feed-back in terms of CD flow and offer another definition of positive or negative feed-back loops. In the neuronal signal transduction network feed-forward and feed-back nature of an edge was independent from an edge being excitatory, inhibitory or neutral. Previous work concentrated on control-related motives which were subnetworks of relatively small size. In contrast, my methodology in its extreme can focus on the whole network. Analysis of aggregated networks revealed connection between functional properties of communities and their size. A possible explanation is that communities performing integrative tasks are highly specialised, and are comprised of relatively small number of elements. Communities performing allocatory and control related tasks perform broader class of more general tasks and are therefore comprised of larger number of elements. Allocation and control is centralised in the sense that the number of communities performing such general tasks is relatively small.

Functional roles and their interrelations are neither exact, nor sharp, they are rather tendencies observable after a suitable form of information reduction. My treatment of the flow representation resembles the phenomenological approach of [5], as nodes are represented in appropriate space, but the space in which I represented the nodes and the way in which nodes were grouped differed substantially. My analysis had three further gains: clarification of the network causality, demonstration of importance of chordless circles and a fresh look to the small-world characterisation of networks. Small-world

property is important and is defined with a generating algorithm which has a clear intuitive meaning. Yet contrasting small-world networks (generated using standard generating algorithms or their combination) with the cerebral cortex revealed that they had different CD statistics.

The cortical network has no pronounced routers, which fact may be related to the evolutionary process that optimised signal processing in the brain for speed. Evolution may also explain the lack of the nodes which only pass signals. Cortex preserved only the minimum number of nodes necessary for performing all the computational steps, i.e. every signal transmission is inseparable from signal processing. I demonstrated similar organisation in other aggregated networks.

My study of the Linux kernel call graph was far from complete, further analysis and inclusion of runtime calls will refine my interpretation of particular nodes at a finer scale. Deeper analysis of the neural signal-transduction network is likely to shed further insight into the low level signal transmission and processing of the cortex.

It was shown that signal processing, transmitting and controlling properties of a given network depend on the definition of a node. By aggregating a community into a single node and applying the same methodology, one can explore signal transmission and processing at the community level. Aggregated networks had different properties from the original networks, thus coarsening the network unit resolution revealed very different community-level information processing, transmitting and control properties. Further analysis of the real-world networks will be given elsewhere.

2.5.3 Local and global structure

In signal and information processing networks global functional organisation was much more random than the local one. This means that global and local organisation principles differ, and stochasticity may play a role on the large scale, while local connectivity is functionally more constrained.

The reason for global functional randomness can be understood as follows. Different processing streams have nodes with similar functional properties, though these properties are exercised over different domains, as it was shown for the cerebral cortex [103]. There is no general rule which would require connection between different integrator nodes in different domains, say. When there is such a connection it is likely to be an important one.

I have also shown a real-world example of a transportation network, which had markedly different properties from the signal processing networks. The finding is not based on comparison of structural, but rather functional properties. This was an example of how the nature of the network constrains its

functional organisation.

2.5.4 Other network measures

Convergence degree and edge betweenness [66] are both computed from the shortest paths. Edge betweenness measures the number of shortest paths passing through a given edge. In accordance with previous results for the macaque visuo-tactile cortex [102], on all the random graphs studied, convergence degree was statistically independent from the notion of edge betweenness, the largest absolute value of the Pearson correlation being 0.02. Besides computing the correlation coefficients, I also visually checked the relation, [7].

One may wonder whether similar conclusions about characteristics of network structure could be drawn from the knowledge of degree distribution. The question is not fully resolved yet, but some conclusions can be drawn. Note that many of the real-world networks under study have unknown degree distributions in the sense that the degree distribution is not known exactly but is rather approximated by some model that is thought to fit the data. These models have to be verified *a posteriori*. Reliable parameter estimation of degree distributions can be made only for very large networks. Many real-world networks (e.g. cortical, metabolic or other) are simply not big enough, so different exploratory methods are needed. Flow representation and its interpretation is meaningful for small networks too, and it may serve as an alternative to methods only applicable in case of large networks.

A study by Bagrow et al. [9] was published with the intention of detecting and representing network structural properties in a unique way. The B-matrices introduced therein use the notion of graph shells, which are graphically represented as matrices and depicted as two dimensional plots. However, interpretation of B matrices in functional terms is hard. The simple concept of convergence degree may fill the interpretation gaps.

Network motifs [3, 4] are subnetworks which appear much more than randomness would suggest. Motifs can be detected with appropriate algorithms, yet their interpretation is difficult, e.g. what is the meaning of two motifs with only slightly different structural characteristics, but with significant difference in their abundance in the original network.

3

Model-based dynamical analysis of functional disconnection in schizophrenia

The discovery of functional subnetworks in the cortex raises very different hardships than that of structural ones. Functional connectivity depends heavily on sensory input and the nature of the task the subject is perceiving to be solved. Thus, a single experiment can provide only partial information about two regions being connected. Of course, structural connectivity can serve as a prior approximation to the functional (where known), but intensities of coupling, and interactions between structurally unconnected areas may vary quite a bit. Such phenomena can only be studied in vivo, for obvious reasons, and the range of means to do so is often limited to non-invasive methods, especially in humans. Electrodes and electrode arrays can be used to record from within the neural tissue of rodents, monkeys with more restrictions and humans only during medical-

---

Related publications:

Bányai M., Diwadkar V., Érdi P.: *Model-based dynamical analysis of functional disconnection in schizophrenia.* NeuroImage **58**(3):870-877, 2011.

Érdi P., **Bányai M.**, Ujfalussy B., Diwadkar V.: *The schizophrenic brain: A broken hermeneutic circle. Some new insights and results.* The 2011 International Joint Conference on Neural Networks (IJCNN), 2011. San José, CA, USA 3024-3027, 2011.

Bányai M., Ujfalussy B., Diwadkar V., Érdi P.: *Impairments in the prefronto-hippocampal interactions explain associative learning deficit in schizophrenia.* BMC Neuroscience **12**(Suppl1):93, 2011.

Gore CD., **Bányai M.**, Grey PM., Diwadkar V., Érdi P.: *Pathological Effects of Cortical Architecture on Working Memory in Schizophrenia.* Pharmacopsychiatry **43**(Suppl1):592-597, 2010.

purpose recording of neural activity before surgeries. These methods are also constrained to record from a few locations at a given time.

The non-invasive, large-scale method to record the electrical activity of neuronal ensembles is EEG. However, the spatial resolution of the electrical imaging is very low, and inference about areas further from the skull is nearly impossible. A partial solution to this problem is functional magnetic resonance imaging (fMRI), which can record with a spatial resolution applicable to make inferences about area-level hypotheses of the entire cortex. In return, the temporal resolution is much lower, rendering the discovery of the fine temporal details of information processing mostly impossible. Nevertheless, fMRI has become the main tool of assessing area-level activation in the cortex, and building connectivity maps based on those. Dealing with the massively multivariate time series requires the prudent application of statistical methods. In the fMRI literature, different types of prevalent erroneous procedures are reported, including non-independent testing [144], insufficient correction for multiple measurements (or, depending on the interpretation, a revolutionary study about the social abilities of dead fish) [20], and other erroneous statistical comparisons [109].

Given the statistical nature of the construction of regions of interest from voxel data, the appropriate tools for network discovery between them are statistical models. These may be quite simple, as correlational maps, directed entropy maps, Granger causalities, coactivations based on linear models and so on. The simplicity and data-driven nature of such methods might allow for naïve network discovery in the sense of identifying connections in an absolute manner, but these findings are epistemologically incomplete, and may only serve as exploratory steps in the procedure of defining more mechanistic models. Hypotheses reflecting procedural understanding of the information processing underlying the observed activity should be expressed in terms of generative models, most likely dynamical ones of the kind described in Section 1.3.3. The definition of such models presupposes extensive knowledge about physiologically plausible connectivity patterns, and requires the hypothesis to be expressed as a difference of two or more distinct models, resulting in model sets to be compared. The inference about the exact probability distribution over the model sets reveals which hypotheses are more probable than others, yielding comparative results to be interpreted strictly within the pre-defined model set (although attempts has been made to extend the methodology to a more uninformed way of network discovery [63]).

It is also important to state, that these models incorporate causal assumptions, in the sense Judea Pearl uses this concept [111]. That is, we do not only assume that there is a conditional dependence or independence between certain variables, in the case of fMRI measurements, the experimental

conditions, neural activities of different areas and receptive BOLD signals. But we assume that there exist physical processes that connect these quantities in a mechanistic chain, establishing an actual cause-effect relationship between them.

The dependence of the results on informed priors can be reduced by taking a differential approach. That is, taking two groups of subjects for measurement that differ in a phenotypical aspect, typically one having been diagnosed with a psychiatric disorder, and the other with negative corresponding diagnoses. This way one can test alterations of the probability distribution over the proposed model sets, which may be easier to detect than absolute phenomena. From an epistemological point of view, it is always easier to spot parts that went wrong, than drawing an accurate picture of the phenomenon in question. And this kind of differential information can be then used to make inferences about the healthy functioning of the altered component, helping to create a more complete answer to the absolute question as well.

Schizophrenia is a psychiatric disease that affects high-level cognitive processes in various ways. In fact, patients may exhibit many types of positive and negative symptoms, forming very different behavioural subtypes of the disease. General attributes of schizophrenia include delusions, in the sense of creating theories detached from actual sensory experience (paranoid symptoms fit here), sometimes hallucinations, and dysfunctions of abilities required by everyday life. While the symptoms can be quite serious, the disease might leave many cognitive aspects rather intact, allowing some patients to conduct even intellectual activities. This characteristic allows a wide range of testing methods to be applied, aiming to discover the biological underpinnings of the disorder. The clinical relevance of such studies is also high, given that more than 0.5 percent of any larger human population is affected by schizophrenia, and up to this day only symptomatic medications exists, rendering the disease generally incurable.

In this chapter I will present a study aimed to discover the alterations of a specific task-dependent functional subnetwork of cortical areas in schizophrenia, utilising fMRI measurements and statistical modelling.

3.1 Overview

Schizophrenia is a complex polygenic disorder with diverse neural correlates. Altered fronto-hippocampal function and interaction [70] is a central aspect of its pathophysiology and may be related to anatomical and/or functional disconnection [55], and altered synaptic plasticity [131], which in turn may result

from a complex expression of genes on the brain's macro-networks [71, 136]. Here I investigated impaired macro-network interactions in schizophrenia by applying Dynamic Causal Modelling (DCM) [129] to the analyses of fMRI data collected during a paired-associate learning paradigm [27]. The aim of modelling neuronal interactions using DCM is to characterise: a) the intrinsic connectivity of the network and b) the contextual modulation of the intrinsic connections by psychological aspects of the task. In the present study I was interested in investigating whether or not there is a plausible difference between patients and controls in the intrinsic connectivity of the learning related macro-network, and in the modulatory effects on these connections by learning related variables such as time. This approach allowed me to simultaneously investigate hypotheses of disconnection (intrinsic connections) and reduced plasticity (modulatory effects of learning) in schizophrenia.

3.1.1 Disconnection hypotheses of schizophrenia

It has been hypothesised that schizophrenia is best understood in terms of pathological interactions between different brain regions. This claim can be formulated on two different levels: on one hand, we can investigate the *structural* organisation of the cortex and look for physical differences between patients and controls, and on the other hand, we can investigate whether there is a difference in the task-dependent *functional* interaction of cortical areas regardless to the underlying neural structure.

The concept of structural "disconnection syndrome" goes back at least to Wernicke [151], who interpreted psychosis as a result anatomically disrupted connections. It was reintroduced by Geschwind [64, 65], and had a crucial role in behavioural neurology and psychiatry [34]. Newer meta-analysis also supported the hypothesis [45].

The relevance of the concept of functional disconnection for interpreting schizophrenia now also seems to be promising: studies showed impairments in functional macro-networks in schizophrenia [91], and based on brain imaging experiments it was suggested that reduced performance of schizophrenic patients in cognitive tasks requiring working memory is related to abnormal prefronto-hippocampal connectivity [55, 59, 150]. The cellular bases of the two types of disconnectivities have also been studied, and two, somewhat interacting mechanisms, i.e. altered anatomical connections and impairments in synaptic plasticity were identified [56, 131, 132].

Specifically, I am interested in the functional reduction, both the qualitative and quantitative nature of it. Consequently, there are two questions to be answered: (i) what are the differences in the model architectures describing the information processing network of healthy and schizophrenia subjects,

and (ii) which connections are significantly impaired in schizophrenia? In technical terms, I analyse effective connectivities, which reflect the causal influence that one brain region exerts over another. Effective connectivity can consist of two components. The first characterises the intrinsic connectivity of the network, and the second models input-dependent changes in them. (Inputs, however, may have effects for brain dynamics not only by modulating connections, but also via direct or indirect influence on specific regions). Pathological connectivities may appear in both components.

Task-related functional connectivity can be investigated with respect to various functions of the brain, e.g. learning, memory, control, etc. I studied associative learning, since this is a cortical function that requires the integration of multiple sensory, representation and cognitive control pathways, making it a useful approach to grasp disordered functional interaction.

3.2 Material and methods

3.2.1 Associative learning: behavioural task and data

Paired-associate learning paradigms that require learning of associations between diverse memoranda over time, have been a cornerstone of the experimental learning literature [31, 32, 154]. These paradigms provide a framework for evaluating the role of the hippocampus in binding as well as in examining the temporal dynamics of neural signals that correlate with changes in performance in learning [50]. In addition to fronto-hippocampal interactions [120], associative learning is based in part on the hippocampal integration of memory streams as diverse memoranda must be bound into integrated associations [43]. It also relies on the consolidation and retrieval of associations between diverse memoranda, sensory inputs and streams of neural activity, particularly by hippocampal neurons. Yet tasks of paired-associate learning and memory have been infrequently applied in understanding dysfunction in schizophrenia, even though they have provided evidence of particularly strong deficits, relative to other tasks [155].

We adopted a paired-associate learning paradigm in which subjects are required to learning arbitrary associations between locations (in space) and objects (with unique identities). The two kinds of memoranda ("where" and "what") are processed by the two components, i.e. spatial (dorsal) and object (ventral), of the forward visual pathway [27, 72, 118]. It is assumed that these information streams converge in the hippocampus, with potential supervisory inputs from the prefrontal cortex [31]. The result macro-network (depicted in Figure 3.2) provides a relative rich framework to estimate model

architectures using DCM [134]. Through the repeated alternation between learning and retrieval epochs using a block design [6] (see Appendix A.2.2), I was able to capture learning dynamics in controls and patients over time (see Figure 3.1).

For details see Appendix A.2.3.

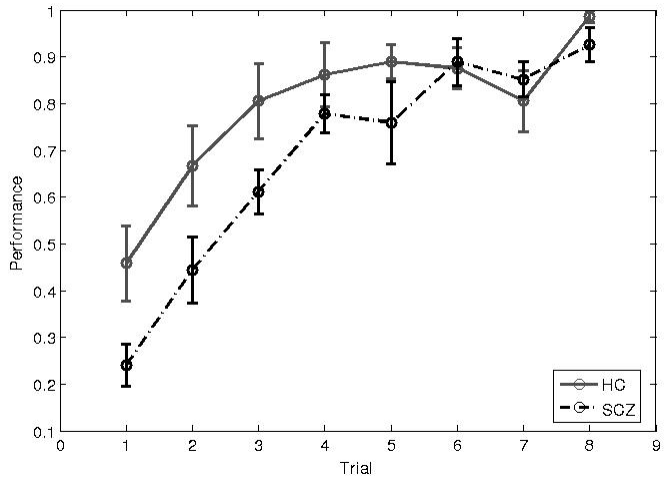


Figure 3.1: Learning dynamics in the associative memory task in controls and schizophrenia patients over time. The data provide evidence of generally asymptotic learning in both groups, with reduced learning rates in patients compared to controls.

3.2.2 Functional network models

Dynamic causal modelling (DCM) is an important method for estimating effective connectivity from neuroimaging data [61]. DCM uses an explicit model of neural dynamics to capture causal interactions between regions within the network (intrinsic connections), modulation of intrinsic connections by the experimental context (e.g., the valence of a face) and driving inputs to regions (e.g., visual stimulation driving face processing regions). Using Bayesian methods, DCM selects from among competing models, i.e., hypotheses, that best capture network interactions during the examined task [133]. Parameter estimates of inter-regional interactions derived from the model reflect measures of effective connectivity that can then be compared to assess significant differences between groups.

Macro-network models were evaluated based on a combination of anatomical and functional assumptions governing interactions between the five regions of interest material to the task. The basic architecture included intrinsic connections between primary and secondary visual cortices (from both dorsal and ventral pathways) that form two basic multi-synaptic pathways of the visual system [97, 98]. Connections from each of the inferior temporal cortex and the superior parietal cortex to the hippocampus reflect its unique anatomical place in a "hierarchy of associativity" [85] allowing it to integrate multi-modal inputs from unimodal areas before redistribution of potentiated associations into the neocortex [42]. This basic model structure was expanded in each of the evaluated models with intrinsic connections from the hippocampus to each of the inferior temporal and superior parietal cortices [85] that may enhance learning over time, and from the prefrontal cortex to the hippocampus, to implement supervisory inputs from this executive region [47, 110].

Elements of the models

DCM assumes a coupled dynamics of the neural activity in the involved brain regions and the external inputs. The connectivity parameters can be obtained by fitting the model to measurement data. The parameters describe effective connectivities, including the intrinsic coupling between brain regions (represented by an $n \times n$ matrix, A, where n is the number of brain regions, 5 in my case), and the effects of the external modulations on these connections (represented by an $m \times n \times n$ tensor, B, where m is the number of inputs, 4 in my case).

The cornerstone of dynamic causal modelling is to define the *external factors*, determined by the experimental paradigm, that modulate the intrinsic connections between brain areas. For the associative learning task I examined the role of four inputs. The first is called Visual, and refers to the presence of any visual stimulus. The second input is Encoding, which refers if the subject is in the encoding period, when the objects are shown in their location, and the third is Retrieval, which indicates if the subject is in the retrieval period, when a cue is presented in a location, and an answer is requested from the subject. The fourth input is Time, which is indicating that which epoch is the subject in, grasping the pass of time. These modulations allows the examination of the change in effective connectivity with respect to different memory functions and temporal development.

3.2.3 Model definitions

To model the *information processing* in the associative learning task, I assumed the presence of two streams connecting the five brain regions. The "forward" or "data" stream propagates sensory information at different levels of processing from the low-level sensory areas towards high-level areas. The "backward" or "control" stream propagates control signals from the high-level areas towards the lower-level ones. In this chapter I examine impairments in cognitive control, so the focus of the investigations is the control stream.

First, multiple models were evaluated by *varying* hypothesis-related *intrinsic connections* between regions, while fixing other connections (A matrix). I included the intrinsic connections of the data stream to all models. These are the causal effects of the primary visual cortex on the inferior temporal and superior parietal areas and the effects of SP and IT on the hippocampus and the prefrontal cortex. The self-connections of all areas are fixed as well. Based on the hypothetical control stream I defined three allowed connections that may extend the basic model in different combinations. These include the intrinsic connections from HPC to IT and SP, and from PFC to HPC. The eight possible combinations of these connections constitute the first model class. All possible intrinsic connections are visualised in Figure 3.2, where the fixed data stream is indicated by black arrows and the varied control stream is indicated by blue arrows. The connection patterns of the models in the first class are summarised in Table 3.1 and depicted in Figure A.4.

A second set of models were evaluated by fixing all intrinsic connections (both data and control streams) and some of the contextual modulations, and *varying* the hypothesis-related *modulatory connections* (B matrix). The fixed modulations were the effects of the Visual input on the connections of the data stream and of the Time input on all connections. I defined the allowed modulatory connections to be the modulatory effects of Encoding and Retrieval on the intrinsic connections of the control stream. I created 16 combinations of the allowed modulatory connections which constitute the second model class. Visualisation of the application of contextual modulations on intrinsic connections is given in Figure 3.2. The connection patterns of the models in the second class are summarised in Table 3.2 and depicted in Figure A.5.

In the first model class all meaningful modulatory connections mentioned above were included, (modulatory effects on nonexistent intrinsic connections were excluded only). The driving inputs of the contextual modulations on the cortical areas were included in all models and are not in the focus of this

study.

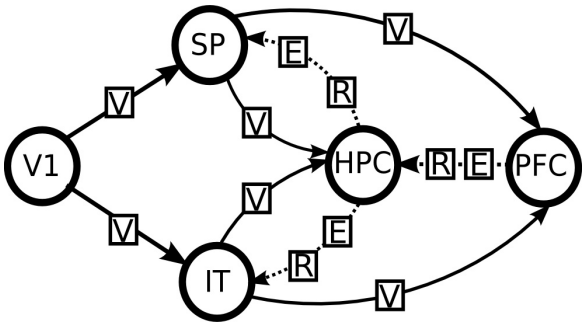


Figure 3.2: The model space for varying intrinsic connections. Connections marked by blue are varied. The effects of external inputs Visual, Encoding and Retrieval are marked by V, E and R respectively. The effect of Time is applied to all connections.

Coding of external modulations

The numerical values of external modulations as functions of time are constructed as follows. For Visual, the value is 1 in encoding and retrieval periods of the task and 0 in resting periods. The second input, Encoding, is 1 in encoding periods and 0 otherwise. The third, Retrieval, is constructed in a similar fashion. The fourth input, Time, is an integer, starting from 1 and increasing by one in the end of each epoch. The direct effects of all four inputs on the neural activity of all brain areas (C matrix) are present in every model I defined.

3.2.4 Dynamic causal modelling

DCM provides a complete phenomenological model framework for the analysis of BOLD data. For a detailed description see [61]. The model consists of two components: a neural state equation and a hemodynamic model. The neural component describes the time evolution of the neural state variables, x, which refer to the neural activity of the ROIs. This is a bilinear formula of the state variables themselves and the input variables, u, which are the conditions defined by the experiment (Eq. 3.1). The parameters of the neural model are the elements of the three matrices, $\theta_n = \{A, B, C\}$. A contains the intrinsic coupling parameters, the causal effects of the ROIs on each other, B contains the modulatory parameters, the effects of the inputs on the intrinsic connections, and C contains the direct effects of the inputs on the ROIs.

$$\dot{x} = (A + \sum_{i=1}^{N} u_j B^j)x + Cu \tag{3.1}$$

$$y = \lambda(x, \theta_h) \tag{3.2}$$

The hemodynamic component describes the nonlinear mapping from the neural activity to the BOLD signal, y, actually measured in the ROIs (Eq. 3.2). For the details of the hemodynamic model see [60]. The two components define a complete generative forward model of the BOLD signal generation, illustrated on Figure 3.3. To estimate the values of the parameter set, $\theta = \{\theta_h, \theta_n\}$ best fitting to measurement data, the "inverse problem" should be solved. One possible procedure to do so is the Bayesian maximum a posteriori (MAP) estimation technique defined by Eq. 3.3, where M denotes the specific connectivity pattern of the model.

$$p(\theta \mid y, M) = \frac{p(y \mid \theta, M)p(\theta \mid M)}{p(y \mid M)} \tag{3.3}$$

To exclude specific connections in my model, we can set the prior probabilities of the corresponding parameters to zero. For all probability distributions in 3.3, I assume that their form is normal, meaning that both the prior ($p(\theta \mid M)$) and posterior ($p(\theta \mid y, M)$) distributions are Gaussians, and the MAP estimation is defined as the mean of the posterior distribution.

3.2.5 Comparison of models

We can compare models with different connectivity patterns in a Bayesian fashion by estimating their model evidence:

$$p(y \mid M) = \int p(y \mid \theta, M)p(\theta \mid M)\, d\theta \tag{3.4}$$

The evidence is the probability of obtaining the actual measurement conditioned on the model form, integrated on the whole parameter space of the model. This way we get the overall probability that my actual model explains the measurement data regardless of the choice of parameters, and also punish models with a larger number of parameters. The computation of the evidence is usually not feasible, but there are several methods available to approximate it. One such method is the variational Bayesian approximation. In this method one defines a functional that is a lower bound of the log-evidence of the model. We maximise this functional to approximate the log-evidence.

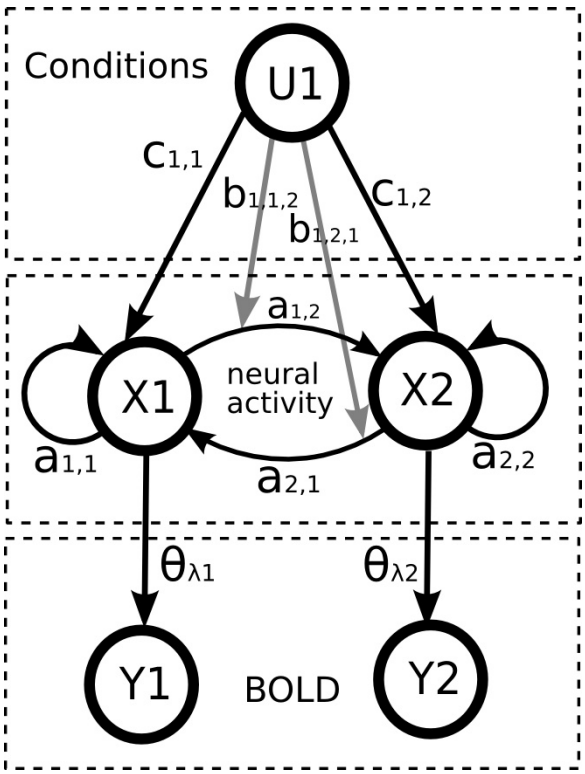


Figure 3.3: Structure of a DCM. U denotes the input variables described in Section 3.2.3. The X variables encode the unobserved neural activity of the regions of interest, according to the different model specifications described in Section 3.2.3. Y variables are the observed BOLD signal. Coupling dynamics of the variables and noise models used in estimation are described in Section 3.2.4

To obtain the expected posterior probabilities of all models in a model class, we can assume a hierarchical model of data generation. On the top level, a Dirichlet distribution describes the occurrence probabilities of the individual models in the population, defined by Equation 3.5.

$$p(r \mid \alpha) = \frac{\Gamma\left(\sum_k \alpha_k\right)}{\prod_k \Gamma\left(\alpha_k\right)} \prod_k r_k^{\alpha_k - 1} \tag{3.5}$$

On the next level, we have multinomial variables, parametrised by the occurrence probabilities of the above level, which describe the probability of a certain model generating the data for a certain subject (Equation 3.6).

$$p(M \mid r) = \prod_{i=1}^{k} r_k^M \tag{3.6}$$

On the bottom level, the actual data is generated by the dynamic model with parameters defined by the multinomial variable. The complete hierarchical model of generating data for a whole group of subjects is depicted in Figure 3.4. We can invert this model using a variational Bayesian method that requires only the estimates of the log-evidences for each subject-model pair. In this way we can obtain the parameters of the multinomial distribution. If we normalise these parameters, we get the expected posterior probability of each model regarding to the subject group, as in Equation 3.7.

$$E(r_k) = E(p(M \mid y_1 \dots y_n)) \tag{3.7}$$

For a detailed description of the comparison method see [133].

3.3 Results

All 24 models were fitted to the measurement data from all 14 subjects, obtaining 336 parameter sets, sorted into two groups, schizophrenia patients (SCZ) and healthy control (HC). The computational procedure applied serves two quantities, the maximum a posteriori (MAP) parameter values and model evidences. They were approximated by the Expectation Maximization algorithm [37], as implemented in the SPM package [62]. The results can be evaluated on multiple levels of abstraction. First, I made an intra-group model comparison to find out which model structures are more probable than others in the subject groups, then I investigated the differences on the level of the individual parameters.

3.3.1 Model comparison

The goodness of a model can be quantitatively described by its posterior probability obtained from Bayesian model selection. I applied the variational (free energy) method to approximate the log-evidence of the models. As model goodness, I calculated posterior probabilities of model structures by random effects analysis, which allows the subjects to be described by different models with the highest probability within a single group, and so is able to capture variability in the information processing structure applied by different subjects. The expected posterior probabilities are listed in the Tables 3.1 and 3.2. The comparison was done within the first and second

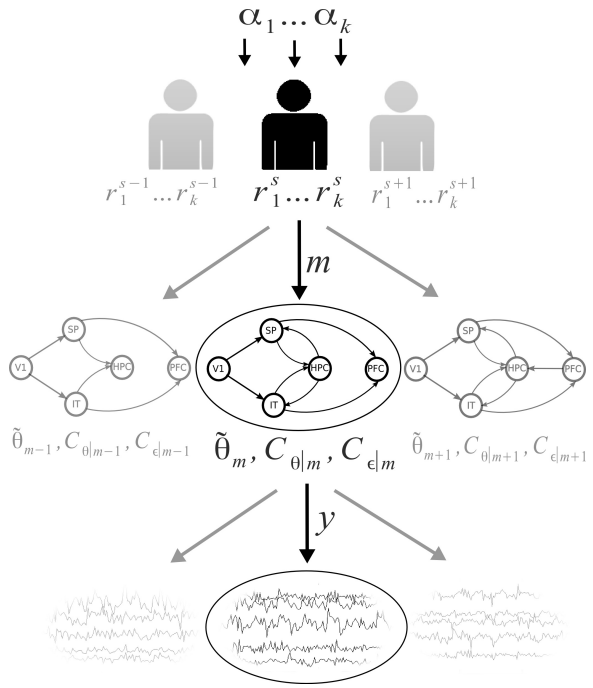
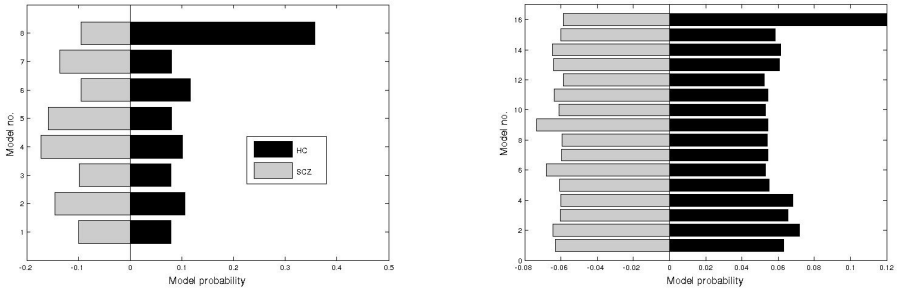


Figure 3.4: Hierarchical model of data generation. For each subject, a mixture model of data generation is assumed, governed by rate parameters r. After sampling the model m, the Gaussian parameters of the connectivity patterns are sampled (hemodynamic parameters are not shown), generating the observed quantities. Hyperparameters α of the prior over rates are shared between subjects, and correspond to the number of "effective occurrences" of individual models in the group.

model classes separately. To find out which models are to be considered probable and which ones improbable (related to the others), I formed two clusters in both subjects groups by k-means clustering. The scores of the models associated to the "probable" cluster are typed in bold face in both tables. The posterior distribution are depicted on Figure 3.5(a) and 3.5(b) respectively. Note that model no. 8 of the first class is identical to model no. 16 of the second class.

The results show that in the control group there is a clear winner for both the intrinsic and modulatory connection patterns, the model that contains the full control stream. In the SCZ group, there is no clear winner, there are several more probable models, and the differences are smaller between model probabilities. It can also be seen that while the winning model in the HC group contains all the connections defined, while the most probable models



(a) Models with varied intrinsic connectivity

(b) Models with varied modulatory connectivity

Figure 3.5: Posterior distributions over the two model classes

in the SCZ group lack more or less connections. This result implies that the information processing network of schizophrenia patients is fundamentally different from the one of controls. However, the model selection does not provide the specific pathways being impaired, so the parameter level analysis is also necessary. The result of the model comparison is depicted in Fig 3.6.

Table 3.1: Model probabilities for varying intrinsic connections

No.	**Additional connections**	**SCZ**	**HC**
1	none	.099	.079
2	PFC→HPC	**.145**	.106
3	HPC→IT	.099	.079
4	HPC→IT, PFC→HPC	**.172**	.101
5	HPC→SP	**.158**	.081
6	HPC→SP, PFC→HPC	.095	.116
7	HPC→SP, HPC→IT	**.136**	.081
8	HPC→SP, HPC→IT, PFC→HPC	.095	**.357**

Connections present in all models: V1→(IT,SP), SP→(HPC,PFC), IT→(HPC,PFC)

3.3.2 Effective connectivities

In the next step of the analysis, I give a more detailed quantitative characterisation of the results. I look at the parameter space level to see if there are significant differences in the effective connectivity in the models fit to

Table 3.2: Model probabilities for varying modulatory connections

No.	**Additional connections**		**SCZ**	**HC**
	Encoding	*Retrieval*		
9	none	none	**.063**	.063
10	none	PF→HC	**.064**	.072
11	PF→HC	none	.060	.066
12	PF→HC	PF→HC	.060	.068
13	none	HC→(SP,IT)	.061	.055
14	none	HC→(SP,IT), PF→HC	**.068**	.053
15	PF→HC	HC→(SP,IT)	.060	.054
16	PF→HC	HC→(SP,IT), PF→HC	.059	.054
17	HC→(SP,IT)	none	**.073**	.054
18	HC→(SP,IT)	PF→HC	.061	.053
19	HC→(SP,IT), PF→HC	none	**.064**	.054
20	HC→(SP,IT), PF→HC	PF→HC	.059	.053
21	HC→(SP,IT)	HC→(SP,IT)	**.064**	.061
22	HC→(SP,IT)	HC→(SP,IT), PF→HC	**.065**	.061
23	HC→(SP,IT), PF→HC	HC→(SP,IT)	.060	.059
24	HC→(SP,IT), PF→HC	HC→(SP,IT), PF→HC	.059	**.120**

Connections present in all models: Time→All, Visual→(V1→(IT,SP), SP→(HC,PF), IT→(HC,PF)

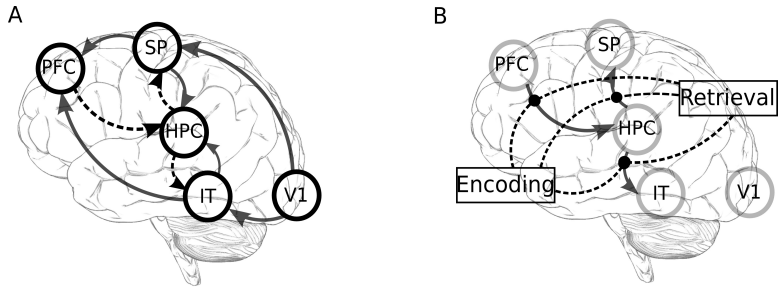


Figure 3.6: (A) Intrinsic connections in most probable fitted DCM models. Solid arrows denote causal connections present in both HC and SCZ groups, dashed arrows denote connections present in the HC group only. (B) Some of the modulatory connections in most probable fitted DCM models.

the two subject groups if we assume fixed model structure. To do so, I selected a reference model for comparison by running the model selection for

all subjects together with no distinction by group. The results can be seen on Fig. 3.7, the winning model is the one containing all hypothesised connections. The means and standard deviations of the intrinsic coupling and modulatory parameters are depicted in Fig. 3.8. To obtain the significance of the differences I applied a *two-sample t-test* on the parameter values in the two groups. For this analysis I used 8 subjects with schizophrenia and 10 healthy controls.

Marginally significant differences between the two groups are in the strength of the intrinsic connections between prefrontal cortex and hippocampus and between hippocampus and inferior temporal cortex. The relatively high p-values are likely to be mostly due to the small number of subjects. Such samples sizes are common for fMRI studies, but from a statistical point of view, they can be rather limiting. All the connections mentioned above are weakened in the SCZ group, which supports the hypothesis about the **impaired effective connectivities** in the control stream in schizophrenia. Both these connections are playing important roles in the cognitive control of the associative memory formation. Furthermore, we see the reduced effects of Time on these causal links meaning reduced excitatory contextual modulation of the pathways by learning. This can be seen as a surrogate of reduced task-related plasticity of a pathway in the illness.

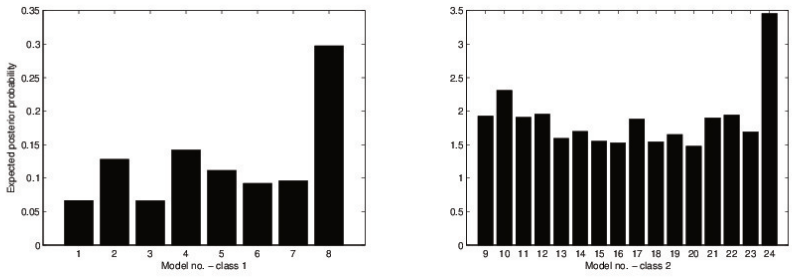


Figure 3.7: Model probabilities for all subjects. Model numbers are drawn from Table 1 and 2. Note that model nos. 8 and 24 are identical to each other.

The comparison of model parameters was also conducted in a Bayesian fashion, which supported the finding presented here. For details, see Section A.2.5.

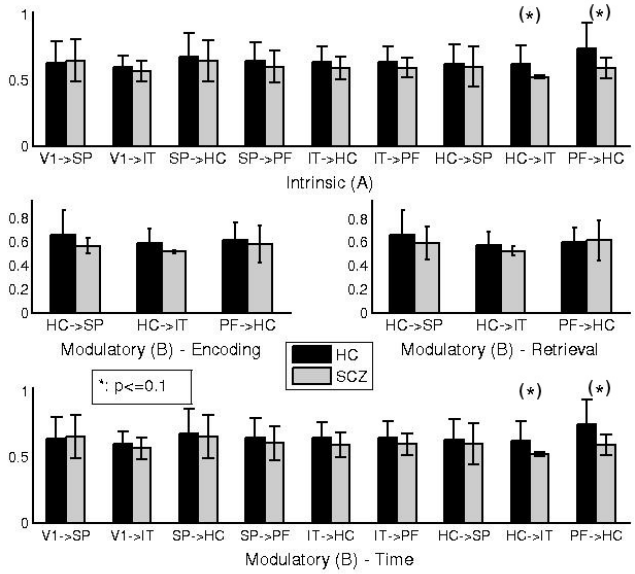


Figure 3.8: Average connectivity parameters for HC and SCZ groups. The most significant differences are in the prefronto-hippocampal and hippocampo-inferior temporal pathways.

3.3.3 Subject-by-subject analysis

To see a more detailed picture of the effects of the associative learning task on the model parameters, I correlated them with the learning rate of the individual subjects. To obtain this rate, k, for each subject, I fitted a learning curve defined by Eq. 3.8 to the behavioural data [27].

$$l(t) = 1 - e^{-kt} \tag{3.8}$$

The Spearman rank correlation coefficients between the DCM parameters and the learning rate was calculated. The results averaged over subjects are shown on Fig. 3.9. The correlations are mostly positive, and high for the hippocampal-superior parietal interaction in the intrinsic and also in the modulatory parameter arrays. This result is not in full correspondence with the inter-group comparison of the effective connectivities, suggesting that the learning rate is not necessarily the key element of the differentiation between schizophrenia and other conditions. This is also leading us to the next step of the analysis.

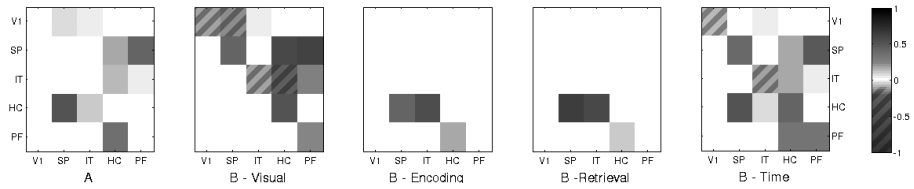


Figure 3.9: Correlations between the learning rate of the subjects and the connectivity parameters of the models fit to their BOLD data.

3.3.4 Illness versus slow learning

A common problem in schizophrenia research is that it is usually hard to separate the effects of the illness on learning skills from the lower performance of naturally slow learners. To address this issue, I selected the subjects from the control group who did not perform better than the SCZ group (there were 3 such subjects in the HC group). I fitted the model space to these subject separately. The resulting posterior model probabilities are shown in Fig. 3.10. It is apparent that the distribution over the model class is similar to the one obtained for the control group and shares no common features with the one obtained for the SCZ group, as one can see in Table 3.1. This result suggests that the methods applied here are independent of the overall learning rate and are able to clearly separate slow learning from schizophrenia by explaining the two phenomena by different model structures.

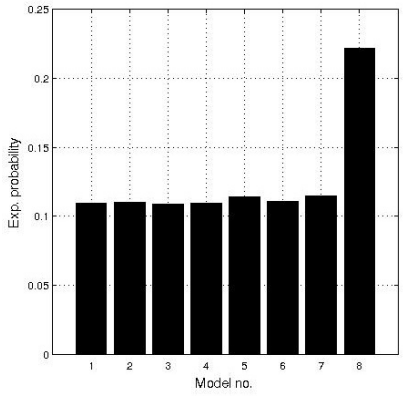


Figure 3.10: Expected posterior probabilities of models from the first class for slow learners. The distribution is similar to the one of the control group and distinctively different from the one of the patient group. Model numbers are drawn from Table 1.

3.4 A synaptic model of working memory in schizophrenia

To assess the possible contribution of intrinsic connectivity alterations in the dorsolateral prefrontal area to the symptoms of schizophrenia, I present a study of a neural network level model of working memory, consisting of 10000 integrate-and-fire neurons, from which 2000 are inhibitory, connected randomly by a probability of 0.2. Synaptic connections are strengthened within certain subpopulations by some form of prior learning (e.g. Hebbian), which will be used to store one-bit memory patterns. The model and its dynamical behaviour is depicted in Figure 3.11. The recallability of a memory pattern with a nonspecific activation of the whole network depends on the synaptic dynamics, detailed in the following section.

3.4.1 The model framework

The synaptic theory of working memory was suggested by [99]. A simple model for the the prefrontal cortex was specified, exploiting the general belief that in this brain region the excitatory synapses are facilitatory. Working memory is therefore generated and maintained by short-term synaptic facilitation.

A reduced short term plasticity model uses two variables, x is the available resource (released transmitter molecules) and u is the utilisation variable (residual calcium level). The increase of u is called the facilitation, the decrease of x is the depression, and the product $u * x$ characterises synaptic change. The process is controlled by two time constants: τ_f and τ_d denoting facilitatory and depressive time constants, respectively. The model is defined by Equations 3.9 and 3.10.

$$\frac{dx}{dt} = \frac{1-x}{\tau_d} - ux\delta(t - t_{sp}) \tag{3.9}$$

$$\frac{du}{dt} = \frac{U-u}{\tau_f} - U(1-u)\delta(t - t_{sp}) \tag{3.10}$$

In [99] the time constants were fixed as $\tau_d = 0.2s$ and $\tau_f = 1.5s$ to express facilitation. The motivation for my experiment came from the ending of [99] : "... The model provides a possible target for a pharmacological interference with WM. In particular, manipulations that modify the facilitation/depression balance in the memory-related cortical areas [...] are predicted to have a strong effect on the stability and duration of memory".

3.4.2 How the time constants might be regulated?

Calcium binding proteins (e.g., neuronal calcium sensor NCS-1) modify short-term plasticity (at least in hippocampal cell cultures) by switching pair-pulsed depression to facilitation [121]. As facilitation in my models appears to signify normal performance but not schizophrenic performance, we might expect that NCS-1 concentration would be large in the normal brain, but low in the schizophrenic brain. This prediction is inconsistent with empirical evidence: NCS-1 is up-regulated in the PFC of schizophrenia patients [81], suggesting that facilitatory synapses should be normal and not schizophrenic. However, the molecular machinery might be much more complicated for the following reasons: (i) NCS-1 might be doubly localised pre- and postsynaptically [101], (ii) NCS-1 is a part of a network of proteins.

Given the complexities mentioned above, a realistic detailed mechanism for changing the balance between facilitation and depression cannot be given at this point of my studies. However, I was able to study the dynamic properties of the system in the two-dimensional parameter space of the time constants. To evaluate the performance of the memory system, I had to define the duration of the memory. The hypothesis was that the shift in balance between facilitation and depression might modify the duration of the memory. Phenomenologically two types of pathology, "too short" and "too long" could emerge. The question is whether the duration of memory depends on the two time constants and reduced connectivity, respectively.

3.4.3 Definition of duration of working memory

Working memory was defined as the time between the end of the write-in signal (the population-specific increase in the background input that loads an item to the memory) and the last point in time when the object can be retrieved from the memory.

Retrieval in the readout signal (a nonspecific increase in the background input), would produce a population spike (PS). This PS codes for the object loaded in the memory previously and refreshes the memory as well. The probability of observing a PS is mostly dependent on the actual level of synaptic efficacy. We can define a threshold value in efficacy that divides the two behaviours of the network (PS or not). So we can define the duration as the time between the endpoint of the write-in signal and the time-point when the efficacy falls under the threshold value.

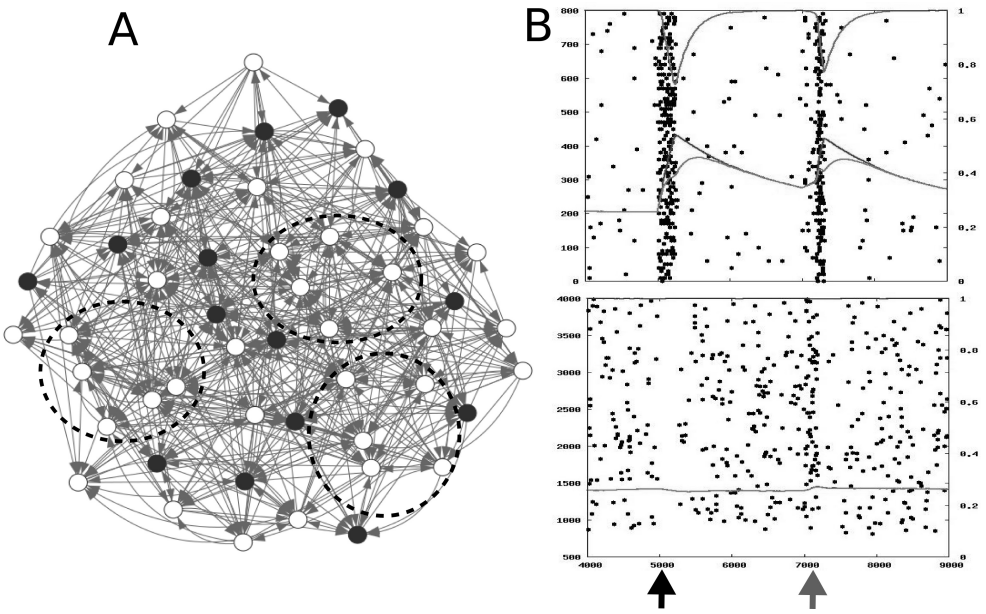


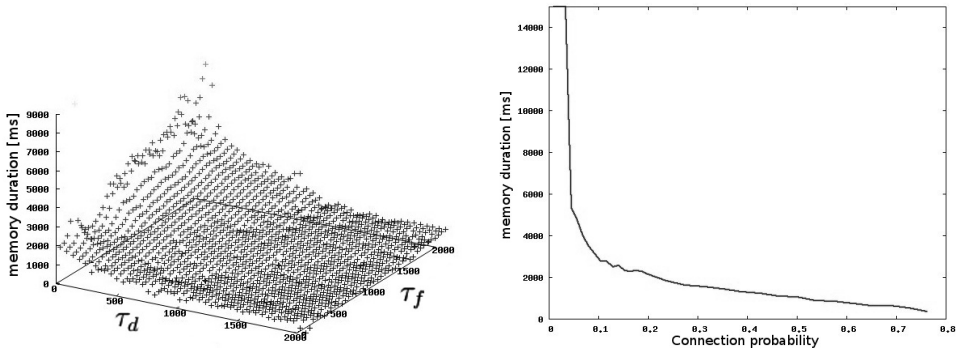
Figure 3.11: Prefrontal working memory model. (A) Architecture of the network, downscaled to 100 vertices. Black vertices depict inhibitory neurons. Dotted circles indicate coding populations with strengthened intrinsic connections. (B) Dynamic behaviour of the model. The above raster plot shows the encoding and retrieval of a memory signal in a coding population. The raster plot below shows a population of non-coding neurons during the same period of time. On both plots, the relative amount of available transmitter, the release probability and the synaptic efficacy is shown by the upper, lower and middle curves respectively. The encoding signal is applied to the coding population at the time point specified by the black arrow, and the read-out signal to the whole network at the grey arrow.

3.4.4 Simulation results

Fig. 3.12(a) shows explorations of the two-dimensional parameter space. These intuitive results indicate that if facilitation relaxes slower and depression relaxes faster, memory duration will increase. One could define a regime to be considered normal, and so there would be two regimes: one for too short and one for too long memory fading time.

Second, the connectivity was changed by setting the overall connection probability from 0 to 0.8. Somewhat counter-intuitively, the duration was reduced by increasing the connectivity, as one can see in Fig. 3.12(b).

However, if we look at the model setting, the cause for this behaviour is



(a) Efficacy and depression parameters (b) Connection density

Figure 3.12: Parameter dependence of memory duration in the synaptic model. Changes in any of the examined parameters may result in abnormally short or long recallability of memory patterns.

obvious. I apply an external input on all the cells, which is modelled by a Gaussian noise with a large mean and small deviation. The mean is actually above the firing threshold of the cells, so if there were no other dynamics, they would fire permanently with a frequency defined by the refractory period. The principal effect of the cells on each other is the inhibition, allowing them to follow different firing patterns.

3.5 Discussion

A widely studied aspect of schizophrenia is the impairment of cognitive control over information processing cortical circuits. The prefrontal cortex is the area that is consensually considered as the centre of cognitive control functions, such as attention, memory and executive functions. Numerous functional imaging studies pointed out decreased influence of the prefrontal cortex on areas which are material in tasks effected in schizophrenia. For a review on pathological functional connectivities related to different cognitive functions see [33]. The architecture of cognitive control in schizophrenia is still not well understood, for preliminary ideas see [16, 35, 115], for the normal and pathological neural circuitry of executive functions, see e.g. [44]. Reduced fronto-temporal functional connectivity associated with auditory hallucinations for schizophrenic patients was extensively studied, see e.g. [86, 141]. A focused strategy training was suggested recently [41] to facilitate cognitive task performance in patients with schizophrenia by changing the dynamics of activity within critical control-related brain regions.

This study targeted the prefrontal control of areas involved in associative learning, primarily the hippocampus. Several model architectures were defined to explain the information processing in the cortex during learning. I compared groups of schizophrenia patients and of healthy controls on two levels, and a fundamental difference between the functional networks implemented in the schizophrenic and healthy brain were found. The parameter level comparison revealed significant impairments in the prefronto-hippocampal and hippocampal-inferior temporal pathways. This finding implies that the lack of cognitive control over the processes of associative learning may underlie the decreased performance of schizophrenia patients in related tasks.

Learning impairments have been considered good markers of hippocampal impairment in schizophrenia [155], and computational models of hippocampal function have been applied to study this cognitive impairment in the disorder [38, 119]. However, the basis of these impairments in terms of network interactions has not been known. My results (the first based on in vivo fMRI data) provide evidence of impaired frontal inputs to the hippocampus, and reduced learning related plasticity of fronto-hippocampal coupling in the disorder.

These results also shed new light on previous studies about the dynamics of the schizophrenic cortex. A previous study from Érdi et al. proposed a model for the cognitive control deficiency in schizophrenia on the neural network level [48]. In this model, the prefrontal cortex acts as a switch that drives the hippocampal formation to learning and recall modes. Slight changes in the accuracy of the control variable reproduces the performance alteration of patients compared to controls. The results presented here are compatible with this mechanistic idea, and together can provide a multi-level picture of the studied phenomena.

3.5.1 Connecting function to structure

The integration of the frameworks discussed in this chapter and Chapter 2 would be highly desirable, as it could provide a consistent account of the organisational principles of cortical macro-networks, and shed new light on the interplay of structure and function in the cortex. To proceed in this direction, one should work on the scaling up of the functional network discovery methods, so that they might produce task-dependent or resting state causal networks with a few ten vertices, possibly also interconnecting multiple tasks and subnetworks. The exact relations of graph theoretical measures to the quantities defined in a probabilistic causal model should also be determined to construct such an integrative approach.

4

Conclusions

4.1 Signal flow in directed networks

The first part of the dissertation introduced the convergence degree as an edge based measure applicable to characterise information flow in directed networks. Based on this measure, a novel method of graph representation was also introduced, which is invariant on the automorphism group of the graph, providing a graphical way to depict network structure. The method to compute probability distributions of the CD for parametrised random network models is given in Section 2.4.3.

The method is applicable in the structural comparison of real-world networks, as presented in Section 2.4.1. From these I conducted the analysis of the visuo-tactile cortical network of the macaque and an intracellular signal transmission protein network, showing a strong hierarchical structure in the latter. I evinced that, where known, the signal transmission roles of the vertices determined by convergence degree is consistent with biological function [92]. For the vertices with unknown function, the method provides a prediction.

To construct a random network model of the cortical area networks, I proposed the preferential small-world graph generation algorithm, described in Section 2.4.2. This algorithm is proven to produce networks with statistical properties much closer to the macaque visuo-tactile cortical network, using the convergence degree distribution as a criterion for similarity.

As shown in Section 2.4.1, I also analysed a prefrontal area network, demonstrating the special role of the dorsolateral prefrontal cortex in the hierarchy of cortical areas, providing a quantitative background to the speculation about the area's prominent role in higher cognitive function.

This part of my dissertation can be summarised as follows:

T 1/1. I showed that the convergence degree measure is suitable for investigating the structurally determined signal flow properties of directed networks, can be used to create more refined classification systems than traditional graph theoretical measures, and also to give a more precise definition of the signal processing roles of individual vertices and edges.

T 1/2. I devised a preferential rewiring graph generation algorithm that gives a better model of the cortical macro-network than previously defined random graph models in the sense of reproducing more structural properties.

T 1/3. I determined the role of the dorsolateral prefrontal area in cortical signal processing quantitatively using the convergence degree: the information flowing between cortical areas shows a strong convergence on the dorsolateral prefrontal cortex.

A further application of the convergence degree is the analysis of aggregated networks (see Section 2.4.1). These are large-scale representations of big networks, where we replace the more strongly connected sets of vertices by a single vertex. I analysed the aggregated version of the protein network mentioned above, the road network of a city and the procedure call graph of the kernel of an operating system. In the latter, I showed that control flow converges on low-level system calls. I showed the statistical connection between the sizes of clusters represented by the vertices and their convergence degree. These results were published in [11, 13, 104].

Possible future directions of research building on the results presented here include the definition of consistent differential geometrical operator sets on directed networks, incorporating structural constraints represented by the convergence degree. Some advancements in this direction has been made by László Lovász in [89]. Such results would greatly further the theoretical basis of dynamical systems models defined on directed networks, pointing towards an integration of network structure and function.

Further analysis of the cortical area network using the convergence degree can be conducted investigating edge importance, as in [105], and possibly other measures of structural constraints on function and dynamics.

4.2 Model-based dynamical analysis of functional disconnection in schizophrenia

The second part of the dissertation presented a model-based data analysis approach to infer functional subnetworks of cortical areas and their alterations in schizophrenia. The analysis was conducted using a dynamic causal modelling framework, fitted to an fMRI measurement of schizophrenic patients and healthy controls performing an associative learning task. Disconnection hypotheses were tested by a definition of two sets of models (see Section 3.2.3), investigating control flow integrity of visual memory formation.

The difference between the functional subnetworks of schizophrenic and control subjects is expressed in a decrease of cognitive control exerted by higher-level areas over the behaviour of the temporal lobe and the hippocampal area, responsible for memory formation, in the patient group. This is proven by Bayesian model selection in Section 3.3.1, and effective connectivity comparison between patient and control groups, shown in Section 3.3.2. With such analyses, a question to answer is whether the inferred inter-group differences in connectivity reflect biological alterations, or the difference of some behavioural measure, such as learning performance. This question is addressed in Section 3.3.4, reassuring the relevance of the statistical results.

The intrinsic, neuronal network level connectivity of the prefrontal cortex is also analysed in Section 3.4, demonstrating that structural changes in the intra-areal network can account for functional differences is working memory dynamics. The coexistence of deficits on multiple levels of cortical hierarchy is possibly contributing to the symptoms of schizophrenia.

These findings are consistent with the results of earlier models described in the literature, like the theory explaining auditory hallucinations by a cognitive deficit in agency determination [146].

The determination of the functional macro-network and the deficit model of the intrinsic connectivity of the prefrontal cortex both support the area interaction model of associative learning [38], in which the prefrontal cortex implements a switching mechanism between encoding and recall modes of memory.

This part of my dissertation can be summarised as follows:

T 2/1. I discovered using dynamic causal models that in patients with schizophrenia, the information flow between prefrontal and hippocampal areas during learning, responsible for cognitive control, is significantly damaged.

T 2/2. I showed that dynamic causal models are able to grasp the physi-

ological differences caused by the illness independently from the task performance of the experimental subjects.

T 2/3. I showed, using a cellular network model of short term plasticity that the alteration of prefrontal synaptic dynamics, as observed in schizophrenia, drives the recall duration of working memory to a pathological regime.

These results were published in [12, 14, 68].

As the mapping of the task-dependent, effective connectivity patterns between cortical areas is far from complete, the most important application of the results described in the theses is the formation of new hypotheses and testing them on data with a similar methodology. For drug discovery applications, the results indicate a promising direction of development, as computational neuropharmacology should be on the rise soon [47, 96].

The results further the emergence of a diagnostic application, but are not sufficient for such themselves, as the use of the described alterations as biomarkers would require the integration of much more data and the implementation of meta-studies. The application of structural and functional networks in the diagnosis of psychiatric diseases is an emerging field, which is expected to yield significant results in the near future [29].

A

Appendix

A.1 Signal flow in directed networks

A.1.1 Statistical analysis of functional organisation

For sake of completeness in Table A.1 we complement Table 2.1 with further results of statistical analysis.

Table A.1: Networks coincide with those of Table 2.1. Shown are omitted entries, two numbers in a cell are the first two empirical moments.

net	VTc	stn	Rome	ER	bench.	kernel a.	stn a.	Rome a.
$n_{G,av}$	9	8	19	3.9 2.47	4.3 2.58	12	7	8
$p_{G,av}$	0.03	10^{-4}	10^{-4}	0.62 0.32	0.23 0.26	0.08	0.18	0.53
$n_{L,tot}$	9	15	14	4.64 2.99	5.14 3.02	10	5	23
$p_{L,tot}$	10^{-4}	10^{-4}	10^{-4}	0.61 0.28	0.10 0.21	0.14	0.04	0.93

Empirical distributions of CD-s and relative overlaps over the excitatory, inhibitory and neutral edge classes in the signal transduction network are shown in Figure A.1.

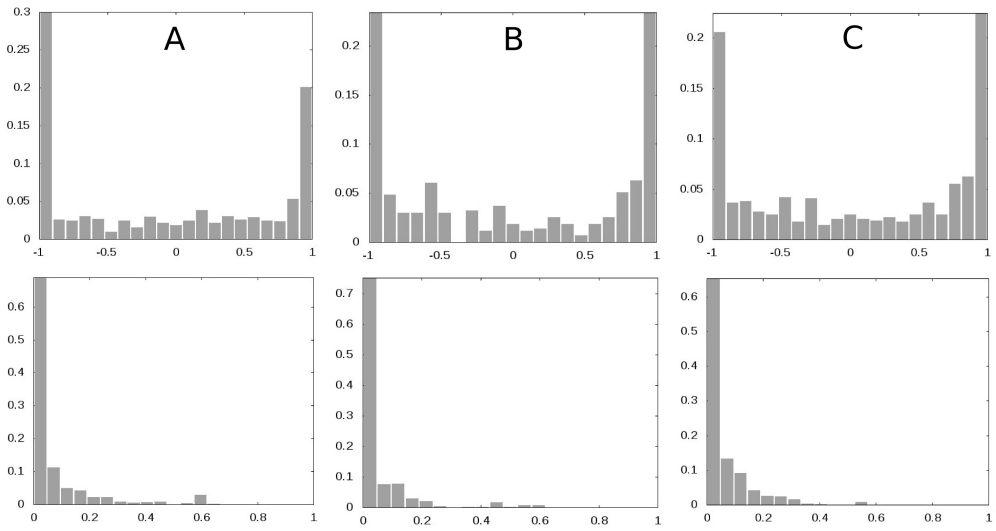


Figure A.1: Empirical distributions of CD-s (first row) and overlaps (second row) for the excitatory (column A), inhibitory (column B) and neutral (column C) edges of the neural signal transduction network.

A.1.2 CD calculation of Erdős-Rényi graphs

Calculation of CD-s for Erdős-Rényi graphs is straightforward, though lengthy. We note that the Erdős-Rényi graphs [46] we work with are *directed*. Furthermore for clarity we note that loop edges and multiple edges are prohibited. First we calculate the probability density function of CD_L, if number of nodes is n and the probability of having an edge between any two nodes is p. Let i denote the in-degree of the tail of the edge, let o denote the out-degree of the head of the same edge, and let l denote the number of nodes in the intersection of the first in-neighbours and out-neighbours of the tail and the head of the given edge. There are two essential terms in formulae below. The first is the one defining how large is the set of nodes we can choose our actual set from, the upper term in the binomial coefficients. The second one is the one defining which edges are prohibited to have the actual set size, the exponents in the $(1-p)$ terms. The exponent of the p terms and the lower terms of the binomial coefficients are simply the sizes of the node sets we choose. The probability of an edge tail having i predecessors is given with binomial density function:

$$p(i) = \begin{pmatrix} n-1 \\ i \end{pmatrix} p^i (1-p)^{n-1-i} \tag{A.1}$$

The probability of an edge head having o successors is given with Equation

(A.1), with i replaced with o.

The probability of having an intersection of the predecessors of the tail and the successors of the head of size l, given the size of the input and output sets, can be calculated as follows. First, if we assume that $i = o = l$, the probability p_l^* of having an overlap of size l is given as follows:

$$p^*(l) = \begin{pmatrix} n-1 \\ l \end{pmatrix} p^{2l}(1-p)^{2(n-1-l)} \tag{A.2}$$

We can take into account the non-overlapping parts of the input and output sets as follows, where the conditional probability of l given o (ranging from l to n) and i (ranging from l to $n-o$) is:

$$\begin{aligned} p(l|i,o) = p^*(l) \quad & \begin{pmatrix} n-1-l \\ o-l \end{pmatrix} p^{o-l}(1-p)^{n-1-o}. \\ \cdot \quad & \begin{pmatrix} n-1-o-l \\ i-l \end{pmatrix} p^{i-l}(1-p)^{n-1-o-i} \end{aligned} \tag{A.3}$$

Let $p(i,o,l)$ denote the joint probability density function of the variables i, o and l, it can be given as:

$$p(i,o,l) = p(l|i,o)p(i,o) = p(l|i,o)p(i)p(o) \tag{A.4}$$

We note that in Equation (A.4) i, o and l can be chosen independently, with l ranging from 0 to $\min(i,o)$. The value of CD_L is given as $(i-o)(i+o-l)^{-1}$. We perform the change of random variables

$$\psi(i,o,l) = (x,y,z), \quad x = \frac{i-o}{i+o-l}, \quad y = o, \quad z = l. \tag{A.5}$$

Changing the variables in the probability density function given with Equation (A.4) and calculating the marginal probability results in probability density function for CD_L:

$$p(x) = \sum_{y,z=1}^{n-1} p\left(\frac{x(z-y)-y}{x-1}, y, z\right) \frac{|z-y|}{(x+1)^2} \tag{A.6}$$

Similarly, to obtain p_O, the probability density function of the relative size of the overlapping set, one proceeds with the following change of variables:

$$\psi(i,o,l) = (x,y,z), \quad x = i, \quad y = o, \quad z = \frac{l}{i+o-l} \tag{A.7}$$

and ends up with the following the probability density function:

$$p_O(z) = \sum_{x,y=1}^{n-1} p\left(x, y, \frac{(x+y)z}{1+z}\right) \frac{x+y}{(z+1)^2} \tag{A.8}$$

Calculation of probability density function for CD_G is recursive. Nodes in the input set are organised into strata according to their distance from the edge head, the cardinalities of the strata being i_k, k ranging from 0 to $n-1$, thus the cardinality of the input set is given as:

$$i = \sum_{k=0}^{n-1} i_k \tag{A.9}$$

When calculating CD_G edges are allowed to the stratum i_{s-1} and all other shortcut edges from stratum i_s to lower strata are prohibited, including head and tail of the edge whose CD_G we are interested in. Loop edges are also prohibited. Strata in the output set are analogously denoted as o_s, meaning the s-th stratum in the output set. We bistratify the overlapping set, so its cardinality can be calculated in the following way:

$$l = \sum_{i \leq j} l_{i,j} \tag{A.10}$$

where l_{ij} denotes the overlap of the i-th stratum of the input set with the j-th stratum of the output set. We note that with probability 1 the cardinality of zeroth stratum in the input and output set is 1. Also, from the definition of zeroth strata it follows $l_{0,0} = 0$ with probability 1.
To shorten the subsequent formulae we use the following notation:

$$I_k = \sum_{r<k} i_r, \quad O_k = \sum_{r<k} o_r, \quad L_{a,b} = \sum_{r<a} \sum_{r \leq m < b} l_{r,m} \tag{A.11}$$

Probability of having i_s nodes in the s-th stratum is:

$$p(i_s | i_{s-1}, \ldots, i_0) = \sum_{a=i_s}^{n-1-I_s} \binom{n-1-I_s}{a} a \sum_{j=1}^{i_{s-1}} p^j (1-p)^{n-1+I_{s-1}} \tag{A.12}$$

We note the restriction on values i_s may have: $0 \leq i_s \leq n - I_s$. The conditional probability in Equation (A.12) was calculated according to the following lines.

The dummy variable a indicates the number of nodes at in-distance s from the tail of the chosen edge. The limit of the first summation is the same term as the upper expression in the binomial coefficient, represents the

number of available nodes to choose the m-th stratum from. The summation and multiplication by a before p^j accounts the fact that every node in the s-th stratum of the In-set can be attached to any number of nodes in the $s-1$-th stratum. The I_{s-1} term in the exponent of $p-1$ represents the prohibition of edges from the s-th stratum to the lower strata except for the one right below it. The complementary term for p^j would be $(1-p)^{n-1-j}$, but the $-j$ in the exponent is compensated by the prohibition of edges to the tail of the given edge from all points of the s-th stratum. All subsequent formulae are derived using similar reasoning.
According to the definition of the conditional probability, we have

$$p(i_s, \ldots, i_0) = p(i_s|i_{s-1}, \ldots i_0) \ldots p(i_1|i_0)p(i_0) \tag{A.13}$$

Probabilities of o_k-s are calculated analogously, with i replaced by o, and a replaced by b denoting the number of nodes at outdistance s from the head of the chosen edge.
Calculation of the conditional probability of having an overlap of size l is recursive. As nodes in the overlapping set share properties of the input and output sets, exponent of the $(1-p)$ term has to prohibit all shortcuts which are prohibited from both sets.
The analogue of Equation (A.2) is:

$$\begin{aligned} p^*(l_{s_1,s_2}|i_{s_1}, i_{s_1-1}, \ldots, i_0; o_{s_2}, o_{s_2-1} \ldots, o_0; l_{s_1-1,s_2}, \ldots, l_{0,0}) \quad = \\ \sum_{a=l_{s_1,s_2}}^{n-1-L_{s_1,s_2}} \sum_{b=l_{s_1,s_2}}^{n-1-L_{s_1,s_2}} \left(\begin{array}{c} n-1-L_{s_1,s_2} \\ a+b-l_{s_1,s_2} \end{array} \right) ab \cdot \\ \cdot \sum_{j_1=1}^{i_{s_1-1}} \sum_{j_2=1}^{o_{s_2-1}} p^{j_1 j_2} (1-p)^{n-1+I_{s_1-1}+O_{s_2-1}} \end{aligned} \tag{A.14}$$

Possible values of l_{s_1,s_2} in Equation (A.14) are restricted as follows: $0 \leq l_{s_1,s_2} \leq \min(i_{s_1}, o_{s_2})$. The conditional probability of having excess over the overlap in the output set is given as:

$$\begin{aligned} p^{\%}(l_{s_1,s_2}|i_{s_1}, \ldots, i_0; o_{s_2}, \ldots, o_0; l_{s_1-1,s_2}, \ldots, l_{0,0}) \quad = \\ \sum_{a=l_{s_1,s_2}}^{n-1-L_{s_1,s_2}-O_{s_2}} \left(\begin{array}{c} n-1-L_{s_1,s_2}-O_{s_2} \\ a \end{array} \right) a \cdot \\ \cdot \sum_{j=1}^{o_{s_2-1}} p^j (1-p)^{n-1+O_{s_2-1}} \end{aligned} \tag{A.15}$$

Analogously, the conditional probability of the input set being larger than the overlap is:

$$p^{\#}(l_{s_1,s_2}|i_{s_1},\ldots,i_0;o_{s_2},\ldots,o_0;l_{s_1-1,s_2},\ldots,l_{0,0}) \;=\; \sum_{b=l_{s_1,s_2}}^{n-1-L_{s_1,s_2}-O_{s_2}-I_{s_1}} \left(\begin{array}{c} n-1-L_{s_1,s_2}-O_{s_2}-I_{s_1} \\ b \end{array} \right) b \,\cdot \sum_{j=1}^{i_{s_1-1}} p^j(1-p)^{n-1+I_{s_1-1}} \tag{A.16}$$

The conditional probability of $l_{s_1 s_2}$ (global analogue of Equation (A.3)) is given as:

$$\begin{aligned} p(l_{s_1,s_2}|i_{s_1},\ldots,i_0;o_{s_2},\ldots,o_0;l_{s_1-1,s_2},\ldots,l_{0,0}) \;=\; & \\ & p^{*}(l_{s_1,s_2}|i_{s_1},\ldots,i_0;o_{s_2},\ldots,o_0;l_{s_1-1,s_2},\ldots,l_{0,0}) \\ & p^{\%}(l_{s_1,s_2}|i_{s_1},\ldots,i_0;o_{s_2},\ldots,o_0;l_{s_1-1,s_2},\ldots,l_{0,0}) \\ & p^{\#}(l_{s_1,s_2}|i_{s_1},\ldots,i_0;o_{s_2},\ldots,o_0;l_{s_1-1,s_2},\ldots,l_{0,0}) \end{aligned} \tag{A.17}$$

Thus, analogously to the Equation (A.4), using Equation (A.13) and its analogue for the output set, the joint probability of i_{s_1}, o_{s_2} and $l_{s_1 s_2}$ is:

$$\begin{aligned} p_J(i_{n-1},\ldots,i_0,o_{n-1},\ldots,o_0,l_{n-1,n-1},\ldots,l_{i_0,o_0}) \;=\; & \\ \prod_{k_1,k_2=0}^{n-1} p(l_{k_1,k_2}|i_{k_1},\ldots,i_0;o_{k_2},\ldots,o_0;l_{k_1-1,k_2-1},\ldots,l_{0,0}) & \end{aligned} \tag{A.18}$$

Based on Equations (A.18, A.9, A.10) one derives the marginal probability function $p_M(i,o,l)$ (which is the global analogue of Equation (A.4)), with $0 \le l \le \min(i,o)$ and $q = n-1$, $r = m-1$, $u_j = u_{j,j}$:

$$\begin{aligned} \mathbf{x} &= \{x_{s_0}, x_{s_1} - x_{s_0}, \ldots, i - x_{s_{n-1}}\} \\ \mathbf{y} &= \{y_{t_0}, y_{t_1} - y_{t_0}, \ldots, o - y_{t_{m-1}}\} \\ \mathbf{u} &= \{u_0, \ldots, l - u_q\} \\ p_M(i,o,l) &= \sum_{s_1=0,\ldots,s_q=0}^{q,\ldots,q} \sum_{t_1=0,\ldots,t_q=0}^{q,\ldots,q} \sum_{u_1=0,\ldots,u_q=0}^{s_1+t_1,\ldots,s_q+t_q} p_J(\mathbf{x},\mathbf{y},\mathbf{u}) \end{aligned} \tag{A.19}$$

then proceeds with the change of variables given in Equations (A.5), and calculates the marginal probability of x resulting in CD_G probability density

of the same form as the one given in Equation (A.6). p_O, the probability density function of the relative size of the overlapping set is calculated using the change of variables given in Equations (A.7), in $p_M(i, o, l)$. Finally, one obtains the probability density function of the same form as the one given in Equation (A.8).

A.2 Model-based analysis of functional disconnection in schizophrenia

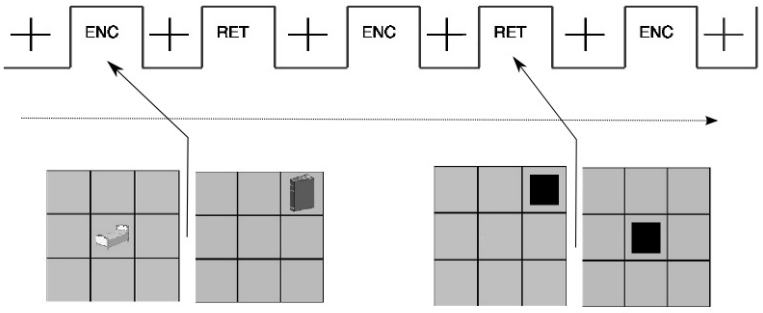


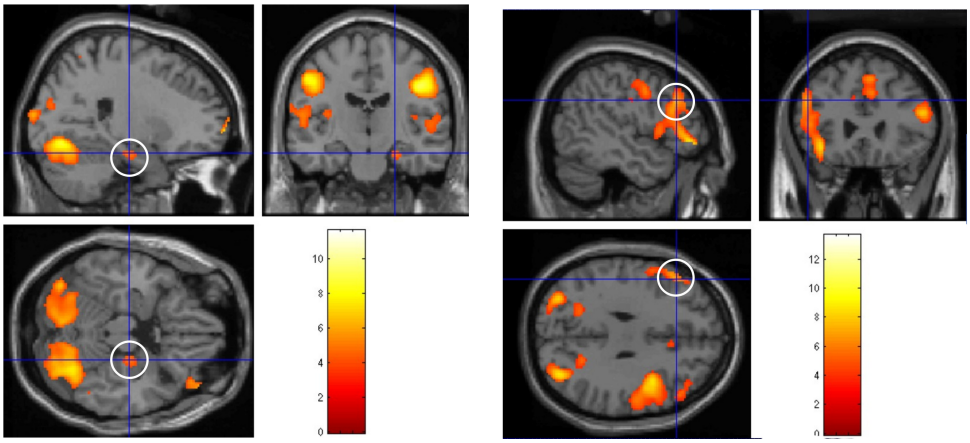
Figure A.2: Structure of the experimental paradigm is depicted with two examples of associations presented during encoding/consolidation ("bed" and "book") and examples of those locations cued during recall/retrieval.

A.2.1 Experimental subjects

Healthy Controls (n=11; mean age=22 yrs, sd=5; 5 females) and stable early course schizophrenia patients (n=11; mean age=26 yrs; sd=5; 3 females) gave informed consent. Groups did not differ in terms of age ($p > .10$). Patients were diagnosed using DSM-IV, SCID and consensus diagnosis. All were on a regimen of atypical anti-psychotics (Risperidone, Olanzapine or Aripiprazole).

A.2.2 Behavioral paradigm

Subjects alternated between blocks of encoding, rest/rehearsal and retrieval. During encoding, nine equi-familiar objects with monosyllabic object names [123] were presented in sequential random order (3s/object; 27 s block length) in grid locations for naming (e.g. "bed" and "book"). Following a rest



(a) Encoding phase (b) Retrieval phase

Figure A.3: Activated voxels are shown in light color in these exemplar slices from data volumes of the fMRI measurement. Picture courtesy of Vaibhav Diwadkar.

interval (R1; 27 s), memory for object-location pairs was tested using cued recall (3s/cue). The cycle ended with an additional rest interval (R2; 27 s). Subjects were instructed to respond by naming the object (or saying "no") between the end of the current and the beginning of the subsequent acquisition (1 s; that is when gradients were turned off). Object names were monosyllabic to minimize head motion. Eight blocks (each cycling between consolidation, rest and retrieval) were employed. The paradigm is illustrated in Figure A.2.

A.2.3 fMRI data acquisition

fMRI was conducted on a Bruker MedSpec 4T system with an 8-channel head coil. 288 T2*-weighted gradient-echo echo-planar images were acquired (TE=30ms; TR=3s; TA=2s; flip angle = 90°; acquisition matrix = 64 x 64 voxels; FOV = 240 mm; 24 slices; 3.75 x 3.75 x 4mm). During scanning, visual stimuli were presented via a projector system controlled by Presentation (www.neurobs.com).

fMRI data were preprocessed and analyzed in SPM2 using a standard processing sequence. Images were manually oriented to the AC-PC line. Following detrending and removal of low frequency components (.008 Hz), images were realigned to correct for head movement, spatially normalized to the MNI (Montreal Neurological Institute) template brain, resliced (2 mm3) and smoothed spatially by a Gaussian filter of 8mm full-width half maximum.

Activations acquired this way are depicted in Figure A.3. DCM's were based on first-level models with eight individual regressors (plus the six motion parameters used as regressors of no interest). Four regressors represented the unique effects of each of the epochs of interest, specifically encoding, R1, retrieval and R2. In addition, time dependent effects on each of these regressors were represented by convolving them with a linear time-ordered component. Time series for DCM analyses were extracted for each subject using a thresholded ($p < .05$) effects of interest contrast applied in each of the five regions of interest using stereotactic region-of-interest maps [93].

A.2.4 Model descriptions

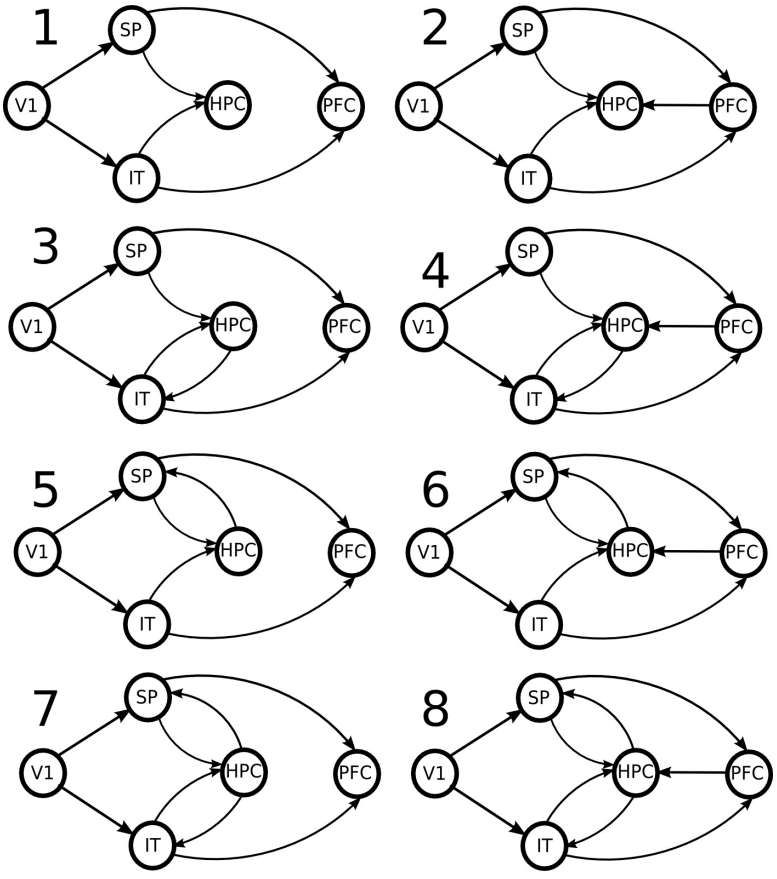


Figure A.4: First set of models with varying control stream connections. Modulatory effects of inputs were applied to all intrinsic connections present.

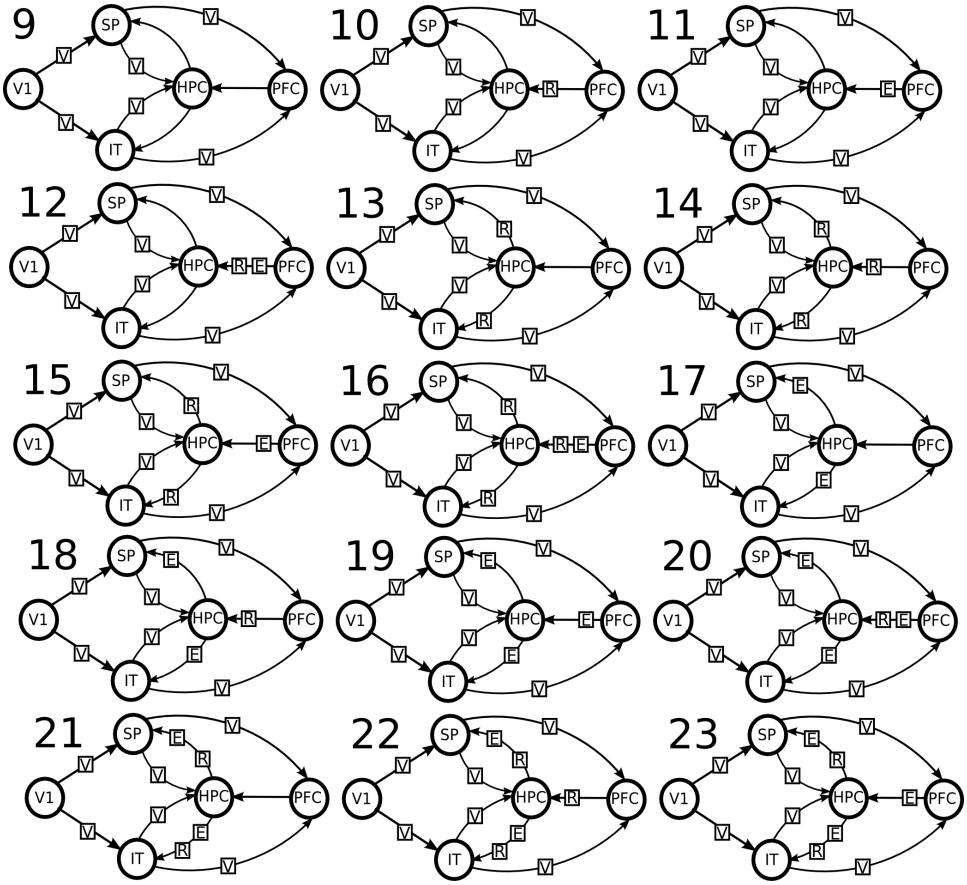


Figure A.5: Second set of models with varying modulatory effect of inputs on control stream connections. Model 24 is equivalent with model 8, and is fully depicted in Figure 3.2.

A.2.5 Bayesian model averaging

The Bayesian way to compare parameter values between two groups is to compute an averaged representative model for each, taking into account posterior model probabilities. The results of parameter comparisons based on Bayesian model averaging are shown in Figure A.6. Parameter matrices of the averaged model A, B and C were constructed by Bayesian averaging applying a fixed effects assumption [112].The significant alteration of the prefronto-hippocampal pathway is also detected by this method, strengthening the results of Section 3.3.2.

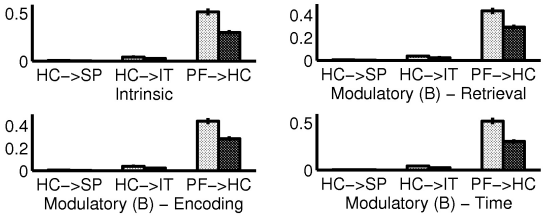


Figure A.6: Bayesian averaging of model parameters. The analysis shows weakened prefronto-hippocampal interaction, consistently with the analysis presented in Section 3.3.2.

A.3 List of abbreviations

Table A.2: All abbreviations used in this dissertation.

Abbreviation(s)	**Expansion(s)**
BAMS	Brain Architecture Management System
BOLD	blood oxygen level dependent (signal)
cAMP	cyclic adenosin monophosphate
CD	convergence degree
CoCoMac	Collations of Connectivity data on the Macaque brain
DCM	Dynamic Causal Modelling, dynamic causal model
DTI	diffusion tensor imaging
DSI	diffusion spectrum imaging
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, Edition IV.
E, ENC	Encoding (experimental condition)
EEG	electroencephalography
ER	Erdős-Rényi (graph)
fMRI	functional magnetic resonance imaging
FOV	field of view (of MRI)
HC	healthy control (subject group)
HF	hierarchical flow
Jacc	Jaccard coefficient
KS	Kolmogorov-Smirnov (statistical test)
MAP	Maximum a posteriori (estimation)
MEG	magnetoencephalography
MNI	Montreal Neurological Institute (coordinates)
MR, MRI	magnetic resonance imaging

Continued on next page

Table A.2 – continued from previous page

Abbreviation(s)	Expansion(s)
NCS-1	neuronal calcium sensor 1
Ovl	overlap
PDF	probability density function
PET	positron emission tomography
PS	population spike
PSW	preferential small world (graph)
R, RET	Recall (experimental condition)
R1, R2	resting period (in experimental paradigm)
ROI	Region of Interest
SCID	Structured Clinical Interview for DSM-IV
SCZ	schizophrenia (subject group)
SP	shortest path
SPM	Statistical Parametric Mapping (software package)
stn	signal transduction network (of proteins in a cell)
T	Tesla (magnetic flux density)
T2*	weighting method for MR image acquisition
TR	acquisition time (of MRI)
TE	echo time (of MRI)
TR	repetition time (of MRI)
V	Visual (experimental condition)
VTc	visuo-tactile cortex (of the macaque)
WM	working memory
WS	Watts-Strogatz (graph)

Table A.3: Names of all abbreviated cortical areas mentioned in the dissertation. For a more detailed description see [49].

Abbreviation(s)	Name(s)
AITv	Anterior inferotemporal (ventral)
AITd	Anterior inferotemporal (dorsal)
CA1	Cornu Ammonis area 1 of the hippocampus
CA3	Cornu Ammonis area 3 of the hippocampus
CITv	Central inferotemporal (ventral)
CITd	Central inferotemporal (dorsal)
DP	Dorsal prelunate
FST	Floor of superior temporal
FEF	Frontal eye field
HPC, HC	Hippocampus
Id	Insular disgranular
Ig	Insular granular
IT	Inferior temporal
MSTd	Medial superior temporal (dorsal)
MSTI	Medial superior temporal (lateral)
LIP	Lateral intraparietal
PFC, PF	Prefrontal (dorsolateral)
PIP	Posterior intraparietal
PITv	Posterior inferotemporal (ventral)
PITd	Posterior inferotemporal (dorsal)
PO	Parieio-occipital
Ri	Retroinsular
SII	Sencondary somatosensory
SMA	Supplementary motor area
SP	Superior parietal
STPa	Superior temporal polysensory (anterior)
STPp	Superior temporal polysensory (posterior)
TF	Temporofrontal
TH	Temporal hippocampal
V1,V2,V3,V3a,V4	Visual areas 1, 2, 3, 3a, 4
V4t	V4 transitional
VIP	Ventral intraparietal
VOT	Ventral occipitotemporal
VP	Ventral posterior

Table A.4: Names of all numbered Brodmann areas mentioned in the dissertation (only by the number or preceded by an 'A').

Number	Name(s)
1, 2, 3a, 3b	Primary somatosensory
4	Primary motor
5, 7a, 7b	Somatosensory association
6	Premotor
8	Frontal eye field
9, 46	Dorsolateral prefrontal
10	Anterior prefrontal
11, 12	Orbitofrontal
13, 14	Insular
24	Ventral anterior cingulate
32	Dorsal anterior cingulate
35	Perirhinal
36	Ectorhinal
45	Pars triangularis Broca's area
47	pars orbitalis, inferior frontal gyrus

Bibliography

[1] L. F. Abbott. Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain Research Bulletin*, 50:303–4, 1999.

[2] R. Albert and A. L. Barabási. Statistical mechanics of complex networks. *Reviews in Modern Physics*, 74:47–94, 2002.

[3] U. Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC Press, 2006.

[4] U. Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8:450–61, 2007.

[5] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America*, 97:11149–52, 2000.

[6] E. Amaro Jr. and G. J. Barker. Study design in fmri: Basic principles. *Brain and Cognition*, 60:220–32, 2006.

[7] F. J. Anscombe. Graphs in statistical analysis. *American Statistician*, 27:17–21, 2004.

[8] M. A. Arbib, P. Érdi, and J. Szentágothai. *Neural Organization: Structure, Function, and Dynamics*. MIT Press, 1998.

[9] J. P. Bagrow, E. M. Bollt, J. D. Skufca, and D. ben Avraham. Portraits of complex networks. *Europhysics Letters*, 81:68004, 2008.

[10] M. Bányai. Az agyműködés hálózatai. *Természet Világa*, 143:505–7, 2012.

[11] M. Bányai, T. Nepusz, L. Négyessy, and F. Bazsó. Convergence properties of some random networks. In *Proceedings of the 7th International Symposium on Intelligent Systems and Informatics*, pages 241–5, Subotica, Serbia, 2009.

[12] M. Bányai, V. A. Diwadkar, and P. Érdi. Model-based dynamical analysis of functional disconnection in schizophrenia. *NeuroImage*, 58: 870–7, 2011.

[13] M. Bányai, L. Négyessy, and F. Bazsó. Organisation of signal flow in directed networks. *Journal of Statistical Mechanics: theory and experiment*, 6:06001, 2011.

[14] M. Bányai, B. Ujfalussy, V. A. Diwadkar, and P. Érdi. Impairments in the prefronto-hippocampal interactions explain associative learning deficit in schizophrenia. *BMC Neuroscience*, 12:93, 2011.

[15] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–12, 1999.

[16] G. Barbalat, V. Chambon, N. Franck, E. Koechlin, and C. Farrer. Organization of cognitive control within the lateral prefrontal cortex in schizophrenia. *Archives of General Psychiatry*, 66:377–86, 2009.

[17] H. B. Barlow. Possible principles underlying the transformations of sensory messages. In W. Rosenblith, editor, *Sensory Communication*, pages 217–34. MIT Press, Cambridge, MA, 1961.

[18] A. M. Bastos, M. Usrey, R. A. Adams, G. R. Mangun, P. Fries, and K. J. Friston. Canonical microcircuits for predictive coding. *Neuron*, 76:695–711, 2012.

[19] E. Ben-Naim and P. L. Krapivsky. Stratification in the preferential attachment network. *Journal of Physics A: Mathematical and Theoretical*, 42:475001, 2009.

[20] C. M. Bennett, M. B. Miller, and G. L. Wolford. Neural correlates of interspecies perspective taking in the post-mortem atlantic salmon: an argument for multiple comparisons correction. *NeuroImage*, 47:S125, 2009.

[21] A. Bigdeli, A. Tizghadam, and A. Leon-Garcia. Comparison of network criticality, algebraic connectivity, and other graph metrics. In *First Workshop on Simplifying Complex Networks: SIMPLEX*, pages 241–5, Venice, Italy, 2009.

[22] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2007.

[23] R. S. Blumenfeld, C. M. Parks, A. P. Yonelinas, and C. Ranganath. Putting the pieces together: The role of dorsolateral prefrontal cortex in relational memory encoding. *Journal of Cognitive Neuroscience*, 23: 257–65, 2010.

[24] B. Bollobás, R. Kozma, and D. Miklós, editors. *Handbook of Large-Scale Random Networks*. Bolyai Society Mathematical Studies. Springer-Verlag, 2009.

[25] M. Bota, H. W. Dong, and L. W. Swanson. Brain architecture management system. *Neuroinformatics*, 3:15–48, 2005.

[26] K. Brodmann. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. J.A. Barth, 1909.

[27] C. Buchel, J. T. Coull, and K. J. Friston. The predictive value of changes in effective connectivity for human learning. *Science*, 283: 1538–41, 1999.

[28] J. W. Buckholtz and A. Meyer-Lindenberg. Psychopathology and the human connectome: Toward a transdiagnostic model of risk for mental illness. *Neuron*, 74:990–1004, 2012.

[29] E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10:186–98, 2009.

[30] E. T. Bullmore, S. Rabe-Hesketh, R. G. Morris, S. C. Williams, L. Gregory, J. A. Gray, and M. J. Brammer. Functional magnetic resonance image analysis of a large-scale neurocognitive network. *NeuroImage*, 4: 16–33, 1996.

[31] S. A. Bunge, B. Burrows, and A. D. Wagner. Prefrontal and hippocampal contributions to visual associative recognition: interactions between cognitive control and episodic retrieval. *Brain Cognition*, 56: 141–52, 2004.

[32] R. Cabeza and L. Nyberg. Neural bases of learning and memory: functional neuroimaging evidence. *Current opinion in neurobiology*, 13: 415–21, 2000.

[33] V. D. Calhoun, T. Eichele, and G. Pearlson. Functional brain networks in schizophrenia: A review. *Frontiers in Neuroscience*, 3:1–12, 2009.

[34] M. Catani and M. Mesulam. What is a disconnection syndrome? *Cortex*, 44:911–3, 2008.

[35] V. Chambon, N. Franck, E. Koechlin, E. Fakra, G. Ciuperca, J. M. Azorin, and C. Farrer. The architecture of cognitive control in schizophrenia. *Brain*, 131:962–70, 2008.

[36] J. DeFelipe and E. G. Jones. Santiago Ramón y Cajal and methods in neurohistology. *Trends in Neurosciences*, 15:237–46, 1992.

[37] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39:1–38, 1977.

[38] V. A. Diwadkar, B. Flaugher, T. Jones, L. Zalányi, B. Ujfalussy, M. S. Keshavan, and P. Érdi. Impaired associative learning in schizophrenia: Behavioral and computational studies. *Cognitive Neurodynamics*, 2: 207–19, 2008.

[39] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Giant strongly connected component of directed networks. *Physical Review E*, 64:025101(R), 2001.

[40] C. Economo and G. N. Koskinas. *Die cytoarchitektonik der hirnrinde des erwachsenen menschen*. Berlin, 1925.

[41] B. T. Edwards BG, Barch DM. Improving prefrontal cortex function in schizophrenia through focused training of cognitive control. *Front Hum Neurosci*, 26:4–32, 2010.

[42] H. Eichenbaum. The long and winding road to memory consolidation. *Nature neuroscience*, 4:1057–8, 2001.

[43] H. Eichenbaum. Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron*, 44:109–20, 2004.

[44] D. P. Eisenberg and K. F. Berman. Executive function, neural circuitry, and genetic mechanisms in schizophrenia. *Neuropsychopharmacology*, 35:258–77, 2010.

[45] I. Ellison-Wright and E. Bullmore. Meta-analysis of diffusion tensor imaging studies in schizophrenia. *Schizophrenia Research*, 108:3–10, 2009.

[46] P. Erdős and A. Rényi. On random graphs i. *Pubclicationes Mathematicae Debrecen*, 5:290–7, 1959.

[47] P. Érdi, B. Ujfalussy, L. Zalányi, and V. A. Diwadkar. Computational approach to schizophrenia: Disconnection syndrome and dynamical pharmacology. In L. M. Ricciardi, A. Buonocore, and E. Pirozzi, editors, *A selection of papers of The BIOCOMP 2007 International Conference*, volume 1028 of *Proceedings of the American Institute of Physics*, pages 65–87, 2008.

[48] P. Érdi, B. Ujfalussy, and V. A. Diwadkar. The schizophrenic brain: A broken hermeneutic circle. *Neural Network World*, 19:413–27, 2009.

[49] D. J. Felleman and D. C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1:1–47, 1991.

[50] P. Fletcher, C. Buchel, O. Josephs, K. J. Friston, and R. Dolan. Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. *Cerebral Cortex*, 9:168–78, 1999.

[51] R. S. J. Frackowiak, J. T. Ashburner, W. D. Penny, and S. Zeki. *Human Brain Function*. Academic Press, 2. edition, 2004.

[52] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40:35–41, 1977.

[53] W. Freeman and J. W. Watts. Prefrontal lobotomy in the treatment of mental disorders. *Southern Medical Journal*, 30:23–31, 1937.

[54] W. J. Freeman. *Mass action in the nervous system*. Academic Press, 1975.

[55] K. J. Friston. The disconnection hypothesis. *Schizophrenia Research*, 30:115–25, 1998.

[56] K. J. Friston. Dysfunctional connectivity in schizophrenia. *American Journal of Psychiatry*, 162:429–32, 2002.

[57] K. J. Friston. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2:56–78, 2004.

[58] K. J. Friston. Functional and effective connectivity: A review. *Brain Connectivity*, 1:13–36, 2011.

[59] K. J. Friston and C. D. Frith. Schizophrenia: a disconnection syndrome? *Clinical Neuroscience*, 3:89–97, 1995.

[60] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price. Nonlinear responses in fmri: The balloon model, volterra kernels and other hemodynamics. *NeuroImage*, 12:466–77, 2000.

[61] K. J. Friston, L. Harrison, and W. D. Penny. Dynamic causal modelling. *NeuroImage*, 19:1273–302, 2003.

[62] K. J. Friston, J. Ashburner, S. J. Kiebel, T. E. Nichols, and W. D. Penny, editors. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 2007.

[63] K. J. Friston, B. Li, J. Daunizeau, and K. E. Stephan. Network discovery with DCM. *NeuroImage*, 56:1202–21, 2011.

[64] N. Geschwind. Disconnection syndromes in animals and man i. *Brain*, 88:237–94, 1965.

[65] N. Geschwind. Disconnection syndromes in animals and man ii. *Brain*, 88:585–644, 1965.

[66] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99:7821–6, 2002.

[67] P. S. Goldman-Rakic and L. D. Selemon. Functional and anatomical aspects of prefrontal pathology in schizophrenia. *Schizophrenia Bulletin*, 23:437–58, 1997.

[68] C. D. Gore, M. Bányai, P. M. Grey, V. A. Diwadkar, and P. Érdi. Pathological effects of cortical architecture on working memory in schizophrenia. *Pharmacopsychiatry*, 43:592–7, 2010.

[69] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns. Mapping the structural core of human cerebral cortex. *PLoS Biology*, 6:e159, 2008.

[70] P. J. Harrison. The neuropathology of schizophrenia. a critical review of the data and their interpretation. *Brain*, 122:593–624, 1999.

[71] P. J. Harrison and D. R. Weinberger. Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Molecular psychiatry*, 10:40–68, 2005.

[72] J. V. Haxby, C. L. Grady, B. Horwitz, L. G. Ungerleider, M. Mishkin, R. E. Carson, P. Herscovitch, M. B. Schapiro, and S. I. Rapoport. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 88:1621–5, 1991.

[73] H. Hirase, A. Czurkó, J. Csicsvári, and G. Buzsáki. Firing rate and theta-phase coding by hippocampal pyramidal neurons during 'space clamping'. *European Journal of Neuroscience*, 11:4373–80, 2008.

[74] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117:500–44, 1952.

[75] J. J. Hopfield and A. V. M. Herz. Rapid local synchronization of action potentials: Toward computation with coupled integrate-and-fire neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 92:6655–62, 1995.

[76] P. J. Ingram, M. P. H. Stumpf, and J. Stark. Network motifs: structure does not determine function. *BMC Genomics*, 7:108, 2006.

[77] Y. Iturria-Medina, R. C. Sotero, E. J. Canales-Rodriguez, Y. Aleman-Gomez, and L. Melie-Garcia. Studying the human brain anatomical network via diffusion-weighted mri and graph theory. *NeuroImage*, 40: 1064–76, 2008.

[78] E. M. Izhikevich and G. M. Edelman. Large-scale model of mammalian thalamocortical systems. *Proceedings of the National Academy of Sciences of the United States of America*, 105:3593–8, 2008.

[79] P. Jaccard. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37:547–9, 1901.

[80] C. Koch and I. Segev, editors. *Methods in Neuronal Modeling: From Ions to Networks*. MIT Press, 1998.

[81] P. O. Koh, A. S. Undie, N. Kabbani, R. Levenson, P. S. Goldman-Rakic, and M. S. Lidow. Up-regulation of neuronal calcium sensor-1 (ncs-1) in the prefrontal cortex of schizophrenic and bipolar patients. *Proceedings of the National Academy of Sciences of the United States of America*, 100:313–7, 2003.

[82] A. Lancichinetti and S. Fortunato. Community detection algorithms: A comparative analysis. *Physical Review E*, 80:056117, 2009.

[83] J. L. Lanciego and F. G. Wouterlood. A half century of experimental neuroanatomical tracing. *Journal of Chemical Neuroanatomy*, 42:157–83, 2011.

[84] L. Lapicque. Recherches quantitatives sur l'excitation électrique des nerfs traiteé comme une polarization. *Journal de physiologie et de pathologie générale*, 9:620–35, 1907.

[85] P. Lavenex and D. Amaral. Hippocampal-neocortical interaction: a hierarchy of associativity. *Hippocampus*, 10:420–30, 2000.

[86] S. M. Lawrie, C. Buechel, H. C. Whalley, C. D. Frith, K. J. Friston, and E. C. Johnstone. Reduced frontotemporal functional connectivity in schizophrenia associated with auditory hallucinations. *Biological Psychiatry*, 51:1008–11, 2002.

[87] T. B. Leergaard, C. C. Hilgetag, and O. Sporns. Mapping the connectome: Multi-level analysis of brain connectivity. *Frontiers in Neuroinformatics*, 6:14, 2012.

[88] N. K. Logothetis and B. A. Wandell. Interpreting the BOLD signal. *Annual Review of Physiology*, 66:735–69, 2004.

[89] L. Lovász. Discrete analytic functions: An exposition. *Surveys in differential geometry*, 9:241–73, 2004.

[90] J. Luo and C. L. Magee. Detecting evolving patterns of self-organizing networks by flow hierarchy measurement. *Complexity*, 16:53–61, 2011.

[91] M. E. Lynall, D. S. Bassett, R. Kerwin, P. J. McKenna, M. Kitzbichler, U. Muller, and E. Bullmore. Functional connectivity and brain networks in schizophrenia. *Journal of Neuroscience*, 14:9477–87, 2010.

[92] A. Ma'ayan, S. L. Jenkins, S. Neves, A. Hasseldine, E. Grace, B. Dubin-Thaler, N. J. Eungdamrong, G. Weng, P. T. Ram, J. J. Rice, A. K. G. A. Stolovitzky, R. D. Blitzer, and R. Iyengar. Formation of regulatory patterns during signal propagation in a mammalian cellular network. *Science*, 309:1078–83, 2005.

[93] J. A. Maldjian, P. Laurienti, R. A. Kraft, and J. H. Burdette. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19:1233–9, 2003.

[94] W. McCulloch and W. Pitts. Logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 7:115–33, 1943.

[95] E. K. Miller and J. D. Cohen. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24:176–202, 2001.

[96] G. Miller. Is pharma running out of brainy ideas? *Science*, 329:502–4, 2010.

[97] M. Mishkin and L. G. Ungerleider. Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural brain research*, 6:57–77, 1982.

[98] M. Mishkin, L. G. Ungerleider, and K. A. Macko. Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, 6:414–7, 1983.

[99] G. Mongillo, O. Barak, and M. Tsodyks. Synaptic theory of working memory. *Science*, 319:1543–6, 2008.

[100] V. B. Mountcastle. The columnar organization of the neocortex. *Brain*, 120:701–22, 1997.

[101] L. Négyessy and P. S. Goldman-Rakic. Subcellular localization of the dopamine d2 receptor and coexistence with the calcium-binding protein neuronal calcium sensor-1 in the primate prefrontal cortex. *Journal of Comparative Neurology*, 488:464–75, 2005.

[102] L. Négyessy, T. Nepusz, L. Kocsis, and F. Bazsó. Prediction of the main cortical areas and connections involved in the tactile function of the visual cortex by network analysis. *European Journal of Neuroscience*, 23:1919–30, 2006.

[103] L. Négyessy, T. Nepusz, L. Zalányi, and F. Bazsó. Convergence and divergence are mostly reciprocated properties of the connections in the network of cortical areas. *Proceedings of the Royal Society B*, 275: 2403–10, 2008.

[104] L. Négyessy, M. Bányai, and F. Bazsó. What makes the prefrontal cortex so appealing in the era of brain imaging? A network analytical perspective. *Acta Biologica Hungarica*, 63:38–53, 2012.

[105] L. Négyessy, J. Minich, D. Meszéna, A. Buzás, B. Jákli, M. Bányai, E. Procyk, P. Barone, and F. Bazsó. From neuronal communication

to the flow of information in the cerebral cortex. In *9th Conference Digital Speech and Image Processing*, Kovacica, Serbia, 2012.

[106] M. E. J. Newman. Models of the small world, a review. *Journal of Statistical Physics*, 101:819–41, 2000.

[107] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[108] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64:026118, 2001.

[109] S. Nieuwenhuis, B. U. Forstmann, and E.-J. Wagenmakers. Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature Neuroscience*, 14:1105–7, 2011.

[110] R. C. O'Reilly and M. J. Frank. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural computation*, 18:283–328, 2006.

[111] J. Pearl. Bayesianism and causality, or, why I am only a half-Bayesian. In D. Corfield and J. Williamson, editors, *Foundations of Bayesianism*, volume 24 of *Kluwer Applied Logic Series*, pages 19–34. Kluwer Academic Publishers, 2001.

[112] W. D. Penny, K. E. Stephan, J. Daunizeau, M. J. Rosa, K. J. Friston, T. M. Schofield, and A. P. Leff. Comparing families of dynamic causal models. *PLoS Computational Biology*, 6:e1000709, 2010.

[113] P. Pons and M. Latapy. Computing communities in large networks using random walks. In P. Yolum, T. Güngör, F. Gürgen, and C. Özturan, editors, *Computer and Information Sciences - ISCIS 2005*, volume 3733 of *Lecture Notes in Computer Science*, pages 284–293, 2005.

[114] J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar, and S. E. Petersen. Functional network organization of the human brain. *Neuron*, 72:665–78, 2011.

[115] J. D. Ragland, A. R. Laird, C. Ranganath, R. S. Blumenfeld, S. M. Gonzales, and D. C. Glahn. Prefrontal activation deficits during episodic memory in schizophrenia. *American Journal of Psychiatry*, 166:863–74, 2009.

[116] R. P. N. Rao, B. A. Olshausen, and M. S. Lewicki, editors. *Probabilistic Models of the Brain: Perception and Neural Function*. MIT Press, 2002.

[117] K. S. Rockland. Five points on columns. *Frontiers in Neuroanatomy*, 4:22, 2010.

[118] R. F. Schwarzlose, J. D. Swisher, S. Dang, and N. Kanwisher. The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 105:4447–52, 2008.

[119] P. J. Siekmeier, M. E. Hasselmo, M. W. Howard, and J. Coyle. Modeling of context-dependent retrieval in hippocampal region CA1: implications for cognitive function in schizophrenia. *Schizophrenia Research*, 89:177–90, 2007.

[120] J. S. Simons and H. J. Spiers. Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, 4: 637–48, 2003.

[121] T. Sippy, A. Cruz-Martín, A. Jeromin, and F. E. Schweizer. Acute changes in short-term plasticity at synapses with elevated levels of neuronal calcium sensor-1. *Nature Neuroscience*, 6:1031–38, 2003.

[122] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich. Network modelling methods for FMRI. *NeuroImage*, 54:875–91, 2011.

[123] J. G. Snodgrass and M. Vanderwart. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6:174–215, 1980.

[124] O. Sporns. Graph theory methods for the analysis of neural connectivity patterns. In R. Kötter, editor, *Neuroscience databases. A practical guide*, pages 171–186. Kluwer, Boston, MA, 2002.

[125] O. Sporns. *Networks of the Brain*. MIT Press, 2011.

[126] O. Sporns, G. Tononi, and R. Kötter. The human connectome: A structural description of the human brain. *PLoS Computational Biology*, 1: e42, 2005.

[127] O. Sporns, C. J. Honey, and R. Kötter. Identification and classification of hubs in brain networks. *PLoS ONE*, 2:e1049, 2007.

[128] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420:190–3, 2002.

[129] K. E. Stephan. Dynamic causal models of neural system dynamics: Current state and future extensions. *Journal of Biosciences*, 32:129–44, 2007.

[130] K. E. Stephan, L. Kamper, A. Bozkurt, G. A. P. C. Burns, M. P. Young, and R. Kötter. Advanced database methodology for the collation of connectivity data on the macaque brain (cocomac). *Philosophical Transactions of the Royal Socitey B*, 356:1159–86, 2001.

[131] K. E. Stephan, T. Baldeweg, and K. J. Friston. Synaptic plasticity and dysconnection in schizophrenia. *Biological Psychiatry*, 59:929–39, 2006.

[132] K. E. Stephan, K. J. Friston, and C. D. Frith. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophrenia Bulletin*, 35:509–27, 2009.

[133] K. E. Stephan, W. D. Penny, J. Daunizeau, R. J. Moran, and K. J. Friston. Bayesian model selection for group studies. *NeuroImage*, 46: 1004–17, 2009.

[134] K. E. Stephan, W. D. Penny, R. J. Moran, H. E. M. den Ouden, J. Daunizeau, and K. J. Friston. Ten simple rules for dynamic causal modeling. *NeuroImage*, 49:3099–109, 2010.

[135] M. Strotzer. One century of brain mapping using Brodmann areas. *Clinical Neuroradiology*, 19:179–86, 2009.

[136] H. Y. Tan, J. H. Callicott, and D. R. Weinberger. Dysfunctional and compensatory prefrontal cortical systems, genes and the pathogenesis of schizophrenia. *Cerebral Cortex*, 17:171–81, 2007.

[137] L. C. Triarhou. Georg N. Koskinas (1885–1975) and his scientific contributions to the normal and pathological anatomy of the human brain. *Brain Research Bulletin*, 68:121–39, 2005.

[138] M. P. van den Heuvel and H. E. Hulshoff Pol. Exploring the brain network: A review on resting-state fmri functional connectivity. *European Neuropsychopharmacology*, 20:519–34, 2010.

[139] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. E. J. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. W. Curtiss, S. Della Penna, D. Feinberg, M. F. Glasser, N. Harel, A. C. Heath, L. Larson-Priork, D. Marcus, G. Michalareas, S. Moeller, R. Oostenveld, S. E. Petersen, F. Prior, B. L. Schlaggar, S. M. Smith, A. Z. Snyder, J. Xu, E. Yacoub, and WU-Minn HCP Consortium. The human connectome project: A data acquisition perspective. *NeuroImage*, 62:2222–31, 2012.

[140] R. VanRullen, R. Guyonneau, and S. J. Thorpe. Spike times make sense. *Trends in Neurosciences*, 28:1–4, 2005.

[141] A. Vercammen, H. Knegtering, J. A. den Boer, E. J. Liemburg, and A. Aleman. Auditory hallucinations in schizophrenia are associated with reduced functional connectivity of the temporo-parietal area. *Biological Psychiatry*, 67:912–8, 2010.

[142] E. S. Vizi and E. Lábos. Non-synaptic interactions at presynaptic level. *Progress in Neurobiology*, 37:145–63, 1991.

[143] E. S. Vizi, J. P. Kiss, and B. Lendvai. Nonsynaptic communication in the central nervous system. *Neurochemistry International*, 45:443–51, 2004.

[144] E. Vul, C. Harris, P. Winkielman, and H. Pashler. Voodoo correlations in social neuroscience. *Perspectives on Psychological Science*, 4:274–90, 2009.

[145] S. Wasserman and K. Faust. *Social network analysis: Methods and applications*. Cambridge University Press, 1994.

[146] F. Waters, T. Woodward, P. Allen, A. Aleman, and I. Sommer. Self-recognition deficits in schizophrenia patients with auditory hallucinations: A meta-analysis of the literature. *Schizophrenia Bulletin*, 38: 741–50, 2012.

[147] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:740–2, 1998.

[148] V. J. Wedeen, P. Hagmann, W. Y. Tseng, T. G. Reese, and R. M. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54:1377–86, 2005.

[149] D. R. Weinberger, K. F. Berman, and R. F. Zec. Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia: I. regional cerebral blood flow evidence. *Archives of General Psychiatry*, 43:114–24, 1986.

[150] D. R. Weinberger, K. F. Berman, R. Suddath, and E. F. Torrey. Evidence of dysfunction of a prefrontal-limbic network in schizophrenia: a magnetic resonance imaging and regional cerebral blood flow study of discordant monozygotic twins. *American Journal of Psychiatry*, 149: 890–7, 1992.

[151] C. Wernicke. *Grundriss der Psychiatrie in klinischen Vorlesungen.* Thieme, 1906.

[152] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner. The structure of the nervous system of the nematode caenorhabditis elegans. *Philosophical Transactions of the Royal Socitey B*, 314:1–340, 1986.

[153] E. P. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics*, 13:1–14, 1960.

[154] S. Wirth, M. Yanike, L. M. Frank, A. C. Smith, E. N. Brown, and W. A. Suzuki. Single neurons in the monkey hippocampus and learning of new associations. *Science*, 300:1578–81, 2003.

[155] S. J. Wood, T. Proffitt, K. Mahony, D. J. Smith, J. A. Buchanan, W. Brewer, G. W. Stuart, D. Velakoulis, P. D. McGorry, and C. Pantelis. Visuospatial memory and learning in first-episode schizophreniform psychosis and established schizophrenia: a functional correlate of hippocampal pathology? *Psychological Medicine*, 32:429–38, 2002.

[156] S. R. y Cajal. *Histologie du systéme nerveux de l'homme et des vertebras.* Maloine, 1909.