

Neural correlates of consciousness in the complexity  
of brain networks



**Ioannis Pappas**

Downing College

Department of Clinical Neurosciences, University of Cambridge

This dissertation is submitted for the degree of

*Doctor of Philosophy*

November 2018

---

*Dedicated to my loving family...*

---

## Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed 60,000-word limit for the Clinical Medicine and Clinical Veterinary Medicine Degree Committee.

Part of the study outlined in chapter 4 has been accepted in a peer-reviewed journal:

Pappas I, Adapa RM, Menon DK, Stamatakis EA (2018) Brain network disintegration during sedation is mediated by the complexity of sparsely connected regions.

*Neuroimage*. (<http://doi.org/10.1016/j.neuroimage.2018.10.078>).

Ioannis Pappas,

November 2018, Cambridge, UK

## Acknowledgements

I would like to acknowledge all those who helped me make this thesis.

First and foremost, I would like to say thank you (ευχαριστώ) to my supervisor Dr Emmanuel Stamatakis. His endless support, motivation and hard work shaped my scientific method and guided me safely to accomplish this thesis. One cannot ask for someone with more dedication and generosity.

I am also grateful to my secondary supervisor Prof David Menon for his insightful support and the warm research environment at the Division of Anaesthesia. Moreover, I would like to thank Dr Adrian Carpenter and Dr Ruma Raha-Chowdhury from the Department of Clinical Neurosciences for allowing me to follow this PhD path.

In addition, I would like to thank all of my friends and colleagues at the Division of Anaesthesia for all their support and the good friendship. A special thank you to Mike Craig for his collaborative spirit and friendship.

Finally, I would like to thank my family. None of this would have been possible without their support.

## Abstract

How do we define consciousness? Besides philosophical endeavours, the development of modern neuroimaging techniques fostered a principled way of quantifying the neural correlates of consciousness. Acquiring and analysing resting-state functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) data, has allowed neuroscientists to noninvasively map the brain's functional interactions (or functional connectivity). Based on data obtained during controlled loss of consciousness and in cases of patients with disorders of consciousness, it has now been suggested that multiple, functionally specialized/segregated areas need to interact and integrate information in order to support consciousness. Thus an emerging idea in neuroscience is that the brain needs to balance the coexistence of functional segregation and integration, a property often termed as brain complexity, in order to produce consciousness. A resulting hypothesis is that consciousness is abolished when the balance between segregation and integration is lost and brain complexity is attenuated.

In that regard, I use complexity of functional connectivity, an aggregate measure of segregation and integration, as a marker of consciousness. This effort consists of two parts. First, I provide evidence that complexity in the healthy, awake brain is critical in the sense that it reflects a critical balance of segregation and integration designed to support efficient information communication. In turn, I provide evidence that loss of consciousness is associated with decreased complexity i.e. that functional connectivity departs from the critical complexity of the healthy, awake brain towards a more segregated configuration.

The structure of this thesis follows accordingly. In the first experimental chapter (3), I show the importance of the critical balance of complexity in the healthy, awake brain by using a structure-to-function association model. Specifically, I show that complexity can be derived upon certain optimal, structural connections (computed as the Nash equilibrium between regions), which promote efficient communication in the brain from the regional to the whole-brain level.

Chapter 4 focuses on capturing alterations of complexity in cases of sedation, anaesthesia and disorders of consciousness. Specifically, I show that as one goes from the awake state to anaesthetic-induced unconsciousness and disorders of consciousness, functional connectivity becomes less complex and more segregated. A refined approach that quantifies complexity in different parts of the brain allowed me to see whether this reduction in complexity is more evident in specific regions and networks. Under this framework, at the regional level I provide evidence that sparsely connected regions linking different parts of the brain play a critical role in whole-brain complexity. At the network level I show the importance of the default mode network in whole-brain complexity.

Even during rest, the brain is not static and displays rich temporal dynamics. Thus it is not only the complexity at each snapshot of time but also how complexity changes across time that can help us understand loss of consciousness. In chapter 5 I use a

dynamic framework to derive and characterize the dynamics of functional connectivity during loss of consciousness. In turn, I provide evidence that brains become less temporally complex as one goes from the awake state to anaesthetic-induced unconsciousness and disorders of consciousness.

Moreover, my goal is to see whether the principle of complexity reduction can be also applied to the developing brain. Towards this direction, in chapter 6 I use complexity on EEG connectivity data to examine anaesthetic-induced loss of consciousness in infants. Specifically, I show that complexity in anaesthetised infants aged 0-3 years is reduced compared to a state of emergence from anaesthesia, indicating its importance in supporting consciousness and brain function since infancy.

Taken together, these findings show that while the complexity of the healthy, awake brain during rest is critically configured, the unconscious brain is characterized by reduced complexity. Based on the results presented in this work, I propose that consciousness can be assessed on the basis of complexity of resting-state functional connectivity data.

## Table of Contents

|   |           |
|---|-----------|
| <b>Declaration .....</b>  | <b>4</b>  |
| <b>Acknowledgements.....</b>  | <b>5</b>  |
| <b>Abstract.....</b>  | <b>6</b>  |
| <b>List of figures and tables .....</b>                             | <b>13</b> |
| <b>Nomenclature .....</b>   | <b>18</b> |
| <b>Chapter 1: Introduction.....</b>                                 | <b>24</b> |
| <b>1.1 What is consciousness .....</b>                              | <b>24</b> |
| <b>1.2 Why is consciousness .....</b>                               | <b>25</b> |
| <b>1.3 Models of consciousness .....</b>                            | <b>26</b> |
| 1.3.1 Early approaches .....  | 26        |
| 1.3.2 Modern approaches-Global workspace .....                      | 27        |
| 1.3.3 Neuroscientific approaches-Global neuronal workspace .....    | 28        |
| 1.3.4 Dynamic core hypothesis: Segregation and integration.....     | 28        |
| 1.3.5 Neural coalitions and thalamocortical loops.....              | 29        |
| <b>1.4 How do we assess consciousness .....</b>                     | <b>29</b> |
| 1.4.1 The clinic as a starting point.....                           | 29        |
| 1.4.2 Sedation/anaesthesia .....                                    | 31        |
| 1.4.3 Disorders of consciousness.....                               | 32        |
| 1.4.4 Functional magnetic resonance .....                           | 33        |
| 1.4.5 Default mode function.....                                    | 34        |
| 1.4.6 Functional connectivity .....                                 | 35        |
| 1.4.7 Seed-based/ICA analysis.....                                  | 35        |
| 1.4.8 Resting-state connectivity during loss of consciousness ..... | 36        |
| 1.4.9 Advantages/limitations .....                                  | 38        |
| 1.4.10 Network/Graph-theoretic approach.....                        | 38        |
| 1.4.11 Dynamic functional connectivity .....                        | 41        |
| 1.4.12 Development and brain networks.....                          | 43        |
| 1.4.13 Development and consciousness.....                           | 44        |
| 1.4.14 EEG.....   | 44        |
| <b>1.5 Complexity .....</b>   | <b>46</b> |
| 1.5.1 Rationale .....   | 46        |
| 1.5.2 Defining complexity .....                                     | 47        |

|   |           |
|---|-----------|
| 1.5.3 Measures of complexity .....                        | 48        |
| 1.5.4 Criticality of complexity.....                      | 50        |
| 1.5.5 Loss of consciousness and reduced complexity.....   | 52        |
| 1.5.6 Temporal complexity.....                            | 52        |
| 1.5.7 Development and complexity .....                    | 53        |
| <b>Chapter 2: Methods .....</b>                           | <b>54</b> |
| <b>2.1 Principles of fMRI .....</b>                       | <b>54</b> |
| 2.1.1 MRI Physics.....                                    | 54        |
| 2.1.2 Spatial encoding.....                               | 59        |
| 2.1.3 BOLD signal.....                                    | 60        |
| 2.1.4 fMRI experiments.....                               | 62        |
| <b>2.2 Preprocessing fMRI data.....</b>                   | <b>63</b> |
| 2.2.1 Slice-timing correction.....                        | 63        |
| 2.2.2 Realignment/motion correction.....                  | 63        |
| 2.2.3 Co-registration and normalization .....             | 64        |
| 2.2.4 Smoothing.....                                      | 64        |
| 2.2.5 Functional connectivity .....                       | 65        |
| 2.2.6 Dynamic functional connectivity .....               | 66        |
| 2.2.7 Software.....                                       | 69        |
| <b>2.3 Principles of diffusion imaging.....</b>           | <b>70</b> |
| 2.3.1 Diffusion in the brain .....                        | 70        |
| 2.3.2 Diffusion weighted imaging.....                     | 70        |
| 2.3.3 Calculating diffusion directionality .....          | 71        |
| 2.3.4 Tractography and structural connectivity.....       | 73        |
| 2.3.5 Software.....                                       | 74        |
| <b>2.4 Network analysis .....</b>                         | <b>75</b> |
| 2.4.1 Network properties.....                             | 75        |
| 2.4.2 Complexity of functional connectivity networks..... | 78        |
| 2.4.3 Software.....                                       | 79        |
| <b>2.5 Electroencephalography .....</b>                   | <b>80</b> |
| 2.5.1 EEG signals .....                                   | 80        |
| 2.5.2 Frequency analysis .....                            | 82        |
| 2.5.3 EEG functional connectivity.....                    | 83        |
| 2.5.4 Source reconstruction .....                         | 85        |
| 2.5.5 Software.....                                       | 88        |
| <b>2.6 Code availability .....</b>                        | <b>88</b> |

|   |            |
|---|------------|
| <b>Chapter 3: Complexity of functional connectivity emerges from optimal structure .....</b>            | <b>90</b>  |
| <b>3.1 Introduction .....</b>   | <b>91</b>  |
| <b>3.2 Materials and Methods .....</b>  | <b>93</b>  |
| 3.2.1 Participants .....  | 93         |
| 3.2.2 Structural connectivity .....   | 94         |
| 3.2.3 Resting-state fMRI and functional connectivity .....  | 95         |
| 3.2.4 Assignment of ROIs to large-scale networks .....  | 96         |
| 3.2.5 Nash equilibrium Network model .....  | 97         |
| 3.2.6 Predicting functional from structural connectivity .....  | 99         |
| 3.2.7 Complexity of functional connectivity networks .....  | 102        |
| 3.2.8 Statistical analysis .....  | 103        |
| <b>3.3 Results .....</b>  | <b>104</b> |
| 3.3.1 Identifying optimal connections in structural connectivity networks .....                         | 104        |
| 3.3.2 Optimal structural connections in large-scale networks .....                                      | 106        |
| 3.3.3 Predicting functional connectivity and its complexity using optimal structural connections .....  | 108        |
| <b>3.4 Discussion .....</b>   | <b>110</b> |
| <b>3.5 Conclusion .....</b>   | <b>113</b> |
| <b>Chapter 4: Reduction of complexity during loss of consciousness .....</b>                            | <b>114</b> |
| <b>4.1 Introduction .....</b>   | <b>115</b> |
| <b>4.2 Changes in complexity under propofol-induced sedation and anaesthesia .....</b>                  | <b>116</b> |
| 4.2.1 Overview .....  | 116        |
| 4.2.2 Materials and Methods .....   | 118        |
| 4.2.2 Results .....   | 134        |
| <b>4.3 Changes in complexity: from sedation to disorders of consciousness .....</b>                     | <b>145</b> |
| 4.3.1 Overview .....  | 145        |
| 4.3.2 Materials and Methods .....   | 147        |
| 4.3.3 Results .....   | 149        |
| <b>4.4 Discussion .....</b>   | <b>151</b> |
| <b>4.5 Conclusion .....</b>   | <b>156</b> |
| <b>Chapter 5: Dynamic functional connectivity and its complexity during loss of consciousness .....</b> | <b>157</b> |
| <b>5.1 Introduction .....</b>   | <b>158</b> |

|  |            |
|--|------------|
| <b>5.2 Materials and methods</b> .....   | <b>160</b> |
| 5.2.1 Participants.....  | 160        |
| 5.2.2 Image acquisition and preprocessing.....   | 160        |
| 5.2.3 Hidden Markov Model.....   | 160        |
| 5.2.4 Statistical analysis.....  | 165        |
| <b>5.3 Results</b> .....   | <b>165</b> |
| 5.3.1 Brain dynamics during sedation.....  | 165        |
| 5.3.2 Brain dynamics in patients with disorders of consciousness.....                    | 169        |
| 5.3.3 Brain dynamics: from sedation to disorders of consciousness.....                   | 170        |
| <b>5.4 Discussion</b> .....  | <b>172</b> |
| <b>5.5 Conclusion</b> .....  | <b>177</b> |
| <br>   |            |
| <b>Chapter 6: Complexity of functional connectivity in the developing brain</b><br>..... | <b>178</b> |
| <b>6.1 Introduction</b> .....  | <b>179</b> |
| <b>6.2 Materials and Methods</b> .....   | <b>183</b> |
| 6.2.1 Participants.....  | 183        |
| 6.2.2 Experimental design and procedure.....   | 184        |
| 6.2.3 Clinical data collection.....  | 185        |
| 6.2.4 EEG data acquisition and preprocessing.....  | 186        |
| 6.2.5 EEG network construction and complexity.....                                       | 186        |
| 6.2.6 Source reconstruction.....   | 188        |
| 6.2.7 Statistical analysis.....  | 190        |
| <b>6.3 Results</b> .....   | <b>191</b> |
| 6.3.1 Power during different ages and conditions.....                                    | 191        |
| 6.3.2 EEG network alterations in infant anaesthesia.....                                 | 192        |
| <b>6.4 Discussion</b> .....  | <b>197</b> |
| <b>6.5 Conclusion</b> .....  | <b>199</b> |
| <br>   |            |
| <b>Chapter 7: Discussion</b> .....   | <b>200</b> |
| <b>7.1 Summary</b> .....   | <b>200</b> |
| <b>7.2 Complexity of functional connectivity</b> .....                                   | <b>201</b> |
| <b>7.3 Alterations of complexity in functional connectivity</b> .....                    | <b>203</b> |
| <b>7.4 Dynamic functional connectivity</b> .....   | <b>205</b> |
| <b>7.5 Complexity of functional connectivity in the developing brain</b> .....           | <b>207</b> |
| <b>7.6 Clinical insights</b> .....   | <b>209</b> |
| <b>7.7 Future work</b> .....   | <b>211</b> |

---

|                             |            |
|-----------------------------|------------|
| <b>7.8 Conclusion .....</b> | <b>213</b> |
| <b>References.....</b>      | <b>214</b> |
| <b>Appendix 1 .....</b>     | <b>252</b> |
| <b>Chapter 3 .....</b>      | <b>252</b> |
| <b>Chapter 4 .....</b>      | <b>255</b> |
| <b>Chapter 5 .....</b>      | <b>258</b> |
| <b>Chapter 6 .....</b>      | <b>260</b> |

---

## List of figures and tables

|  |          |
|--|----------|
| <b>Figure 1.1 - Simplified representation of different states of consciousness.</b><br>.....                       | Page 30  |
| <b>Figure 1.2 - Different levels of segregation/integration and brain network complexity.</b> .....                | Page 48  |
| <b>Figure 1.3 - Network degree entropy as a measure of brain complexity.</b><br>.....                              | Page 50  |
| <b>Figure 2.1 - Principles of MRI image acquisition.</b> .....   | Page 58  |
| <b>Figure 2.2 - A double-gamma model of the haemodynamic response function.</b><br>.....                           | Page 61  |
| <b>Figure 2.3 - Preprocessing functional MRI data.</b> .....   | Page 65  |
| <b>Figure 2.4 - Diffusion image acquisition and structural connectivity.</b><br>.....                              | Page 74  |
| <b>Figure 2.5 - Functional connectivity network analysis.</b> .....  | Page 77  |
| <b>Figure 2.6 - From EEG data to EEG networks.</b> .....   | Page 85  |
| <b>Figure 2.7 - Source reconstruction: from EEG data to cortical networks.</b><br>.....                            | Page 89  |
| <b>Figure 3.1 - Identifying optimal connections in structural connectivity networks using the NNG model.</b> ..... | Page 105 |
| <b>Figure 3.2 - Intra- and internetwork optimal connections in structural brain networks.</b> .....                | Page 107 |

---

|   |                 |
|---|-----------------|
| <b>Figure 3.3 - Predicting functional connectivity and its complexity from structural connectivity. ....</b>  | <b>Page 110</b> |
| <b>Figure 4.1 - Degree entropy of functional connectivity networks as a measure of complexity: a more refined approach. ....</b>                                  | <b>Page 126</b> |
| <b>Figure 4.2 - Complexity is decreased during moderate sedation. ....</b>  | <b>Page 135</b> |
| <b>Figure 4.3 - The functional role, location, and connectivity of sparsely connected regions during different levels of sedation. ....</b>                       | <b>Page 138</b> |
| <b>Figure 4.4 - Complexity of functional connectivity after removing networks. ....</b>   | <b>Page 140</b> |
| <b>Figure 4.5 - Individual differences in reaction times during moderate sedation correlate with rare-low and residual entropies during the awake state. ....</b> | <b>Page 142</b> |
| <b>Figure 4.6 - Complexity is decreased during anaesthesia. ....</b>  | <b>Page 144</b> |
| <b>Figure 4.7 - Complexity of sparsely connected regions from sedation to disorders of consciousness. ....</b>  | <b>Page 150</b> |
| <b>Figure 5.1 - Deriving brain dynamics from resting-state data using the Hidden Markov Model. ....</b>   | <b>Page 162</b> |
| <b>Figure 5.2 - Brain dynamics during sedation. ....</b>  | <b>Page 168</b> |
| <b>Figure 5.3 - Brain dynamics in patients with disorders of consciousness. ....</b>  | <b>Page 170</b> |
| <b>Figure 5.4 - Brain dynamics from sedation to disorders of consciousness. ....</b>  | <b>Page 172</b> |

---

|   |                 |
|---|-----------------|
| <b>Figure 6.1 - Spectral power analysis for infant brain dynamics under anaesthesia and emergence from anaesthesia. ....</b>                                    | <b>Page 192</b> |
| <b>Figure 6.2 - Complexity of alpha functional connectivity at the sensor level in the infant brain during anaesthesia and emergence from anaesthesia. ....</b> | <b>Page 193</b> |
| <b>Figure 6.3 - Complexity of alpha functional connectivity at the sensor level in the infant brain at different ages. ....</b>                                 | <b>Page 194</b> |
| <b>Figure 6.4 - Complexity of alpha functional connectivity at the source level in the infant brain during anaesthesia and emergence from anaesthesia. ....</b> | <b>Page 195</b> |
| <b>Figure 6.5 - Complexity of alpha functional connectivity at the source level in the infant brain at different ages. ....</b>                                 | <b>Page 196</b> |
| <b>Figure 7.1 - Complexity as an objective marker of consciousness. ....</b>  | <b>Page 211</b> |
| <b>Appendix Figure 3.1 - Predicting functional connectivity from structural connectivity matrices using different orders of polynomial transformation. ....</b> | <b>Page 253</b> |
| <b>Appendix Figure 3.2 - Predicting functional connectivity from structural connectivity matrices at the 129-ROI resolution. ....</b>                           | <b>Page 254</b> |
| <b>Appendix Figure 4.1 - Alterations in sparsely connected regions' connectivity during sedation. ....</b>  | <b>Page 255</b> |
| <b>Appendix Figure 4.2 - Residual entropy during moderate sedation. ....</b>  | <b>Page 256</b> |
| <b>Appendix Figure 4.3 - Residual rare entropy during moderate sedation. ....</b>   | <b>Page 257</b> |

---

**Appendix Figure 5.1 - Brain states extracted from the Markov model for the sedation experiment. .... Page 258**

**Appendix Figure 5.2 - Brain states extracted from the Markov model for the DOC experiment. .... Page 259**

**Appendix Figure 6.1 - Spectral power analysis for infant brain dynamics under anaesthesia. .... Page 260**

**Appendix Table 3.1 - Optimality statistics for the structural networks at the 129-ROI resolution. .... Page 254**

---

## Nomenclature

AAL - Automated Anatomical Labelling

A/D - Analog to Digital

AIMS - Anaesthesia Information Management System

ANOVA - Analysis of Variance

BCT - Brain Connectivity Toolbox

BEM - Boundary Element Method

BOLD - Blood Oxygen Level Dependent

CB - Cerebellar Network

CRS - Coma Recovery Scale

CRS-R - Coma Recovery Scale Revised

CSF - Cerebrospinal Fluid

DA - Dorsal Attention Network

dFC - Dynamic Functional Connectivity

DMN - Default Mode Network

DOC - Disorders of Consciousness

DSI - Diffusion Spectrum Imaging

DWI - Diffusion Weighted Imaging

eAAL - extended Automated Anatomical Labelling

EEG - Electroencephalography

EPI - Echo Planar Imaging

FA - Fractional Anisotropy

fMRI - Functional Magnetic Resonance Imaging

FO - Fractional Occupancy

FP - Fronto-Parietal network

FWHM - Full Width at Half Minimum

GFA - Generalized Fractional Anisotropy

GNW - Global Neuronal Workspace

GQI - Generalized Q-sampling Imaging

HARDI - High Angular Resolution Diffusion Weighted Imaging

HCP - Human Connectome Project

HMM - Hidden Markov Model

HRF - Haemodynamic Response Function

ICA - Independent Component Analysis

IIT - Integrated Information Theory

L - Limbic network

LCMV - Linearly Constrained Minimum Variance

LSN - Large-Scale Network

MCS - Minimally Conscious State

MEG - Magnetoencephalography

MNI - Montreal Neurological Institute

MOSSA - Maintenance Of Surgical Anaesthesia

MPFC - Median Prefrontal Cortex

MR - Magnetic Resonance

MRI - Magnetic Resonance Imaging

NIRS - Near-infrared spectroscopy

NNG - Nash-equilibrium Network Game Model

ODF - Orientation Distribution Functions

OEF - Oxygen Extraction Fraction

PCA - Principal Component Analysis

PCC - Posterior Cingulate Cortex

PET - Positron Emission Tomography

QBI - Q-ball imaging

RF - Radiofrequency

ROI - Region Of Interest

SEVO - Sevoflurane

SM - Somato-Motor network

SDF - Spin Distribution Function

SOA - Stimulus Onset Asynchrony

SPM - Statistical Parametric Mapping

SUB - Subcortical network

TE - Echo Time

TMS - Transcranial Magnetic Stimulation

TR - Repetition Time

VA - Ventral Attention network

VB - Variational Bayes

VIS - Visual Network

VS - Vegetative State

---

---

## Chapter 1: Introduction

Consciousness is the most familiar aspect of our lives but yet the most mysterious. There is something particularly compelling about consciousness. Although it is our most vivid experience we cannot reconcile it with what we know about the world. We understand how the autumn trees smell but we do not understand the experience that this smell creates for us. Despite this conundrum, conscious experience is part of the natural world, and like other phenomena that we observe around us, it necessitates an explanation. The three basic pillars for understanding consciousness can be summarized as follows: First what is consciousness and how can we pinpoint it to certain phenomena and experiences. Second why consciousness exists at all. The fact that consciousness can be generated by complex systems such as the brain is something that cannot be predicted by the brain's computational features alone. The final question pertains to how consciousness emerges. If conscious experience is part of the natural world then can it be attributed to some kind of physical process, for example neuronal firing in the grey matter of the brain? In the next sections I will discuss these questions and how these relate to the development of the neuroscience of consciousness.

### 1.1 What is consciousness

When talking about consciousness, it is notoriously difficult to pin down any complete definition. We can say that a being is conscious if it is something it is like to be that being (Nagel, 1974). In other words, a mental state is conscious if there is an associated quality of experience or qualitative feel of what it is that we experience. The collections of these experiences or feels have been also phrased as qualia (Jackson, 1982). The properties of qualia describe the phenomenology of consciousness. First they are subjective as they are private to the person experiencing them. This is also referred as the subjective quality of the experience. The qualia have also an internal property; there is something about “feeling like” a conscious agent. These and other properties are far from complete in describing the properties of consciousness. While many philosophers endorse the existence of qualia as valid phenomena (Chalmers, 1995), others dismiss qualia as introspective illusions (Dennett, 1988). Regardless, the

diversity of qualia suggests that the phenomenology of consciousness is associated with a vivid, multifaceted experience. However, describing this phenomenology of consciousness does not entail why these phenomena exist at all.

## 1.2 Why is consciousness

Questions regarding why there should be consciousness at all are characterized as the “hard” problem of consciousness (Chalmers, 1995; 1996). The fact that brains as complex systems can produce information and react to stimuli in sophisticated ways is no philosophical mystery. If one sees a red object, information will be passed from the neural pathways of retina all the way to the visual cortex invoking a certain neural activation that, in turn, can lead to reporting the colour of the object as red. However, these physical processes cannot entail why we are ever accompanied by the phenomenal experience of “redness”-how it “feels” to see a red colour. A commonly used example to address the distinction between how it feels and the physical process of “feeling/experiencing” regards a neuroscientist who, for years in complete isolation, became proficient at knowing the complete physical truth - everything in completed physics, chemistry, neuroscience, etc. After finally leaving the room, she experiences the red colour for the first time and that constitutes a completely new experience for her that goes beyond the physical truth she learned all these years (Jackson, 1982).

To this end, answers for the hard problem of consciousness mostly address the question of whether we should even attempt to intertwine consciousness with physical phenomena and neuronal assemblies. This has been formulated as the explanatory gap, referring to the arduous task of explaining the phenomenology of consciousness in terms of physical phenomena (Alter and Walter, 2007). Some philosophers believe that this gap cannot be fulfilled and they use mental experiments to reject the idea of physicalism in consciousness (Chalmers, 1996). However, other philosophers claim that consciousness is functionally definable, meaning that if we understand everything about *how* consciousness emerges then we can understand why consciousness is produced and there will be nothing else to explain (Dennett, 2005). This leads me to the next section where I will focus on how consciousness is produced in the brain. I

will leave the philosophical endeavours for explaining why there should be consciousness at all as something that can be further discussed outside this thesis.

### **1.3 Models of consciousness**

How consciousness is produced is usually referred to as the “easy” problem of consciousness not because it is easily tangible but because we know how to approach it using the empirical method (Chalmers, 1995). The goal is to create an accurate description of how the neural assemblies in the brain produce what one perceives as conscious experience. As I mentioned previously, bridging the phenomenology of consciousness with how the brain processes information to produce consciousness might provide new avenues for understanding consciousness; thus I will focus on this approach more extensively.

#### **1.3.1 Early approaches**

Early accounts regarded consciousness as part of the mind and discussed how its properties relate to its physical substrate and the body (also called the “mind-body problem”). Starting with ancient western philosophers, Aristotle in the 4<sup>th</sup> century BC believed that the mind could not be separated from the capacities of the body (Sihvola, 2007). In the late 4<sup>th</sup> century/beginning of 3<sup>rd</sup> century BC, Democritus and Epicurus proposed that the mind arises from matter stating, “there is only atoms and the void” thus providing a fully physical representation of the mind on the premises of atoms and their interactions (Leszl, 2006). In the seventeenth century, Descartes took a different approach on the mind-body problem and argued that mental substance is distinct from matter. He considered consciousness as a part of the non-physical mind and thus something that can be separated from the body (“mind-body dualism”)(Hatfield, 2000). On the contrary, and perhaps based on the early physical theories proposed by the ancient Greek philosophers, Spinoza argued that anything that can be inferred from the mind will have its parallel in the body that it represents (Curley, 1986). Thus, Spinoza regarded consciousness as something related to the way the mind is represented in its physical substrate (like the human body) and not a separate entity. At the same time, Leibniz expanded on Spinoza’s notions of

consciousness as being a representation of the mind within the body and developed a theory of mind that was fully naturalized. In that regard, he claimed that the way distinct conscious experience arises relates to the functional organization of the body (Remnant and Bennett, 1996).

### **1.3.2 Modern approaches-Global workspace**

The mind-body theories by Spinoza and Leibniz developed during the seventeenth century treated consciousness as a phenomenon to be explained in natural ways, and so they endeavoured to explain consciousness in terms of the underlying physical representations. This fostered functional accounts of consciousness that continue to this day, relating consciousness to its physical substrate and the brain. However, even after the onset of modern scientific psychology in the mid-nineteenth century, the mind was still largely equated with consciousness. Psychological theories argued that the conscious mind is characterized by a “stream” of thoughts i.e. a stream of consciousness. For example, William James, the father of modern psychology, argued that:

“Consciousness, then, does not appear to itself chopped up in bits. Such words as 'chain' or 'train' do not describe it fitly as it presents itself in the first instance. It is nothing jointed; it flows” (James, 1890)

Despite some advancement in the psychological aspect of consciousness and its association with the mind, the relationship of consciousness to the brain remained very much a mystery. In the 1980s and 1990s, with the renewed emphasis on explaining cognitive capacities such as memory, perception and language comprehension in the brain, there was a major resurgence of scientific interest in the relationship of consciousness to the neural machinery of the brain. In that regard, Posner (Posner and Snyder, 1975; Posner and Rothbart, 1998) and Shallice (Shallice, 1972; 1988; Norman and Shallice, 1980) proposed that information is conscious when it is represented in an “executive attention” or “supervisory attentional” system. A more refined theory, involving how information processing at the brain-wide level

produces consciousness, was proposed by Baars. The basic idea is that information becomes conscious because it is globally available to the brain (Baars, 1988). Functionally specialized modules produce information that can become part of the global workspace and thereby shared by other modules in the brain. What we perceive as conscious content is what is being globally shared. One interesting account of this idea is that there is considerable amount of neural processing occurring at any given time in the brain that we are not aware of because it is not globally available.

### **1.3.3 Neuroscientific approaches-Global neuronal workspace**

Fuelled by technological advances, interest in the neuroscientific study of consciousness substantially increased from around 1990 onwards. Publications of studies on neural correlates of consciousness appeared, in particular by Francis Crick and Christof Koch (Crick and Koch, 1990). They proposed that looking for neural correlates alone allows the neuroscience of consciousness to progress independently from philosophical disputes. Further discussions on the neuronal basis of consciousness emerged leading to the specific neuronal theories of consciousness. One of them is the global neuronal workspace (GNW) theory proposed by Dehaene and Naccache (Dehaene and Naccache, 2001) that emphasizes the role of integrated connectivity in order to mobilize the relevant neuronal modules that become available in the global workspace. This proposal claims the existence of long-range excitatory axons, particularly dense in prefrontal and cingulate regions alongside thalamocortical connections, that interconnect multiple specialized processed. GNW neurons amplify and maintain a specific neural representation in the global workspace, making it conscious. Then long-distance axons of GNW neurons broadcast information to other specialized modules for brain-wide processing (Dehaene and Naccache, 2001).

### **1.3.4 Dynamic core hypothesis: Segregation and integration**

Alternative neuroscientific approaches of consciousness relate not to whether there are specific groups of neurons but to what are the properties these should have in order to produce consciousness. The dynamic core hypothesis proposed that the

ability of the brain to sustain consciousness is tantamount to its ability to integrate information from different neuronal systems (Tononi and Edelman, 1998). In that regard, Tononi and Edelman proposed that information encoded by a group of neurons attains a conscious level only if it achieves a level of segregation (i.e. sufficiently specialized) and a level of integration with other neuronal groups (i.e. sufficiently functionally connected with the rest of the brain) (Tononi and Edelman, 1998). Expanding on this idea, the integrated information theory (IIT) argues about the brain's capacity to produce consciousness as a result of its ability to integrate information. The theory has been evolving since 2004 (Tononi, 2004), heralding the idea that the brain (and perhaps any biological system) can produce consciousness on the premise of the effect it has on itself (intrinsic cause-effect power). The level of this causal effect is defined over complicated spatiotemporal scales and can discriminate between different states of consciousness (Tononi, 2004).

### **1.3.5 Neural coalitions and thalamocortical loops**

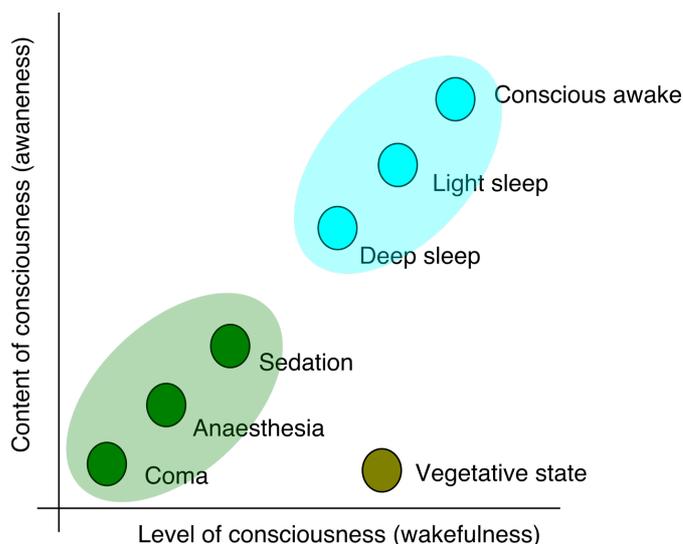
Other theories of consciousness emphasize the thalamocortical system as the one generating consciousness. Llinás and colleagues proposed, on the basis of earlier studies using MEG, that synchronous oscillations/rhythms in thalamocortical loops in different frequency bands create the conscious state (Llinás et al., 1998). In addition, these thalamocortical responses are modulated by brainstem inputs that regulate arousal pathways (Llinás et al., 1998). Synchronous oscillations between neurons have been involved in other theories of consciousness. For example, Crick and Koch suggested instead that consciousness might require competition among “coalitions” of neurons, in which winning coalitions determine the contents of consciousness at a given time (Crick and Koch, 2003). They pointed out that neuronal coalitions carry a similar notion to that of the dynamic core theory where all neurons need to be integrated.

## **1.4 How do we assess consciousness**

### **1.4.1 The clinic as a starting point**

In parallel to the proposed theories describing how consciousness is produced, there is

a clinical need to assess and measure consciousness in order to help patients—for example assessing whether a patient is aware of her surroundings after brain injury or is unconscious during surgical anaesthesia. Consciousness, as measured and defined by clinicians, is a practical starting point in the discussion about measures and theories of consciousness. Early clinical studies of the level of consciousness started in the 1940s when scientists investigated the neural components regulating the brain’s sleep-wake mechanisms. Since then, new empirical studies emerged with the goal of capturing consciousness in a principled way that builds upon but moves beyond the study of wakefulness. Based on observations from anaesthesia and disorders of consciousness studies, the spectrum of conscious states proved to be complex, which led to the idea of systematizing the conscious state into two distinct components: arousal/wakefulness and awareness (Laureys, 2005). Arousal refers to one’s state of vigilance and sleep-wake cycle, whereas awareness indicates the ability to be aware of reporting something consciously. The combination of these two aspects of consciousness has defined a “map” of states of consciousness outlined in a two-dimensional space: one axis being the level of wakefulness and the other axis being the level of awareness (Fig 1.1).



**Figure 1.1 - Simplified representation of different states of consciousness.** States of consciousness can be mapped on a two-dimensional space: the first dimension being the level of awareness (the content of consciousness) and the second dimension being the level of wakefulness (the level of arousal/alertness). Adapted from Laureys (2005).

States range from low wakefulness and awareness, such as coma and general anaesthesia, to the highest level of awareness and wakefulness-the conscious awake. In this thesis, I will restrict my focus on two specific states of this consciousness map: pharmacologically-induced sedation/anaesthesia and disorders of consciousness and I will compare them to the conscious awake state.

#### **1.4.2 Sedation/anaesthesia**

Anaesthesia is defined as the controlled and reversible induction of unconsciousness (hypnosis), analgesia, and muscle relaxation (Mashour, 2004). At a small dose, anaesthetics first suppress thinking, focused attention, and working memory. As the dose increases, consciousness and voluntary responsiveness begin to fade signifying a transition to a sedative state and, if the anaesthetic dose is high enough, to a fully anaesthetized state. One example of anaesthetic widely used in the clinical setting is propofol, an intravenous anaesthetic agent with putative actions on the inhibitory GABA<sub>A</sub> receptor (Jevtovic-Todorovic, 2016). Anaesthetics like propofol hyperpolarize neurons by increasing inhibition and suppressing cerebral blood flow in multiple cortical regions (Alkire et al., 1995; Fiset et al., 1999) as well as blood flow in the thalamus and insula (Xie et al., 2011; Warnaby, et al., 2017). Evidence also exists that anaesthetics act on sleep-promoting pathways, thereby impairing arousal levels (Brown et al., 2010; Vacas et al., 2013).

A question of particular clinical importance relates to how anaesthetics abolish consciousness. For example, anaesthetics are believed to act on midline areas (such as the premotor cortex, the medial prefrontal cortex, the anterior cingulate cortex and the thalamus) a process that has been associated with abolishing behavioural responsiveness but not necessarily with losing consciousness (Alkire et al., 2008). To this end, theoretical accounts have suggested that anaesthetics abolish consciousness because they alter the functional cross-talk between brain regions (i.e. its functional connectivity)(Alkire et al., 2008; Hudetz, 2012). I will return to this point at a later part of this chapter.

### 1.4.3 Disorders of consciousness

Understanding how consciousness is lost during anaesthesia can help us understand clinical cases where brain injury has caused loss of consciousness. Such clinical cases can be classified in different states of unconsciousness based on the levels of arousal and awareness (Laureys, 2005). Coma is the state where patients exhibit minimal arousal and awareness. Patients that have awoken from a coma, but still have not regained awareness are considered to be in a vegetative state (Laureys, 2005). Patients in vegetative state can recover and attain a minimally conscious state when there is reproducible evidence of awareness (Boly and Seth, 2012).

An initial way to assess consciousness in these patients is by collecting psychological and behavioural data (Naccache, 2017). These could include probing behavioural properties, the presence of which would unquestionably signal a conscious state (Naccache, 2006). These approaches are useful, particularly for assessing the level of awareness in the presence of some form of wakefulness. However, the difficulty in probing consciousness in cases where obtaining subjective reports is challenging, makes the use of these measures impractical (Naccache, 2017). To circumvent this problem, passive methods of approximating awareness during impaired wakefulness are sometimes used. In this context, the next significant progress was made with the introduction of more advanced behavioural measures. Giacino and his colleagues defined, in 2002, the minimal conscious state as being a state in which: “cognitively mediated behaviour occurs inconsistently, but is reproducible or sustained long enough to be differentiated from reflexive behaviour” (Giacino et al., 2002). On this basis, a more refined behavioural measure was devised: the coma recovery scale (CRS). A revision of this score (CRS-R) provided a useful tool to clinicians to differentiate between minimally conscious state patients and other cases of disorders of consciousness (Gerrard et al., 2014). Despite its success (for example in cases that were mistakenly misclassified as vegetative or coma-Schnakers et al., 2009), studies have rendered the use of this measure important but not sufficient. This claim is based on the fact that a large set of cortically generated complex behaviours might escape conscious reports such as those collected by the CRS-R score (Naccache, 2017). On that note, recent approaches suggest that studying spontaneous patterns of brain

activity during rest, irrespectively of the behavioural capacity of the patient, might provide more information about the unconscious brain (Soddu et al., 2011). For example, investigating potential functional disconnections, in which case brain regions cannot communicate information, is now considered to be a fruitful approach for looking at different cases of disorders of consciousness (Monti, 2012).

Together, the results from anaesthesia and disorders of consciousness suggest that, in order to help clinicians quantify and evaluate different states of consciousness, a fruitful approach maybe that of studying patterns of brain activity “at rest” (Raichle et al., 2010). Of particular importance is that this approach can reveal neural correlates of unconsciousness without relying on the behavioural capacity of each individual or other cognitive markers. In the next chapter I will focus on obtaining and analysing these resting-state patterns and describe some of their applications to loss of consciousness.

#### **1.4.4 Functional magnetic resonance**

The last decade has seen an unprecedented emergence of high-resolution neuroimaging techniques that have allowed us to non-invasively map the functional and structural core of the human brain (Raichle, 1998). Neuroscientists can now manipulate the conditions of an experiment and analyse neuroimaging data in order to identify regions engaged during this manipulation (Posner and Raichle, 1994). For example, early studies by Petersen and colleagues used positron emission tomography (PET) to show regions activated during semantic processing (Petersen et al., 1988). An important milestone in this effort was the development of functional magnetic resonance (fMRI) imaging as an indirect measure of neural activity-it measures haemodynamic changes under the presence of a strong magnetic field (Logothetis et al., 2001). Soon after the development of fMRI, Ogawa and colleagues found that slow (<1 Hz) haemodynamic fluctuations (or blood oxygen level dependent–BOLD signal) have robust, reproducible oscillatory dynamics (Ogawa et al., 1990). These slow fluctuations have been further linked with electrophysiological recordings (Nir et al., 2007) and neuronal spiking recordings in animals (Shmuel and Leopold, 2008) and humans (Fox and Raichle, 2007), showing that fMRI can be used as a tool to map the

neural correlates of brain function. Early applications of fMRI focused on looking at differences in brain activity (activations) only in the context of a “task condition”. Brain activity was calculated by contrasting task conditions to a non-task or control condition i.e. subtracting the two conditions. The control condition to which task data would be compared was initially a resting condition (or resting-state) but became more sophisticated as the technique developed (Price, 2012).

#### **1.4.5 Default mode function**

Despite the prevalence of the subtracting technique in the early fMRI experiments, it was soon realized that rest was not just a state where brain dynamics during tasks could be compared. In an almost unexpected manner, studies subtracting task from rest showed that the activity of a set of regions was higher in rest compared to task. The significance of these was initially neglected. Interest was revitalized when fMRI studies, propelled by a positron emission tomography (PET) meta-analysis (Shulman et al., 1997), showed that a set of regions, later termed the “default mode network” (DMN), was actually relatively deactivated during task (Raichle, 2015). The question remained of whether the dynamics of these regions—apparently evident during rest—had specific neurobiological basis and was not an artefact of subtraction analysis. To show this, Raichle and colleagues utilized two important pieces of information from PET imaging. First, it had been previously shown that the transient increases in brain activity from a resting/baseline condition to a task activation, resulted in an increased amount of blood oxygen as the ratio of oxygen consumed to oxygen delivered (OEF) decreased (Fox and Raichle, 1986). Thus OEF was regarded as a marker of activation: if no activation existed, then OEF was uniform across the brain. In turn, a PET study showed that, during resting-state scanning, oxygen consumption and delivery depicted the highest level of deviation in the regions of the DMN while, at the same time, OEF was uniform across the brain (Gusnard and Raichle, 2001). Thus, although there was no activation during rest, as indicated by the uniform OEF, there was still high metabolic activity in the default mode regions. Further studies verified that resting-state fMRI is associated with the “activation” of the default mode regions and has been termed as so-called default mode function of the brain (Raichle et al., 2001;

Raichle, 2015). In light of this, the analysis of activity during rest paved the way for a large area of research known as resting-state fMRI and the corresponding analysis technique called functional connectivity.

#### **1.4.6 Functional connectivity**

As evident by the previous remarks, until the middle of the 1990s the major corpus of neuroscience research focused on task-induced activations. Although low frequency oscillatory dynamics had been observed since the first fMRI studies, their value was largely neglected (Frahm et al., 1993). Analysis of spatially coherent BOLD fluctuations during resting-state made a significant contribution towards understanding functional organization between regions. Functional connectivity became popular when Biswal and colleagues showed that functionally relevant regions in the motor cortex during rest were coupled in terms of their BOLD responses (Biswal et al., 1995). This showed that BOLD signals between these regions were functionally connected, potentially producing a motor “system” similar to the one obtained by the conventional subtractive activation analysis. Incentivized by these approaches, functional connectivity is now understood as a tool for looking at correlated BOLD responses at resting-state without the need of conventional subtractive techniques.

#### **1.4.7 Seed-based/ICA analysis**

Expanding on the technique used by Biswal and colleagues, seed-based functional connectivity can now assess the functional connectivity between the average signal from a set of predefined voxels (seed) and the rest of the brain. Interestingly, using this seed-based analysis technique Greicius and colleagues also reported the temporal coherence of the default mode brain regions, suggesting their functional connectivity during resting-state scanning, and introduced the default mode network (DMN) terminology (Greicius et al., 2003).

One problem with this technique is that it needs an *a priori* definition of a seed.

Although this can be particularly compelling during hypothesis-driven research, it can restrict the interpretation of the results or make false assumptions about the data (Cole et al., 2010). Independent component analysis (ICA), on the other hand, is a multivariate, data-driven technique, which aims to segregate the temporal signal into spatially and statistically distinct components with similar temporal dynamics. In comparison with seed-based functional connectivity, ICA requires few *a priori* hypotheses usually related to the number of components (Beckmann et al., 2005).

Using ICA techniques it was found that different parts of the brain, at rest, were making up different functional systems or large-scale networks (LSNs) (Smith et al., 2009; Bressler and Menon, 2010; Yeo et al., 2011). These include sensory networks like the visual and auditory networks, central executive networks (Cole et al., 2014a), and the DMN (Raichle, 2015). Interestingly, since its initial discovery the DMN is now robustly reproduced using these techniques across different cohorts showing a distinct spatial pattern including the ventral medial prefrontal cortex (MPFC), the dorsal medial prefrontal cortex, the posterior cingulate cortex (PCC) and adjacent precuneus plus the lateral parietal cortex (Binder et al., 1999; Mazoyer et al., 2001). The rich dynamics of resting-state activity have been also used to study relationships between LSNs. Fox and colleagues found that networks responsible for internal processing were anti-correlated with executive networks during resting-state, suggesting a specific toggling of information between different specialized systems (Fox et al., 2005). This led to the stratification of networks into task-positive/executive and task-negative/default mode networks that perform antithetic roles in information processing during different cognitive contexts (although this claim has been further refined-Spreng, 2012; Vatansever et al., 2017).

#### **1.4.8 Resting-state connectivity during loss of consciousness**

As mentioned previously, identifying different states of consciousness has been linked with studying resting-state activity and its functional connectivity. Similar to the trajectory that led to the emergence of resting-state connectivity, initial studies into the functional correlates of loss of consciousness started with PET. In cases of propofol-induced anaesthesia studies used PET to measure cerebral blood flow and

they showed reductions in a variety of regions such as the thalamus, the cuneus and precuneus, the PCC and the angular gyri (Alkire et al., 1995; Fiset et al., 1999). In similar fashion, earlier studies using PET showed significant alterations in patients with disorders of consciousness especially in the precuneous (Laureys et al., 2004) and the thalamus (Laureys et al., 2000b).

With the emergence of resting-state connectivity, understanding brain dysfunction through aberrant functional connectivity has emerged as a promising method of translating fMRI into the clinical setting (Fox and Greicius, 2010; Rubinov and Bullmore, 2013). Studies investigating loss of consciousness have shown a radical reorganization of functional connectivity between certain regions or LSNs. During sedation, alterations in functional connectivity have been reported (Barttfeld et al., 2015b), especially in higher-order networks such as the DMN (Stamatakis et al., 2010) while anaesthesia seems to also impair fronto-parietal connectivity (Boveroux et al., 2010). On the contrary, the functional connectivity of primary sensory networks, such as the auditory or the visual network, seem to be left intact from the effect of the anaesthetic (MacDonald et al., 2015). The thalamus is another important region under consideration for altered functional connectivity during loss of consciousness. In that regard, thalamic functional connectivity with major higher-order LSNs appears to decrease during anaesthesia (Guldenmund et al., 2013).

In cases of disorders of consciousness a variety of alterations in functional connectivity have been also observed. Studies have reported changes in DMN connectivity compared to healthy controls (Vanhaudenhuyse et al., 2010; Di Perri et al., 2016) as well as the fronto-parietal network (Crone et al., 2014). Disconnection between primary sensory areas and higher-order networks has been observed in patients with disorders of consciousness (Demertzi et al., 2015). Thalamocortical functional connectivity is also impaired in patients with disorders of consciousness while restoration of thalamocortical connectivity has been shown to align with emergence from the vegetative state (Laureys et al., 2000a).

### 1.4.9 Advantages/limitations

While important pieces of information, analysing functional connectivity using seed-based analysis or ICA analysis might have specific problems. Seed-based analysis comes with an *a priori* definition of a seed region whereas ICA has been criticized for the subjectivity in selecting the number of components to be produced (Beckmann et al., 2009). Contributing to this ambiguity, the previous results on functional connectivity during loss of consciousness show that changes in functional connectivity involve alterations between multiple regions or LSNs. Taken together, restricting the focus of functional connectivity analysis to a certain number of regions or LSNs might hinder understanding of multiple alterations that can explain loss of consciousness.

### 1.4.10 Network/Graph-theoretic approach

A different approach constructs appropriate models that can capture the connectivity of *multiple* regions rather than focusing on a restricted scope of connectivity. Having roots in the fields of mathematics and physics, one popular way of modelling multiple interactions is using a complex network or graph (Bullmore and Sporns, 2009). Analysis of network properties offers a new way of looking at functional connectivity through the lens of network organization between multiple regions or LSNs. In the next section I will describe networks and their properties that have been used for modelling functional connectivity and I will provide examples of their applications in studying the organization of functional connectivity in cases of anaesthetic-induced unconsciousness and disorders of consciousness.

### *Networks*

Theoretical approaches have discussed the idea of the “connected brain” where brain function is mediated by the way each brain component is connected to each other. This approach was influenced by connectionism and other theories of philosophical and psychological content (Catani et al., 2013). The study of brain networks finds its origins in the works of Santiago Ramon y Cajal who used light microscopy to discover that distinct anatomically defined units were well interconnected (Cajal, 1995). The

emergence of non-invasive neuroimaging techniques and the acquisition of structural and functional data made it easier to quantify structural connections and functional connections across the whole brain and investigate their properties.

Analysis of networks and their properties has its roots in systems and graph theory, where a network is a complex system comprising a collection of nodes and edges (Bullmore and Sporns, 2009). Network properties capture their underlying organization and are defined in specific terms pertaining to the domain of graph-theoretic analysis (Sporns, 2013). For example, the degree of each node represents the extent to which each node is connected to the rest of the network (number of connections of each node). Other examples include modularity, that quantifies how well a group of nodes can be grouped together to form a distinct module, and global efficiency that reflects how a network is communicating information efficiently across all its nodes (Fornito et al., 2016). In the case of brain networks, nodes correspond to the brain components under investigation and edges correspond to the relationships between these components. One particular example for defining nodes is by using a brain parcellation. Since the original work of Brodmann in 1909 (Brodmann, 1909), the brain has been parcellated into a certain number of regions of interest (ROIs) spanning different parts of the brain that are believed to have similar anatomical or functional basis. Other cortical and subcortical parcellations have since been developed using different data-driven techniques (Makris et al., 2005). Thus brain networks can be considered as networks connecting multiple cortical and subcortical regions covering the entire brain.

### ***Structural connectivity networks***

Brain networks can be broadly classified into two categories based on the type of connectivity: structural and functional connectivity networks (Fornito et al., 2016). Structural networks are considered when edges correspond to the underlying anatomical connections between regions. Structural connectivity networks were constructed as early as the 1990s in the primate cortex. Felleman and Van Essen compiled a structural connectivity matrix from the prior tract-tracing literature that summarized 305 axonal connections between 32 areas of the visual cortex in the

macaque monkey (Felleman and Van Essen, 1991). Similar analysis of the cat cortex showed that cortico-cortical connections were often organized into distinct systems (Scannel et al., 1995). With recent advances in diffusion imaging and tractography methods, white matter tracts between brain regions of interest have now been used to represent the human brain's structural connectivity network. Converging evidence using diffusion imaging has shown that structural connectivity networks have a specific organization (Hagmann et al., 2008; Bullmore and Sporns, 2012). Fundamentally associated with this organization is the “heavy-tailed” shape of the degree distribution. Its shape shows that the degree property is not normally distributed but it rather includes the existence of highly connected regions (or hubs) and other less connected (sparsely connected) regions that, together, enable whole-brain communication (Hagmann et al., 2008). In addition, structural networks have been shown to have a “small-world property” by combining high clustering and increased global efficiency (Watts and Strogatz, 1998; Bullmore and Sporns, 2009). This property allows whole-brain communication to be efficient while, at the same time, utilizing minimum wiring (Bassett and Bullmore, 2006).

### ***Functional connectivity networks***

In the case of functional connectivity networks, edges stand for correlations between the BOLD time series obtained from certain ROIs. The realization that functional connectivity also has a specific organization came later than structural connectivity networks. Initial studies of epileptiform activity in the macaque cortex showed that its network organization has small-world properties (Kötter and Sommer, 2000). With the increasing use of fMRI, functional connectivity networks began to emerge as means of representing whole-brain functional connectivity. Multiple studies have shown that fMRI functional connectivity networks have also specific organization (Bassett and Bullmore, 2006). Their degree distribution has been shown to have a “heavy-tailed” shape with the existence of highly and sparsely connected regions (Vértes et al., 2012).

### ***Networks and loss of consciousness***

The value of network neuroscience emerged when neuroscientists started to ask how network properties alter between different conditions. Changes in network properties have been shown to capture differences in functional organization from task execution in health individuals (Vatansever et al., 2015a) to neurodegenerative disorders such as schizophrenia (Lynall et al., 2010). Networks also change their organization during loss of consciousness. The degree of highly connected regions (hubs) is decreased in patients with disorders of consciousness (Crone et al., 2014) and anaesthesia (Schröter et al., 2012). The degree of non-hub regions has also been shown to change in patients with disorders of consciousness (Achard et al., 2010; Crone et al., 2014). In addition, modularity and global efficiency of functional connectivity networks is altered during loss of consciousness with networks becoming more fragmented (Schröter et al., 2012; Crone et al., 2014). Taken together, network analysis has revealed a diverse spectrum of changes in the organization of functional connectivity networks during loss of consciousness not previously seen with seed-based or ICA methods.

#### **1.4.11 Dynamic functional connectivity**

Functional connectivity networks change not only between different conditions but also *within* conditions. Brains dynamically change even when they do not have to deal with incoming information reflecting the dynamic neuronal signalling that underlies brain activity (Rabinovich et al., 2012). Accordingly, functional connectivity networks and their properties are expected to change across time, showing that regions can dynamically connect or disconnect. Studies have used dynamic functional connectivity (dFC) methods in order to demonstrate changes in functional connectivity across time (Hutchison et al., 2013). These methods usually calculate multiple functional connectivity networks corresponding to different time windows. In turn, in order to find out whether network properties change across time, properties of each “windowed” network are calculated and contrasted against the properties of other “windowed” networks (Hutchison et al., 2013).

Identifying changes in functional connectivity in the temporal domain in experiments of loss of consciousness is important for several reasons. First theoretical models suggest that consciousness is also associated with the temporal repertoire of states the brain can access (Tononi, 2004; Carhart-Harris et al., 2014; Tagliazucchi et al., 2014). In that regard, loss of consciousness is associated with shrinking this repertoire and thus restricting brain access to a limited amount of information (Alkire et al., 2008). Second, temporal changes in functional connectivity can help us identify the dynamic transitions into and emergence from unconsciousness (Långsjö et al., 2012; Hudson et al., 2014). Third, it can allow us to scrutinize the state of unconsciousness at each time point. This is important, for example, in cases of anaesthesia where brain responses need to be monitored constantly under certain anaesthetic doses (Hudetz et al., 2015). In cases of disorders of consciousness, characterization of changes in functional connectivity might elucidate dynamics not previously seen by the static method. These dynamics might be able to sustain some kind of residual and behaviourally covert awareness that can help assess the state of consciousness in these patients (Naro et al., 2018).

Under this dynamic connectivity framework, previous work has shown that the dynamics of functional connectivity change over time during anaesthesia indicating a reduced repertoire of functional configurations compared to the awake brain (Hudson et al., 2014; Barttfeld et al., 2015a). On a similar note, studies of dynamic functional connectivity in disorders of consciousness have shown a reconfiguration of between and within LSN connectivity across time (Di Perri et al., 2018).

Although they provide important pieces of information, one major weakness of these dFC methods relates to whether they can fully capture the repertoire of brain dynamics (Hutchison et al., 2013). Recent approaches show that functional connectivity has much richer dynamics and can be described by the ongoing reconfiguration of connectivity and switching between different states that go beyond the temporal resolution of the windowed dynamic connectivity (Bressler and Tognoli, 2006). One idea for unveiling the underlying rich dynamics regards finding coherent dynamic states without the need for an *a priori* definition of windows of time (Baker

et al., 2014). Appropriate mathematical analysis can provide a temporal decomposition of brain activity into a number of states/networks that characterize the recurrence of brain dynamics during the experiment (Vidaurre et al., 2017a). Thus, these alternative dFC models can identify how brain activity in the unconscious brain temporally organizes into discrete states with the potential of studying their rich dynamics not previously seen with windowed connectivity (Baker et al., 2014). In that regard, applying these techniques to characterize loss of consciousness is an important approach.

#### **1.4.12 Development and brain networks**

Functional connectivity networks do not only change in a short time scale but also with age. Starting from birth, symmetric brain regions of primary sensorimotor and visual networks are shown to be functionally synchronized (Gao et al., 2015). In contrast, the fronto-parietal and DMN networks are still in a premature, “scattered” stage and become functionally coherent later in the first year of life (Gao et al., 2009) propelled by cortical and subcortical maturation (Petanjek et al., 2011). Overall, the development of functional connectivity in the first year of life starts with the development of primary sensory systems followed by the development of higher-order and default mode systems; however different networks demonstrate unique timings and developmental trajectories. Thalamocortical connectivity is also different during infancy. For example, studies have found that neonatal thalamic functional connectivity is dominated by connections to the primary sensory sensorimotor/auditory/visual networks (Alcauter et al., 2014). At later stages in the first year of life, thalamocortical connectivity includes DMN and fronto-parietal networks (Alcauter et al., 2014). Beyond the first year, reorganization of functional connectivity persists and it does not attain adult-like level until later in childhood. For example, studies have shown that networks in younger children are less integrated compared to older children, suggesting that their inter-network communication is not fully developed (Fair et al., 2007b; Vértes and Bullmore, 2015).

### **1.4.13 Development and consciousness**

The developmental trajectory of the human brain can provide a unique window of information into studying the features of consciousness. At the very early stages of life, newborn infants display features characteristic of what may be referred to as basic consciousness. Based on evidence regarding the increasing integration of higher-order and default mode LSNs as well as the development of thalamocortical connectivity, theories suggest that consciousness develops from its basic level to an adult-like level during the first years of life (Lagercrantz and Changeux, 2009). Thus, it is now believed that the developmental trajectory is also reflective of a consciousness trajectory where the postnatal maturation of the brain allows the development of more complex circuitry that supports increased conscious processing (Rochat, 2003). So far I have discussed only one point in this trajectory, namely the adult consciousness. Investigating consciousness at the other end of the spectrum, the developing brain, might provide validation or extension of the results for the adult brain. For example, certain questions emerge as to whether loss of consciousness in the developing brain is associated with functional connectivity reorganization, even though it has not fully attained adult-like level. Due to ethical and practical reasons, studying loss of consciousness in the developing brain is an arduous task; thus only a limited number of studies have looked at loss of consciousness at that age range (Mongerson et al., 2017). Adding to this, the previous differences reported by others between adult and infant functional connectivity networks suggest that translating results from the adult to the infant brain is not trivial. In that regard, more tailored techniques and conclusive results are required to fully understand loss of consciousness in these stages of life.

### **1.4.14 EEG**

#### *Frequency decomposition*

Although dynamic functional connectivity methods applied to fMRI can capture the temporal dynamics of functional connectivity, they are inherently limited by the low temporal resolution of fMRI acquisition (commonly one data point per 2 seconds). Electroencephalography (EEG) is another modality that records brain activity with a

high temporal resolution from a set of predefined number of electrodes/sensors spatially distributed across the scalp. EEG is believed to record postsynaptic potentials from deep layers of pyramidal cells (Cohen, 2017). Importantly, the high temporal resolution of EEG acquisition is able to capture the different frequencies at which neurons oscillate. Neuronal networks in the human brain demonstrate oscillatory activity believed to serve a variety of cognitive processes (Hutcheon and Yarom, 2000; Buzsáki and Draguhn, 2004). Neuronal oscillations measured by EEG usually cover a broad frequency spectrum. Current literature has defined canonical frequency bands as 1-4Hz (delta), 4-8Hz (theta), 8-12Hz (alpha), 12-30Hz (beta), and >30Hz (gamma) oscillations (Niedermeyer and da Silva, 2004).

### *EEG networks*

Neuronal oscillations at different parts of the brain can be coherent either in terms of amplitude or phase and are believed to reflect communication. By oscillating at a specific frequency a neuronal group rhythmically opens the group's windows for communication with another neuronal group oscillating at the same frequency, and closes it for another neuronal group oscillating at a different frequency (Fries, 2005). This is regarded as the functional coherence between different brain regions and is a similar concept to functional connectivity between different BOLD signals.

In EEG studies, measures of signal coherence across electrodes are used to capture brain communication at different frequencies (Nolte et al., 2004). Similar to functional connectivity networks, EEG networks can be used to show how coherence is organized across the brain (Micheloyannis et al., 2009). As the definition of coherence is frequency-specific, EEG networks at different frequencies can represent brain connectivity that corresponds to different communication mechanisms.

### *EEG and loss of consciousness*

Under several consciousness frameworks, integration of information across neural systems in different frequencies plays an important role in the emergence of conscious states (Crick and Koch, 2003). Studies have shown that scalp EEG during conscious, waking behaviour demonstrates low-amplitude “desynchronized” patterns (Berger,

1929). However, during loss of consciousness these patterns oscillate at alpha frequencies and become coherent especially in the forebrain (Purdon et al., 2013). Simulations have shown that this is associated with coherent alpha oscillations between the thalamus and frontal parts of cortex (Contreras et al., 1996; Brown et al., 2010; Ching et al., 2010). These alpha, coherent oscillations in the thalamocortical circuitry have been discussed in the context of consciousness because it is thought that awareness depends partly on how frontal and thalamic areas oscillate at these alpha rhythms (Llinás et al., 1998). Thus it is now believed that a relationship between the loss of consciousness and alpha oscillations exists (Gugino et al., 2001; Hughes and Crunelli, 2005). Expanding on this idea, studies using EEG connectivity have shown alterations in alpha frequency networks in anaesthesia (Chennu et al., 2016) and disorders of consciousness (Chennu et al., 2017) with network becoming more fragmented and less globally efficient. Thus the potential of using EEG networks can be important for further understanding of changes in functional connectivity during loss of consciousness otherwise not observed with fMRI techniques.

## **1.5 Complexity**

### **1.5.1 Rationale**

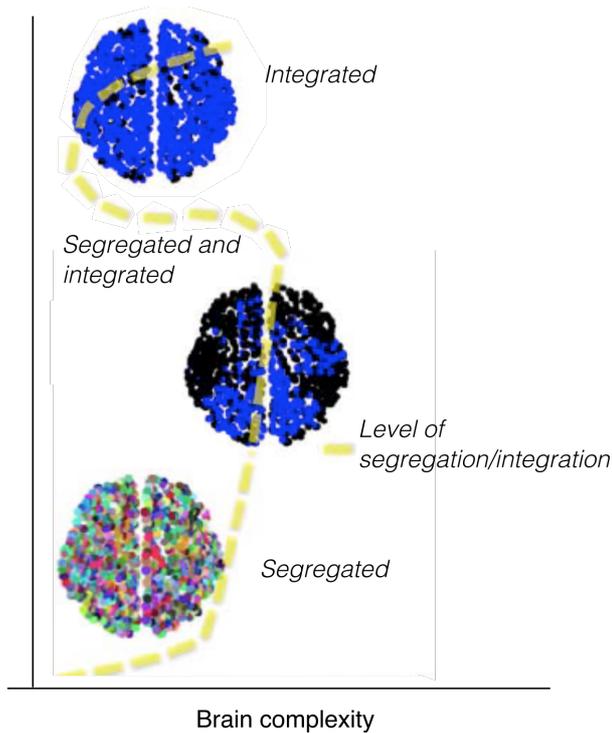
The previous parts suggest that in order to advance the neuroscience of consciousness one could start from existing theories to inform the design of potential measures of consciousness such as the ones that build upon resting-state functional connectivity. How can one make this transition feasible? The described theories of consciousness in part 1.3 converge to a certain point: conscious processing requires neural activity in segregated networks and the projections among them in order to integrate information. This balance between functional integration and functional differentiation otherwise known as brain *complexity* is an important measure that can quantify the state of consciousness (Tononi et al., 1998; Tononi and Edelman, 1998; Sporns et al., 2000). Towards this direction, theoretical and practical indices based on this principle have been designed to assess the joint presence of differentiation and integration in neural systems (Tononi et al., 1994). These metrics are only applicable, however, to simple systems of simulated elements or under highly restrictive

assumptions and have not been fully translated to testing human brain complexity. At the same time, in part 1.4 I argued that functional connectivity and graph theoretic approaches could provide a suitable scaffold for examining neural interactions in the brain as a whole. Taken together, these two remarks provide the framework for a testable hypothesis: *The complexity of functional connectivity networks can be used as a measure of the states of consciousness*. In that regard, I introduce a novel framework by examining the complexity of functional connectivity and how it changes during loss of consciousness.

### 1.5.2 Defining complexity

Complexity, in simple terms, quantifies the diversity and non-uniformity that arises from the interactions of similar units (Bak and Paczuski, 1995). When applied to networks (where units correspond to nodes and interactions correspond to connections), complexity can be regarded as an aggregate measure of the richness (or repertoire) of network organization as it looks at the (statistical) *distribution* of network connectivity across different regions (Zamora-López et al., 2016).

Complexity also captures the coexistence of integration and segregation by looking at the distribution of connectivity of specialized regions and how these integrate to shape network communication (Tononi et al., 1994; Sporns et al., 2000; Zamora-López et al., 2016). If a network is composed of functionally segregated elements, then integration is low and complexity is biased towards a more segregated configuration. On the other hand, if the system shows cooperative behaviour at the global level, then integration is high and the system has a complexity biased towards a more integrated configuration (Fig. 1.2). The complexity repertoire in between these two ends defines a landscape of network organization associated with different levels of segregation and integration. Complexity in the human brain reflects a specific balance of these two as functionally specialized groups of neurons distributed across the brain are integrated to produce coherent information (Tononi et al., 1994; Tononi et al., 1998). The importance of brain complexity lent the idea of developing practical concepts and measures for quantifying it, as I will describe more extensively in the next sections.



**Figure 1.2 - Different levels of segregation/integration and brain network complexity.** When each functional system processes information in a specialized way and independently from other systems, then brain is in segregated configuration. On the other hand, if functional systems show increased cooperative behaviour then they communicate information at the whole-brain level and the brain is at an integrated configuration. Brain network complexity is believed to be at a “sweet” spot by keeping a balance between segregation and integration. Adapted from Tagliazucchi and Chialvo (2012).

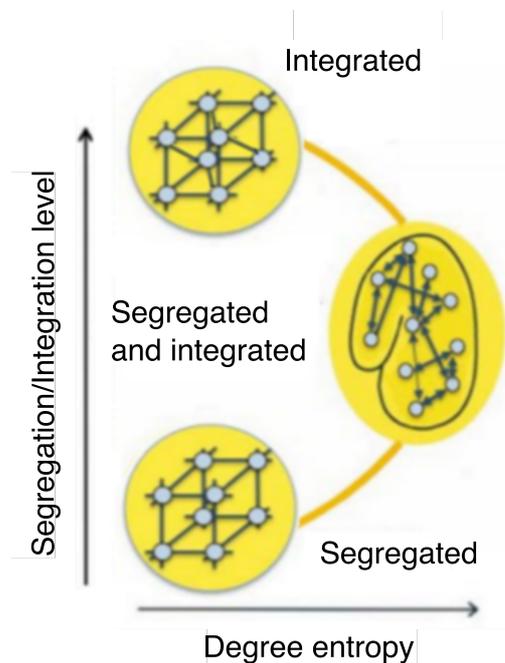
### 1.5.3 Measures of complexity

Previous theoretical and practical research towards quantifying brain complexity has led to the development of different measures. Some examples are neural complexity (Tononi et al., 1994), causal density (Seth et al., 2011),  $\Phi$  from integrated information theory (Tononi, 2004; Balduzzi and Tononi, 2008; Oizumi et al., 2014), stochastic integrated information (Barrett and Seth, 2011) and Lempel-Ziv complexity (Casali et al., 2013). Many of these measures quantify complexity as follows. First, they define a partition of the brain network into segregated modules alongside their associated complexity characterized in terms of statistical connotation (for example distribution of connectivity in some specified neuronal groups). Second, they calculate integration of these modules in order to look at how these integrate to produce a complex brain network (Tononi et al., 1994). However, due to the vast number of possible partitions, depending on what is considered a segregated module, this approach makes the

computational cost for calculating complexity extremely high for networks of large size.

To this end, one can overcome the obstacle of calculating segregation and integration over multiple partitions by looking at the randomness/predictability of the whole system (van Emden 1975; Grassberger, 2012), i.e. when talking about brains, by looking at how predictable/homogenous the distribution of brain connectivity is. A principled way of capturing brain complexity, in light of this randomness approach, is by quantifying the entropy of the degree distribution (Zhao et al., 2011; Mowshowitz and Dehmer, 2012). Entropy is a measure from information theory and captures the randomness/predictability of the degree distribution by looking at the coexistence of hubs and low-degree regions—a crucial feature of segregation and integration in functional connectivity networks (Bullmore and Sporns, 2012; Zamora-López et al., 2016). Thus the previous repertoire of complexity can be now translated to a network degree entropy repertoire: from homogeneously connected networks/fully segregated networks to random networks/fully integrated networks. Specifically, a homogeneously connected network is a network in which all nodes have similar numbers of edges and connected with their nearest neighbours in a lattice-like structure. Because every node has approximately the same number of edges, the entropy of the degree distribution is low. At the same time, few edges are needed to communicate between nearby nodes, but many edges would be required to communicate with distant parts of the network i.e. it supports more segregated communication. The other extreme is the random network, in which edges are distributed randomly throughout the network. In this case, degree entropy is also low as one can easily predict their degree distribution. At the same time, in a random network information could reach distant nodes using a small number of edges, but reaching neighbouring nodes requires many more edges than on a lattice-like network. Thus a random network is biased towards an integrated configuration. Above and beyond these two configurations, brain networks keep a *balance* of segregation and integration and have specific degree entropy that is higher and more unpredictable compared to homogeneously wired or random networks i.e. they show a

“sweet spot” of complexity (Tononi et al., 1998)(Fig. 1.3). I will return to this point in the next section.



**Figure 1.3 - Network degree entropy as a measure of brain complexity.** When each functional system is uniformly connected to each other, then the degree distribution is homogenous thus having low degree entropy. Similarly, if functional systems show a very high number of connections between each other, then they are connected in random fashion-thus networks have low degree entropy. When a balance between integration and segregation exists, entropy is high as the degree distribution is highly unpredictable with the existence of a diverse pattern of connected systems. Adapted from Tagliazucchi and Chialvo (2012).

#### 1.5.4 Criticality of complexity

Not any level of complexity can sustain conscious processing (Tononi, 2004). For example homogeneously wired networks can perform less operations and do not possess the same capacity for information as brain networks (Brunel, 2016). Brain network complexity is critical in the sense that it can foster the richness of conscious experience through a specific balance of segregation and integration of information i.e. there is a particular “sweet spot” of complexity at which the brain functions (Tononi and Edelman, 1998; Deco and Jirsa, 2012). By this token, alterations in complexity signify departure from criticality that can potentially lead to brain dysfunction (Shew and Plenz, 2013; Tagliazucchi et al., 2016).

How does one show that complexity of functional connectivity is critical? Theories on criticality come from statistical physics. Critical systems exist on a state of “bistable” transition i.e. a product of competition between ordering and disordering collective tendencies (Christensen and Moloney, 2005). At that state, the parts of the system give rise to the system’s macroscopic properties that cannot be understood by considering only one part’s behaviour. Thus criticality can be viewed as the cooperative behaviour of all parts of the system and how this gives rise to system properties such as complexity (Bak, 1996). By transposing these ideas to brain networks, one can argue that the cooperative behaviour of brain regions can be associated with the emergence of segregation and integration tendencies and how these compete and balance each other to produce a complex brain (Tononi et al., 1998).

Previous studies have shown that complexity of functional connectivity is derived from criticality by making a link to the brain’s underlying structural network (Tagliazucchi and Chialvo, 2012). It has been shown that if structural connectivity is endowed with certain dynamical properties reflecting these “bistable” transitions, then one can reproduce resting-state functional connectivity at a critical point of complexity (Sporns et al., 2000; Haimovici et al., 2013).

Expanding on this approach, recent studies argue that the criticality of complexity can be shown if one understands the cooperative dynamics developed upon the underlying structural communication paths (Avena-Koenigsberger et al., 2017). Towards this direction, complexity is critical because it uses structural paths that promote communication information in a critically balanced and efficient way. Thus in the first experimental chapter of this thesis (chapter 3), I provide evidence that the complexity of functional connectivity in the healthy brain is critical because it is developed upon a balance (Nash equilibrium) of “optimal” structural connections. These are connections that represent the most efficient structural communication paths in the brain. Thus the criticality of complexity of functional connectivity can be explained on the premises of a balance of optimal connections that together support efficient communication in the brain from the local/regional level to the global/whole-brain level i.e. a balance of segregation and integration.

### 1.5.5 Loss of consciousness and reduced complexity

Alterations in complexity would imply departure from the balanced integration and segregation. My hypothesis is that any *decrease* in complexity would signify loss of consciousness. Why is this the case? Decrease in complexity can affect consciousness by biasing brain networks towards a more segregated/less integrated configuration. When integration is lost, specialized neuronal groups cannot communicate rapidly and effectively thus information cannot be consciously accessed (Tononi and Edelman, 1998; Tononi and Koch, 2015). Initial evidence supporting this statement comes from magnetoencephalographic (MEG) studies of binocular rivalry indicating that awareness of a stimulus occurs when increased integration between brain regions takes place (Tononi et al., 1998). In addition, the GNW model suggests that information from functionally specialized regions fails to reach the global workspace (i.e. become conscious) unless it is sufficiently integrated via long-range excitatory axons (Dehaene and Changeux, 2011). The hypothesis that a segregated configuration is evident of the unconscious brain also finds fruitful ground in theories regarding the importance thalamocortical connectivity in maintaining consciousness. Due to its extensive connectivity with the rest of the brain, alterations in thalamic connectivity cause the brain to become more fragmented, thus impairing whole-brain integration (Alkire et al., 2008). Together, the previous remarks suggest that a decrease in complexity might provide a marker for loss of consciousness in light of the loss of the brain's ability to integrate information. In-depth investigation of this hypothesis will be performed in chapter 4.

### 1.5.6 Temporal complexity

Similar rationale is followed for understanding the temporal changes in the properties of functional connectivity networks. Instead of focusing on functional connectivity networks and their properties at each time point, in chapter 5, I consider the distribution of functional connectivity across time (Baker et al., 2014). This distribution will reflect the probability of the brain being active at different states that I will derive from the BOLD data (Vidaurre et al., 2017a). I use this approach to characterize dynamic changes in functional connectivity as occupancy and switching

between different brain states, thus providing the means for exploring complexity of functional connectivity in the temporal domain.

### **1.5.7 Development and complexity**

Finally, developmental changes show that network properties that have been observed to change during adult loss of consciousness cannot be directly translated to infant loss of consciousness. The lack of development in the functional connectivity of the infant brain is believed to restrict the infant brain from attaining a state of consciousness similar to the adult brain (Rochat, 2003). However, certain theoretical accounts suggest that, although consciousness in infants is not similar to that of the adults, it can still be conferred, to some extent, to the same balance of segregation and integration (Lagercrantz and Changeux, 2009). Investigation of this hypothesis necessitates a separate study focusing on the effects of anaesthesia in the complexity of the developing brain. Thus, in chapter 6 I will investigate the complexity of functional connectivity networks derived from EEG data in infants undergoing surgical anaesthesia. My goal is to demonstrate that decreases in complexity can discriminate between anaesthetized infants and infants emerging from anaesthesia in a similar way as in the adult anaesthesia.

Taken together, extensive investigation of these hypotheses not only provides evidence for the importance of complexity in the neural correlates of consciousness, but also paves the way for future studies that can focus on clinical measures and applications for classifying different states of consciousness.

## Chapter 2: Methods

In this chapter I briefly outline the methods used in the experimental chapters of this thesis. Details for specific analyses and datasets used are provided in the corresponding experimental chapters.

### 2.1 Principles of fMRI

Functional connectivity will be extensively examined in this thesis as a means of understanding loss of consciousness. Functional connectivity is based on computing BOLD signal correlations. BOLD signals are obtained from fMRI data using MRI scanners. Here I describe the technology and methods involved in obtaining such data.

#### 2.1.1 MRI Physics

The origins of MRI go back to Nobel Prize winning discovery by Felix Bloch in 1946 on the properties of the atomic nucleus (Bloch et al., 1946). He was the first to show that a charged particle such as the hydrogen atom possesses a magnetic field due to its spin around the axis. Based on this property, nuclear magnetic resonance spectrometers were developed to study the molecular basis of materials by looking at how local magnetic fields affect atomic nuclei. However, it was not until the 1970s where it was realized that different tissues in the human body have different magnetic resonance profiles (Damadian, 1971). MRI scanners were developed in order to produce images of the human body. Such scanners contain a large wire coil that can produce high magnetic fields through the application of an electrical current. In clinical and research settings the most commonly used scanners are limited to 1.5 or 3 Tesla of magnetic field strength.

Most of the human body is made up of water molecules which contain hydrogen. The images produced by an MRI scanner rely on the specific properties of the protons in the nucleus of the hydrogen atom. Based on Bloch's theories (Bloch et al., 1946),

protons can be viewed as dipole magnets with an electrical charge and spin. In normal conditions, the protons in our body spin at different directions in a manner that equalizes their magnetic forces. Under the presence of a magnetic field of constant strength  $B_0$ , protons jump into two, so-called energy states: one parallel to the magnetic field (low energy state) and one anti-parallel to the magnetic field (high energy state). As lower energy is parallel to the direction of the field, a slight majority of protons are parallel to the magnetic field. Because protons are unable to completely align with the magnetic field, they precess about their axis at a characteristic frequency known as the Larmor frequency. Each precession is characterized by a longitudinal and a transverse component showing the different directions of the precession. The net magnetization of all the protons is the sum of the longitudinal and transverse components. The net longitudinal magnetization is the sum of the magnetizations of the protons parallel to the magnetic field minus the sum of those that are in the anti-parallel direction. The net transverse magnetization component relates to whether protons precess in the same phase. Net transverse magnetization of out phase protons is zero because protons cancel each other out whereas the net transverse magnetization of “in phase” protons is non-zero (Blink 2004; Diechmann 2009; Stamatakis et al., 2017). I will simply refer to the transverse magnetization of “in phase” protons as the transverse magnetization.

The net magnetization that occurs when the states of the spins are at equilibrium cannot be measured directly. What can be measured however is how net magnetization changes after it has been perturbed. To perturb the two magnetizations, a magnetic field  $B_1$  that pulsates at a specific Larmor frequency is applied i.e. a radiofrequency signal (RF signal). This frequency is defined as

$$\omega_0 = \gamma B_0$$

where  $\gamma$  is the gyromagnetic ratio (the ratio between the magnetic moment and angular momentum of a spin). An explanation for why the RF pulsates at this frequency will be provided later. In a typical MR experiment, a 60-90° pulse is applied, which shifts the magnetisation vector from the longitudinal plane (parallel to  $B_0$ ) to the transverse plane. Consequently, protons jump to the antiparallel direction (loss of

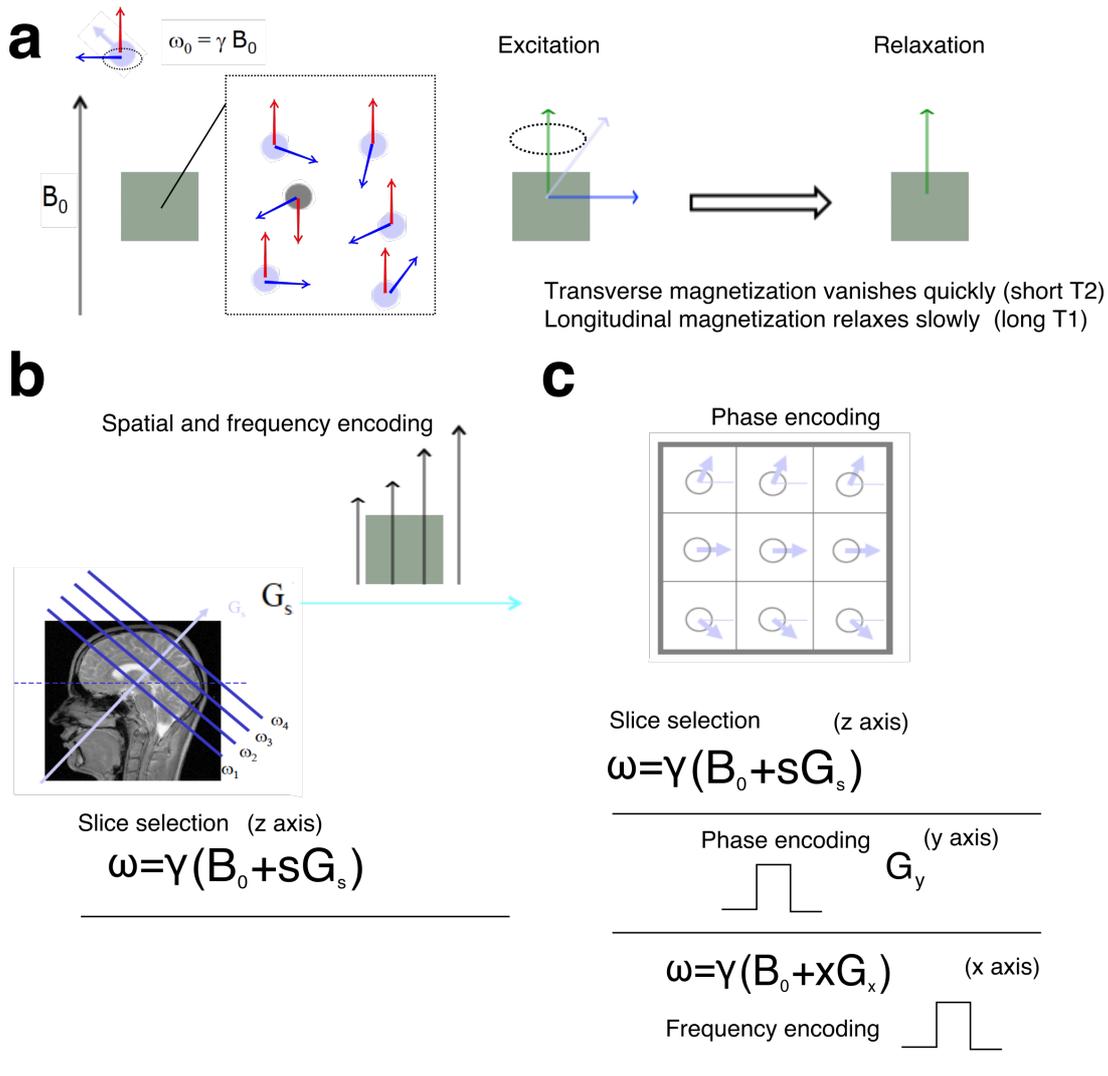
longitudinal magnetization) and they become coherent with the RF signal (increase in transverse magnetization). After the RF pulse stops, protons fall out phase corresponding to loss of transverse magnetization and protons lose their additional energy corresponding to recovery of longitudinal magnetization. These two processes take place at different time scales due to several environmental factors (Blink 2004; Diechmann 2009; Stamatakis et al., 2017). Most relevant for fMRI is the loss of the transverse component because it is recorded as an electromagnetic signal by using a receiver coil in the transverse plane (MR signal). If the excitation frequency of the RF pulse equals the Larmor frequency  $\omega_0$ , then the signal has an exponential decay. The pace at which the transverse magnetization is lost (T2) is fast due to inhomogeneities in the magnetic field and other effects. This is usually quantified by using a decay constant called T2\*. Quantitatively, the relationship between T2 and T2\* is given by

$$1/(T2^*) = (1/T2) + (1/T2')$$

where T2' represents the dephasing effects caused by magnetic field inhomogeneities (as those induced by  $B_0$  or from magnetic fields with varying strength/gradients that will be introduced later). The lost signal can be recovered using a new application (echo) of the RF signal. However, tissue-specific factors also contribute to signal loss, and their effect cannot be entirely reversed by the new application of the RF signal. In addition, some tissues produce larger and some produce smaller differences between the original and the echo signal intensities. It is on this basis that different tissues can be imaged. The effect of decay transverse magnetization is called transverse relaxation and depends on the tissue type. If the RF signal is applied in sufficient time intervals (TE or echo time), the intensity will be different in different tissues (Blink 2004; Diechmann 2009; Stamatakis et al., 2017). The recovery of T1 magnetization also depends on tissue and environmental factors and can be used for imaging different types of tissue. T1 magnetization is enhanced when successive excitation occurs at time intervals not long enough to allow full recovery of longitudinal magnetization, i.e. short repetition times (TR).

The T2\* transverse magnetization decay depends on tissue susceptibility and, most importantly for fMRI studies, on the magnetic properties of haemoglobin. Since the

seminal work of Ogawa and colleagues, it was realized that exploiting the magnetic susceptibility of haemoglobin could provide a measurable signal at regions where neuronal activation would occur (Ogawa et al., 1990). The haemoglobin molecule has magnetic properties that differ depending upon whether or not it is bound to oxygen (Pauling and Coryell, 1936). Deoxygenated haemoglobin has paramagnetic properties meaning that it causes a loss in the MR signal around its area. On the contrary, oxygenated haemoglobin has diamagnetic properties meaning that it does not affect the MR signal. However, if deoxygenated haemoglobin results in signal dropout how can we obtain increased signal when neuronal activation occurs? When neuronal activation occurs within a region this causes an increase in cerebral blood flow and the use of glucose but *not* a commensurate increase in the oxygen consumption rate (Fox and Raichle, 1986). This results in a decreased oxygen extraction and lower deoxygenated haemoglobin content per volume unit of brain tissue (Fox and Raichle, 1986). Thus there is a decrease in the relative amount of deoxygenated haemoglobin and, due to the magnetic properties of deoxygenated haemoglobin, there is a higher T2\* signal intensity (Huettel et al., 2004). Therefore, at a sufficient TE, increased signal will be obtained where there is high neuronal activity and lowered signal where there is low neuronal activity. This change in intensity is an indirect measure of neuronal activity or what is referred to as the blood oxygen level dependent (BOLD) signal (Poldrack et al., 2011).



**Figure 2.1 - Principles of MRI image acquisition.** (a) Under the presence of a homogenous magnetic field with strength  $B_0$  the protons align themselves to the magnetic field in a parallel (low energy state) or antiparallel manner (high energy state). Because they cannot align fully, they precess about their axis at a specific frequency  $\omega_0$  (Larmor frequency) that depends linearly on the strength of the magnetic field by the gyromagnetic ratio ( $\gamma$ ). Longitudinal magnetization is due to a difference in the number of spins in parallel and anti-parallel state. Transverse magnetization is due to spins getting into phase coherence. Applying an RF signal with a specific frequency  $\omega_0$  excites the protons to jump to a higher energy state (excitation phase). During excitation, longitudinal magnetization decreases and transverse magnetization appears. After excitation, transverse magnetization vanishes quickly (T2) while longitudinal magnetization (T1) recovers slowly (relaxation phase). Due to tissue-induced inhomogeneities in the magnetic field, these two processes allow imaging different types of tissue. The T2 decay is used to produce a measurable signal (MR signal) that is useful for functional imaging of neuronal activity. (b) The superimposition of other magnetic fields with varying magnetic strength allows parsing the signal from the three dimensions of the brain. A magnetic field with inhomogeneous intensity across different slices can be applied in the brain ( $G_s$  gradient) to encode spatial information

at each slice of the brain (z-axis). (c) In turn, an additional gradient-named phased encoding gradient ( $G_y$  gradient)-is applied that forces protons to precess at the same frequency but different phases. This allows encoding protons along the second dimension of the brain (y-axis). Finally a frequency-encoding gradient ( $G_x$  gradient) is applied to force protons to precess with different frequencies thus allowing encoding protons along the third dimension of the brain (x-axis). Parts of the figure was adapted from the MRC-CBU tutorial slides available here (<http://imaging.mrc-cbu.cam.ac.uk/methods/IntroductionNeuroimagingLectures>).

### 2.1.2 Spatial encoding

One remaining problem is that of obtaining spatial information that can be eventually transformed to a 3-dimensional image. Under the presence of the magnetic field  $B_0$ , different parts of the brain will have different Larmor frequencies slightly deviating from  $B_0$ . Image formation in MRI requires the addition of spatially *varying* magnetic fields, known as gradients that will cause spins at different locations (in the three dimensional space of the brain) to precess at different frequencies and phases.

The first gradient applied pertains to encoding the protons in specific slices of the brain i.e. encoding the z-axis of the three dimensional brain (slice-selection gradient). The scanner can select the particular slice to image by turning on the slice-select gradient and then altering the frequency of the RF signal to match the frequency at the desired slice position. Protons not in the slice will not get excited since their Larmor frequency will not match the frequency of the pulse and they will not jump to a higher energy state. After applying the slice-selection gradient (in other words knowing one spatial dimension), two further gradients are applied: one for phase and one for frequency encoding. The phase encoding gradient forces protons at all y locations to precess at the same frequency but different phases thus allowing encoding along the y-axis. The frequency gradient is forcing to protons to precess at different frequencies in the x-axis thus allowing encoding their x spatial position. Together, these two gradients encode the protons in the two remaining spatial dimensions (Blink 2004; Diechmann 2009; Stamatakis et al., 2017). The total MR signal recorded associated with different frequencies and phases is mathematically represented in k-space. An inverse Fourier transform is then used to convert k-space data into images that we can

later process. As mentioned before, when this sequence ends, one needs to wait for the longitudinal magnetization to partially recover and reapply this sequence (TR, repetition time). Because this sequence is rather slow, advanced sequences like the echo-planar imaging (EPI) have been developed to obtain functional data faster (Mansfield, 1977). Moreover, EPI is susceptible to local changes in blood oxygenation due to the paramagnetic properties of haemoglobin, making it an ideal candidate for functional MRI (Ogawa et al., 1990).

Taken together, the described procedure can result in either T1-weighted (i.e. relying on T1 relaxation) with more anatomical structural information or T2-weighted (i.e. relying on T2 relaxation) images that are more suited for fMRI.

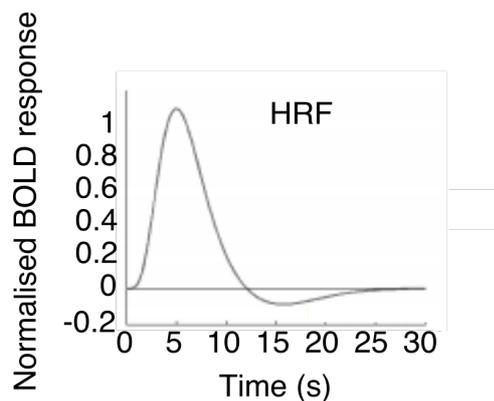
The effective spatial resolution of fMRI is at the order of millimetres and is related to the specificity of the spatial encoding gradients described above. Spatial resolution is described by what is known as “voxel size” meaning the size of the rudimentary volume elements or pixels that comprise the three-dimensional MRI images. The temporal resolution is limited by the delay between the neuronal activation and the peak of the neuronal response as well as the temporal resolution applied in MR sequences. In practice, there are trade-offs between spatial resolution and temporal resolution that can determine the final image size and resolution. Overview of the MRI acquisition principles and spatial encoding is described in Fig. 2.1.

### **2.1.3 BOLD signal**

The dynamics of neuronal activity and their relationship to oxygen consumption and the BOLD signal is not straightforward. In this section I will describe shortly our understanding of the BOLD signal.

During neuronal activation the brain reacts by increasing local blood flow, thus bringing in more oxygen and causing hyperoxygenation. Simultaneously, the local blood volume increases thus increasing the oxygen consumption and the level of deoxyhaemoglobin. These two processes are non-commensurate meaning that the increase in deoxyhaemoglobin is slower than the increase in local blood flow, resulting

in a net hyperoxygenation that lasts for about 4-6 seconds, and in an increased  $T2^*$  intensity, thus a positive BOLD response. Subsequently, the oxygen consumption and blood flow returns to pre-activation conditions. Because blood flow response is slower, there is an undershoot in the observed BOLD signal after the event that triggered the haemodynamic response. The dependence of BOLD signal on time after neuronal activation is called the haemodynamic response function (HRF) (Poldrack et al., 2011, Fig. 2.2). The HRF is slow, with a peak around 5 seconds followed by a 15-20 seconds undershoot until it returns to baseline. In addition, the HRF can be treated as a linear time-invariant system. In other words, the response to a long sequence of neuronal activity can be determined by linearly adding the response to shorter trains of neuronal activity. This is particularly useful in explaining the statistical models and image processing employed later (Poldrack et al., 2011).



**Figure 2.2 - A double-gamma model of the haemodynamic response function.** In response to a stimulus, the corresponding BOLD signal shows an initial dip, followed by a peak at around 5 seconds, and a prolonged return to baseline, that lasts at around 30 seconds.

However the BOLD signal is a rather indirect measure of neuronal activity, being dependent on the coupling of vascular activity to metabolic demand that, in turn, is linked to neuronal activity (Logothetis et al., 2001). Current models suggest that presynaptic release of glutamate induces a BOLD response that generates local field potentials in post-synaptic neurons, meaning that BOLD signals reflect pre-synaptic activity, not post-synaptic firing (Goense and Logothetis, 2008). The post-synaptic targets include glial cells, the messenger systems of which cause arteriolar muscle to relax and initiate haemodynamics. However different accounts suggest that local blood flow can be altered prior to neuronal activation perhaps in anticipation of a

potential stimulus-response (Sirotin and Das, 2009). Although the controversy regarding the origins of BOLD signal is an important topic, in this thesis I will use the BOLD signal under the assumption that it is well-described proxy of neuronal activity.

#### **2.1.4 fMRI experiments**

Understanding BOLD responses associated with given experimental conditions consists the major aim of fMRI experiments. FMRI experimental designs can be broadly categorized in two categories: i) categorical referring to those experiments that compare conditions by subtracting one condition from another ii) parametric referring to the experiments looking for alterations in BOLD responses based on tailored experimental stimulus. Depending on stimulus delivery, there are broadly three subcategories of fMRI experiments: blocked, event-related, mixed blocked and event-related, and participant-dependent designs (Amaro and Barker, 2006). Blocked designs include experiments that compare task blocks to non-task blocks in order to isolate the BOLD response related to a certain task. Event-related experiments allow a more refined temporal decomposition of the BOLD usually in relation to a particular event of interest (for example correct responses vs. erroneous ones). In mixed blocked and event related experiments, the previous two types can be inter-mixed to produce an even finer temporal decomposition of BOLD response, by averaging responses during events and contrasting them between task and non-task blocks. Finally, a participant-dependent design is used in cases where the experimenter has no control, such as resting-state scanning. In summary, the major criterion in choosing a particular experimental design depends on the question of interest. In this thesis I focused only on resting-state scanning in order to see how spontaneous BOLD responses during rest can inform on the state of consciousness in anaesthetized individuals and patients.

## 2.2 Preprocessing fMRI data

Prior to any fMRI statistical analysis a series of preprocessing steps are required in order to ensure the validity of the results.

### 2.2.1 Slice-timing correction

Because fMRI data is acquired slice by slice, there are slight differences in the timing the slices were obtained at within a single image. Thus it is the case that a functional slice is obtained at a point that is later in the HRF compared to other slices. Using interpolation techniques, slice-timing correction ensures that every point in a given functional image is the signal from the same point in the HRF (Sladky et al., 2011).

### 2.2.2 Realignment/motion correction

The realignment process corrects head movement or other types of movement caused by respiration or cardiac signals. Correcting for motion is important as it can lead to false activations as well as spurious functional connectivity estimations (Parkes et al., 2018). The correction algorithm deals with the displacement of the image in 6 directions, x, y, z translations and x, y, z rotations. For this purpose a rigid-body transformation is used. Rigid-body transformations assume that the size/shape of the two images that are to be co-registered are identical (it is the same brain). One image can be superimposed upon the other by a combination of three translations and three rotations (Ashburner, 2009). Realignment is usually done in two steps: First, images are realigned to the first image and a mean image of these is computed. Second, images are realigned again in a second pass to the mean image. Computer algorithms can identify the set of parameters that provide the best match to the reference image by using a cost function (for example using a voxel-by-voxel intensity subtraction criterion) (Ashburner, 2009).

Even *after* realignment a considerable amount of the variance in the data can be accounted for by the effects of movement. Thus the translational and rotational (motion) parameters used to realign the images are incorporated as confounds in any statistical models used later.

### 2.2.3 Co-registration and normalization

The next preprocessing step involves the spatial transformation of fMRI images to a standard space, a process known as spatial normalization. In order to allow for within and between group comparisons the individual functional images need to be warped to a standard Euclidean space (Fox, 1995). The two most widely used spaces in the neuroscience community are Talairach (Talairach and Tournoux, 1988) and Montreal Neurological Institute (MNI) spaces (Evans et al., 1993). The Talairach space is based on a stereotaxic atlas of the human brain (postmortem brain of a single subject) published by Talairach and Tournoux, whereas the MNI space is based on an average brain template coming from the MRI scans of several hundred individuals.

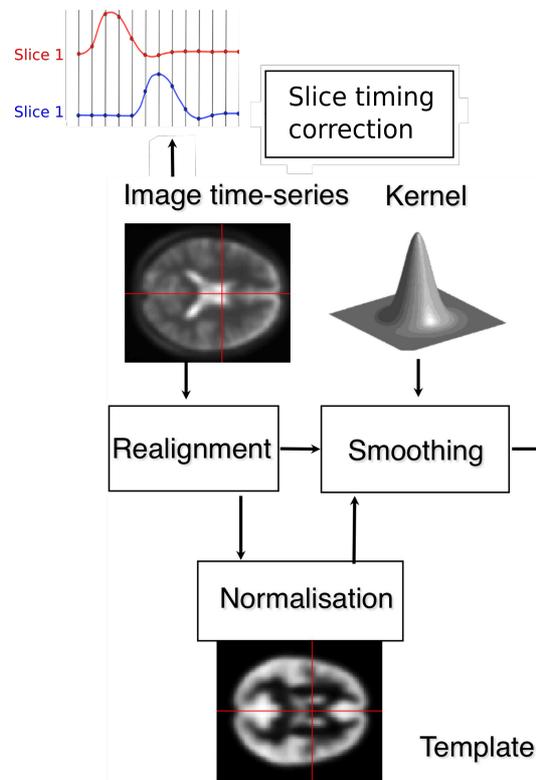
Since structural images such as T1-weighted images contain more spatial information (higher resolution) they are utilised when spatially normalising fMRI images. A preliminary step to assist with this process is to transform (or co-register) the T1 structural image of each individual to their lower resolution, mean functional image discussed earlier. This registration is similar to the rigid-body transformation employed for motion correction but it requires a more complicated cost function due to the different nature of these images (high-resolution structural versus low-resolution functional (Wells et al., 1996).

In turn the co-registered structural image is segmented into different tissue types (grey matter, white matter, cerebrospinal fluid/CSF). The grey matter segmentation, alongside tissue probability maps in standard space, is used for better normalisation of the co-registered structural image. The normalisation parameters estimated are then applied to realigned functional images. This has been labelled as the “unified segmentation method”, in which case the segmentation and normalisation stages are combined into a single process (Ashburner and Friston, 2005).

### 2.2.4 Smoothing

After normalization, a standard process is to smooth the functional data. This is done in order to normalize the statistical distribution of the data at the group level, thus

increasing the likelihood of finding meaningful statistical results (as in the case of parametric statistics that are often used in fMRI analysis). Usually a 3-dimensional Gaussian kernel with a specified full-width (FWHM) in mm is used (Ashburner, 2009). Overviews of the preprocessing pipeline is described in Fig. 2.3.



**Figure 2.3 - Preprocessing functional MRI data.** Slice-timing correction accounts for the fact that each slice in a functional slice is obtained at a point that is later in the HRF compared to other slices. After slice-timing correction, functional images are realigned within each subject to account for motion correction. The images are then normalized to a standard template in order to wrap all individual brains to the same space. Finally images are smoothed using a kernel (here a 3-dimensional Gaussian kernel is shown with a specified full width) in order to normalize the data for subsequent cross-subject comparisons.

### 2.2.5 Functional connectivity

Functional connectivity regards the temporal correlations between the BOLD data derived from different seeds or regions (Biswal et al., 1995). As addressed in the introduction, the functional connectivity work presented in this thesis considered a whole-brain approach by regarding the time courses from a number of regions of interest (ROIs)/parcellation spanning the entire brain. After defining ROIs (I refer the reader to the experimental chapters for more details on which parcellation was used),

sophisticated toolboxes such as the CONN toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2011) can be used to extract and correlate BOLD signals from each ROI using the preprocessed functional data. Prior to this and besides the aforementioned preprocessing steps, the functional data requires additional processing in order to deal with noise and spurious correlations. In that regard, it is common that the preprocessed functional data are highpass filtered for dealing with scanner noise. Physiological and motion related noise are dealt with by using the anatomical *CompCor* technique, which removes the first 5 principal components of the signal from white matter and CSF masks, as well as the motion parameters and their first-order temporal derivatives (Behzadi et al., 2007). Functional connectivity values are then derived as Pearson's  $r$  pairwise linear correlations between the BOLD signals of each ROI and are  $z$ -fisher transformed in order to make the distribution of correlations normal. This process results in functional connectivity data for each individual.

### **2.2.6 Dynamic functional connectivity**

The previous analysis of functional connectivity assumes that the strength of interactions between regions is constant over time. As described previously, correlation coefficients between regions are calculated from the time series of the entire scan. Although a convenient framework, given the dynamic nature of brain (Rabinovich et al., 2012) it might be of interest to obtain a more refined representation of the *temporal* changes in functional connectivity (Hutchison et al., 2013). The idea that the temporal analysis of functional connectivity might provide information about the underlying brain dynamics comes from studies combining fMRI and high temporal resolution modalities such as EEG (Allen et al., 2012). These studies have shown that functional connectivity changes are commensurate to changes in the electrophysiological recordings indicating a temporal variability on a time scale of seconds to tens of seconds (Chang et al., 2013).

In that regard, the question of what is the most appropriate method in order to measure the temporal variations of functional connectivity gains high importance.

The majority of literature in this area uses the “sliding window” method to estimate variations in functional connectivity (Hutchison et al., 2013). By using this method, functional connectivity is estimated within pre-specified and overlapping “windows of time” that have fixed length. Expanding on this framework, other methods looking at windowed connectivity but not requiring *ad hoc* definitions of windows and window lengths have been developed. For example, dynamical conditional correlation analysis uses maximum likelihood methods to estimate the appropriate parameters that can inform us on time-varying variances and correlations between BOLD signals (Lindquist et al., 2014).

These are important studies and have provided useful evidence for the brain’s rich dynamics. However, building upon literature from other modalities having much higher resolution than fMRI (for example EEG “microstates”, Khanna et al., 2015 or attractor neural networks, Conklin and Eliasmith, 2005), new methods have emerged that talk about a richer repertoire of dynamics not previously seen by windowed connectivity. Broadly speaking, these methods suggest that brain dynamics are characterized by a certain number of discrete *states* that recur during different time points during the scan (Baker et al., 2014). These models also assess the *transitions* between these states allowing us to quantify the brain’s switching rate between different states (Baker et al., 2014).

### ***Hidden Markov model***

One representative model that is going to be used in this thesis to study the dynamics of loss of consciousness (chapter 5) is the Hidden Markov model (HMM)(Vidaurre et al., 2017a; 2017b). Compared to other models that also assume the existence of discrete states (for example as these derived from some form of clustering analysis, Barttfeld et al., 2015a), this model was chosen due to its sophistication on capturing the occurrence of states and the between-state transitions, its tractable inference, and its reduced computational workload.

The HMM uses the BOLD data obtained using standard preprocessing methods. Using notation, HMM assumes that for the BOLD data of length  $T$  and for an *ad hoc*

defined number of states  $K$ , there is a number of hidden state variables  $s = \{s_1, \dots, s_T\}$  (and their parameters  $\theta$ ) that describe the dynamics of the BOLD data.

A posterior distribution model is used to represent this assertion. It assumes that the BOLD data  $y$  at each time point is obtained as the joint probability of being on each state and their relative transition probabilities between them. Formally

$$P(y, s, \theta) = P(s_0 | \pi_0) \times \prod_t^T P(s_t | s_{t-1}, \pi_t) P(y_t | s_t, \theta) P(\pi_t) P(\pi_0) P(\theta)$$

where  $P(\pi_t)$ ,  $P(\pi_0)$ ,  $P(\theta)$  are chosen to be non-informative priors,  $P(s_0)$  is the initial state probability,  $P(s_t | s_{t-1}, \pi_t)$  is the state transition probability, and  $P(y_t | s_t, \theta)$  is the posterior observation probability. The HMM parameters  $\Theta = \{\pi_t, \pi_0, \theta\}$  consist of  $\pi_0$  which parameterizes the initial state probability,  $\pi_t$  that determines the  $K \times K$  state transition probability matrix, and  $\theta$  which describes the observation probability. The state transition probability is assumed to depend only on the previous state (Markovian property) i.e.

$$P(s_t | s_1, \dots, s_{t-1}) = P(s_t | s_{t-1}) = \pi_t$$

The observed model probability for each state  $k \in K$  is assumed to follow a Gaussian distribution with mean  $\mu_k$  and covariance  $\Sigma_k$  that represent the activity and connectivity of each state respectively. This can be denoted as

$$P(y_t | s_t = k, \theta) \sim \mathcal{N}(\mu_k, \Sigma_k)$$

The prior distributions over the HMM parameters  $\theta = \{\pi_0, \pi_t, \theta\}$  and the posterior distributions are chosen to be conjugate, i.e. they come from the same type of probability distribution (usually Gaussian). This choice makes the model tractable to certain kinds of inference.

The next goal is to infer the hidden variables and their parameters. One method for doing this is to use variational Bayes (VB) inference that deals efficiently with complicated probabilistic models (Wainwright and Jordan, 2008). Using VB one can obtain the hidden variables and parameters and, in turn, the posterior observed probability  $P(y_t | s_t = k, \theta)$  that fully describes the BOLD data. Practically speaking, the inference allows us to compute at which state each time the brain is and the state's parameters. This gives us the methodological toolbox to quantify useful quantities about the temporal dynamics of the brain: for example, how much time is spent on each state or how often does the brain switch between states. Details will be given in chapter 6.

### **2.2.7 Software**

The aforementioned preprocessing pipeline for all the experiments described in this thesis was implemented in the Statistical Parametric Mapping (*SPM*) Version 12.0 toolbox publicly available here

(<http://www.fil.ion.ucl.ac.uk/spm/>)

The toolbox utilized the *MATLAB* Version 12a platform

(<http://www.mathworks.co.uk/products/matlab/>)

Functional connectivity was calculated using the *CONN* toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012) publicly available here

(<http://www.nitrc.org/projects/conn>)

The HMM model was built on top of the *HMM-MAR* toolbox (Vidaurre et al., 2017a; 2017b) publicly available here

(<http://github.com/OHBA-analysis/HMM-MAR/wiki>)

### **2.3 Principles of diffusion imaging**

Diffusion imaging estimates white matter local diffusion direction from measurements of water diffusion. Importantly, white matter tractography pieces together this information to infer structural connections between brain regions. Thus diffusion imaging and tractography offer unique avenues for understanding the brain's structural connectivity. In chapter 3, I used structural connectivity in order to provide a suitable model for explaining the critical complexity of functional connectivity. Thus in this section I will describe the basic premises of diffusion imaging and structural connectivity.

#### **2.3.1 Diffusion in the brain**

Diffusion imaging profiles the motion of water molecules in the brain. Each molecule experiences a “random walk,” known as Brownian motion, as it moves through a substrate. Though the motions of individual molecules are unpredictable, the position of the molecules on average can be modelled according to Einstein's law (Einstein, 1956). However, Einstein's law works only in the case where the motion of the molecules is unimpeded i.e. isotropic diffusion. If molecules are impeded in one or more directions, the diffusion is anisotropic and Einstein's law fails to predict the position of the molecules. Examples of anisotropic diffusion can be found in the human body, with the most relative to the brain example being the diffusion in the axons. Specifically, because of the hydrophobic and tubular structure of the axons, water diffuses much more rapidly along the axon than across its membrane (Sen and Basser, 2005).

#### **2.3.2 Diffusion weighted imaging**

Imaging the diffusion of water molecules created new challenges for MRI. To sensitize MRI images to diffusion a modified MR sequence needs to be used. As introduced previously with MR imaging, an adequately applied strength-varying magnetic field (or gradient) can influence the phase of the spins, with the degree of influence

depending on the strength of the field. Obtaining diffusion weighted imaging (DWI) data requires an additional gradient in order to encode the diffusion information (Hagmann et al., 2006). This gradient consists of two pulses separated by a certain time interval in the acquisition process (diffusion time interval). The first of the two gradient pulses in this sequence introduces a phase shift depending on the strength of the gradient at the position of the spin. Before the application of the second gradient pulse, which induces a phase shift dependent on the spin position, a  $180^\circ$  RF pulse is applied to reverse the phase shift induced by the first gradient pulse. Since the diffusion-encoding gradient causes the phase shift to vary with position, the spins that remain in the same location along the axis of the gradient during the two pulses will return to their initial state. However, spins that have shifted will experience different field strength in the second pulse. Thus they will not return to their position but they will experience a total phase shift resulting in a decreased intensity of the measured MR signal. The longer the displacement distance is, the higher the phase shift and the greater the decrease in signal will be. Hence, the resultant image shows low signal intensity in regions where diffusion along the gradient is high (Hagmann et al., 2006).

Similar to the k-space of the MRI data, the result of the application of gradients in different directions and with different intensities at specific moments of the acquisition, results in different values of the signal that can be represented in a coordinate system (q-space). Inverse Fourier methods applied to the q-space result in images depicting diffusion in different directions at each brain position. To describe the parameters applied in sampling from the q-space, the term “b value” is often used. The b value is proportional to the product of the diffusion time interval and the square of the strength of the diffusion gradient (Hagmann et al., 2006).

### 2.3.3 Calculating diffusion directionality

DWI allows the estimation of the diffusion direction in each voxel in the brain. One particular problem DWI faces is that it is not able to fully capture the diffusion characteristics. Diffusion is a three-dimensional phenomenon with a direction and shape. In that regard, diffusion can be described by a  $3 \times 3$  diffusion tensor  $D$  that fully

characterizes diffusion in 3-dimensional space, under certain Gaussian assumptions (Batchelor et al., 2005).

$$D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}$$

where each number  $D_{ij}$  corresponds to the diffusion rates in each combination of direction  $(i, j)$ . If the diffusion data is sampled in sufficient directions, then a diffusion tensor model can be fitted to the diffusion data by solving a model tensor equation with 6 degrees of freedom (equal to the number of free variables in the diffusion tensor). Diffusion tensor imaging (DTI) is a specialized diffusion imaging technique that implements this approach. In DTI, diffusion gradients are applied in multiple directions to fully sample the diffusion tensor in space. Specifically, a series of diffusion-weighted images are obtained in which the q-space is sampled in at least six (equal to the number of unknowns in the diffusion tensor  $D$ ) different directions plus a non-diffusion-weighted reference image ( $b_0$ ). In general, a b value of approximately 1000 sec/mm<sup>2</sup> is used for sampling from the q-space. It is worth noting that, whereas only 6 encoding directions are needed to estimate this diffusion vector, a larger number of spatially uniformly distributed encoding directions is usually acquired in order to mitigate the noise effects on the diffusion parameters. One major limitation of DTI is its inability to describe fibre directionality in regions where two or more fibre populations with different orientations are present (e.g., crossing fibres regions). The fibre crossing confound in DTI has prompted efforts to develop diffusion imaging methods capable of resolving fibre crossing (Alexander, 2011). Using q-space imaging (QSI), investigators have found that in regions of fibre crossing the diffusion function possesses significant multimodal structure (Tuch et al., 2003). They have suggested that, in order to resolve this complicated diffusion process, a gradient sampling on a three-dimensional lattice is required (Tuch et al., 2003). However, there are some disadvantages associated with these techniques. First QSI is time-intensive. Second, QSI requires large pulsed field gradients in order to perform sufficient sampling of diffusion. To address the sampling burden of QSI, investigators have proposed an alternative approach based on sampling on a spherical shell (or combination of shells) in a so-called diffusion wave-vector space, an approach

referred to as high angular resolution diffusion imaging (HARDI) (Tuch et al., 2002). Methods such as q-ball imaging (QBI), have been developed to reconstruct diffusion directionality from HARDI data (Tuch, 2004). Other approaches that also use sophisticated tensor models to deal with fibre crossing include diffusion spectrum imaging (DSI) (Wedeen et al., 2008) or hybrid diffusion imaging (Wu and Alexander, 2007).

### 2.3.4 Tractography and structural connectivity

Once the diffusion tensor is obtained from the data, quantitative measures can capture the degree of anisotropy of the diffusion. One such measure is the fractional anisotropy (FA), a value between 0 and 1 that is calculated using the eigenvalues of the diffusion tensor (Alexander, 2011). A value of zero means that diffusion is unrestricted in all directions whereas a value of one means that diffusion is restricted among all directions. Values in between 0 and 1 imply that diffusion is restricted among specific directions.

Thus, FA estimates the principal direction of diffusion in the 3-dimensional space. In that regard, white matter tractography techniques estimate the connectivity patterns between different brain regions from the continuity in the local estimates of diffusion direction at each voxel. These algorithms include the selection of a starting point or seed, which can be either a voxel or a precise location defined by Cartesian coordinates in the brain space, and then iteratively reconstruct structural connectivity by estimating the diffusion directions towards other regions or voxels. Tractography algorithms can be classified largely into deterministic and probabilistic categories. Deterministic algorithms provide a unique trajectory for each seed point (Jones et al., 2005) whereas probabilistic algorithms generate multiple possible trajectories and select the “best” fit using a cost function (Koch et al., 2002). The limited number of diffusion gradients and degrees of freedom make tractography prone to error especially in the case of crossing fibres. To this end, more advanced methods have been implemented that deal with the uncertainty of fibre detection more efficiently (Behrens et al., 2007).

It is worth noting that the fibres quantified with tractography are often considered to represent individual axons or nerve fibres. However, a more precise account would regard these as lines of fast diffusion that follow the local diffusion maxima and that only generally reflect the axonal architecture. Notwithstanding these limitations, tractography algorithms have been extensively used these in order to assess white matter connectivity between different parts of the brain (Lazar, 2010).

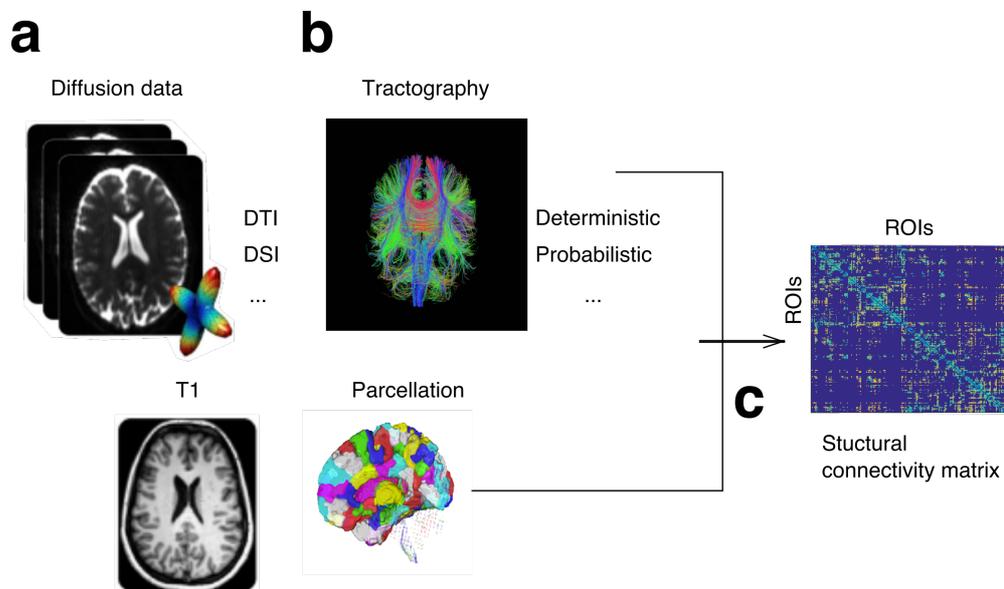
Together, diffusion data in combination with tractography algorithms can provide information about the structural connectivity between brain regions. An overview of this process is described in Fig. 2.4.

### 2.3.5 Software

I used *dsi-studio* (Yeh et al., 2013) publicly available here

(<http://dsi-studio.labsolver.org/>)

in order to perform diffusion image preprocessing and tractography.



**Figure 2.4 - Diffusion image acquisition and structural connectivity.** (a) Diffusion data is obtained by exploiting the diffusion properties of water along the white matter axons. Diffusion can be described by a diffusion tensor that fully characterizes the diffusion direction in 3-dimensional space. Based on the complexity of the diffusion tensor, different techniques are applied in order to reconstruct diffusion

directions in white matter (DTI, DSI, etc.). (b) White matter tractography links this information together to infer structural connectivity between different brain locations. Tractography algorithms construct diffusion-based trajectories from one brain region to another potentially resembling the underlying white matter connections. (c) Tractography methods can be combined with a brain parcellation (for example, as the one that comes from T1 data) in order to produce structural connections between ROIs.

## 2.4 Network analysis

### 2.4.1 Network properties

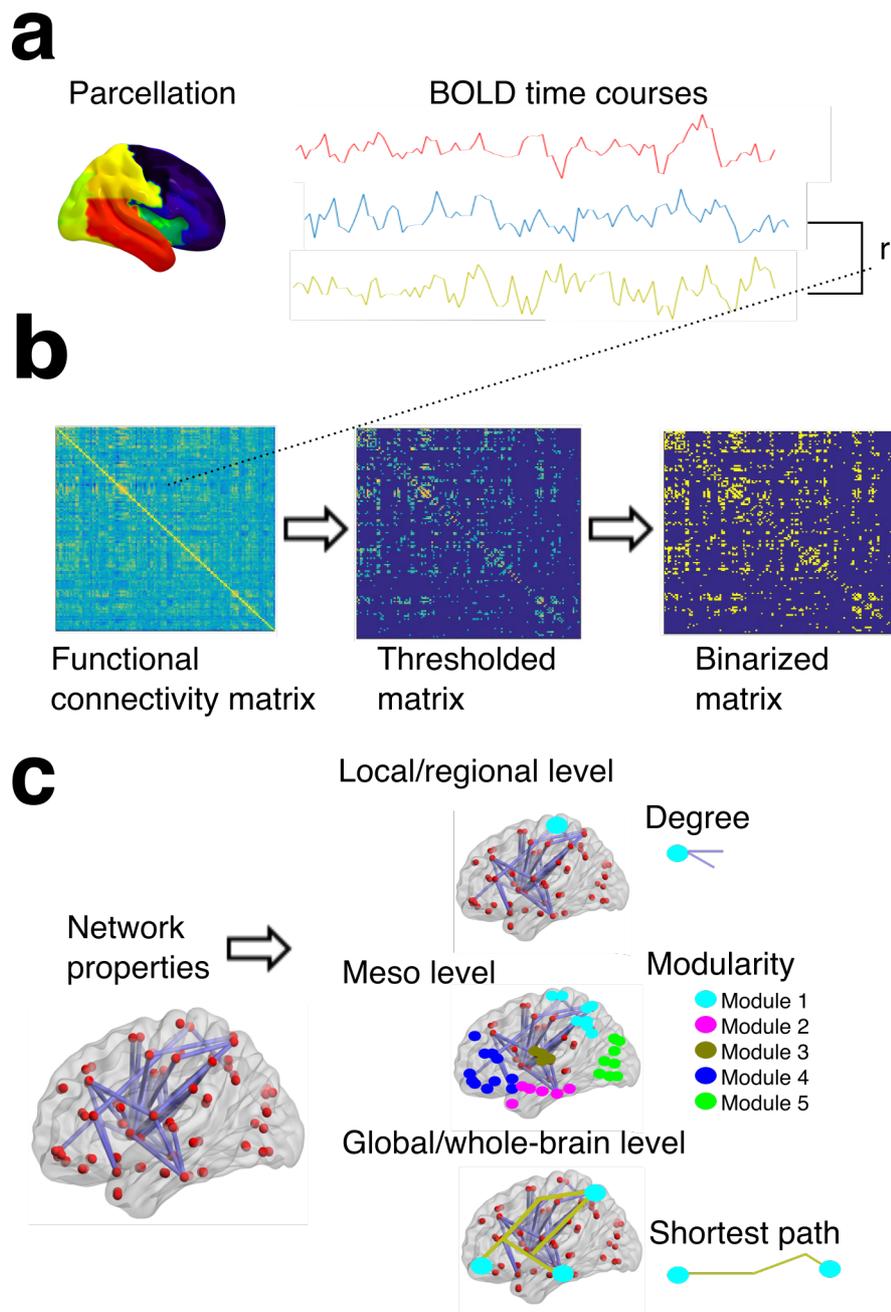
After computing functional connectivity values, I applied network analysis techniques in order to study network organization and its complexity. Here I will describe their basic premises.

A network (graph) is a collection of nodes and edges. Nodes can be defined arbitrarily (for example regions of interest or EEG electrodes/sensors) and edges stand for connections between nodes that can be either structural or functional (Fornito et al., 2016). A useful representation of a network is by using a two-dimensional matrix with the size of the two dimensions equal to the number of nodes. Each entry  $(i, j)$  in the matrix reflects an edge between nodes  $i$  and  $j$ . Prior to analysis of its properties, a network is usually thresholded. This is done in order to remove spurious edges as well as to make them sparse, a prerequisite for the computation of many network properties (Fornito et al., 2016). There is no consensus as to what is the most appropriate technique for thresholding a network. A popular technique uses the proportional thresholding technique meaning that a percentage of the most important connections are maintained for each individual. This way, the number of connections is maintained the same across individuals thus avoiding the problem of finding any statistical differences in network properties just because of the different number of edges. Choice of the percentage number is usually arbitrary. Studies suggest a lower threshold to prevent networks from being severely fragmented (Power et al., 2011) and an upper threshold to avoid introducing more, potentially noisy, edges into the network (Bassett and Bullmore, 2006). In turn, subsequent network properties are averaged over this range of thresholds. After thresholding, the resulting network matrix consists only of the strongest connections. Subsequently, this weighted

correlation matrix can be converted to a binary matrix in order to improve computational power.

After these processing steps, network properties can be investigated using tools from graph theory. The most fundamental property is the degree of the network representing the number of connections of each node towards the rest of the network (Rubinov and Sporns, 2010; van den Heuvel and Sporns 2013). Degree is a fundamental property of interest because it correlates with most of the network properties (Orsini et al., 2015). In addition, the shape of the degree distribution can discriminate between types of networks having different organization. For example networks that have random organization are fundamentally different in their degree distribution from networks that exhibit some form of clustering (for example modular or brain networks) (Newman et al., 2001; Vértés et al., 2012).

Other properties can be defined depending on which aspect of network organization is examined. For example, modularity is a measure expressing how well a network can be decomposed into different modules or communities. Modularity is defined as the number of links within a module over the links in the network. In order to identify network modules, different algorithms are applied usually by maximizing a cost function (Lancichinetti and Fortunato, 2012). Additional properties include the average shortest path length that is related to the global efficiency of the network in the sense that the smaller the shortest path length the quicker information will be transmitted globally in the network (Fig. 2.5).



**Figure 2.5 - Functional connectivity network analysis.** (a) BOLD signals from different ROIs are pairwise correlated to obtain a value  $r$  of functional connectivity between them. (b) Repeating this process for all the pairs of ROIs results in a functional connectivity network with nodes being ROIs and edges being pairwise correlations. A thresholding technique is applied in order to keep the strongest correlations and remove any spurious correlations. Depending on the analysis planned, the matrix can be also binarized in order to represent an unweighted network where the presence of an edge is denoted by 1 and its absence by 0. (c) After the network is obtained, different properties can be quantified. At the local/regional level, degree is the number of connections each region has towards the rest of the network. In turn, modularity is a measure of how well nodes are organized in modules that are strongly connected within them and sparsely connected with each other. At the global/whole-brain level the

length of the shortest path between every region in the brain is the minimum number of steps it takes to reach one region from another and is related to the global efficiency of the network.

### **2.4.2 Complexity of functional connectivity networks**

Complexity of functional connectivity networks is central to this thesis and will be introduced here.

#### ***From complexity to entropy***

Quantifying complexity in systems has their roots in the physics literature. Complexity has been used by physicists to ascribe a quantitative property to phenomena of which a complete description is infeasible due to their large number of parameters and properties (van Emden, 1975). Phenomena which in the physics literature are considered as complex are, among others, chaotic dynamical systems, fractals, networks and cellular automata. A common feature of these and other examples is the following: their properties show a specific pattern somehow situated between randomness and rigidity (Grassberger, 2012). In that regard, complexity measures have been developed in order to quantify how close a system is to randomness (Bai et al., 2014). A measure of randomness complexity, (or simply complexity) broadly used in physics and information theory is Shannon entropy (or simply entropy) (Shannon and Weaver, 1949). Increased entropy implies high complexity as the system cannot be easily predicted whereas low entropy implies less complexity and more predictability (Shannon, 1948). Formal definitions follow below.

#### ***Complexity of networks***

The notion of complexity (and its measure entropy) can be applied to networks, and, by extension, to functional connectivity networks. Since the initial investigations of network properties in functional connectivity networks, it has been shown that the shape of the degree distribution is not random. On the contrary, it emphasizes the coexistence of highly connected regions (hubs) and other sparsely connected regions, reflecting a balance between segregation and integration. Thus one way of capturing

the complexity of functional connectivity networks is by measuring the entropy of the degree distribution (Mowshowitz and Dehmer, 2012; Zamora-López et al., 2016). Formally, the degree distribution  $P$  can be obtained from the histogram of the degree sample of each network. In turn, entropy of the degree distribution can be obtained by calculating the following sum i.e.:

$$H = - \sum_{i \in P} p_i \log p_i$$

where  $p_i$  reflects the probability of a specific degree value occurring in the degree sample. Estimating entropy in an unbiased way, i.e. estimating the true entropy of the distribution based on a limited-size sample is an important problem with many applications (Harris, 1975). Standard estimators calculate entropy by using the frequency estimates of the sample. Let  $F = (F(1), F(2), \dots)$  the fingerprint of the sample of size  $k$  where  $F(1)$  is the number of degree values appearing once,  $F(2)$  is the number of degree values appearing twice and so on. Then entropy can be defined as  $H = - \sum_{i \in p_i} p_i \log p_i = - \sum_i \frac{i}{k} F(i) \log i/k$ .

In chapter 3, I will use entropy to show how complexity can be derived from structural connectivity. In chapter 4, I will refine quantification of entropy by focusing on the degree distribution of hubs and sparsely connected regions separately in order to see if there are more localized changes in complexity during loss of consciousness.

### 2.4.3 Software

I used the Brain Connectivity toolbox (BCT) publicly available here

(<http://sites.google.com/site/bctnet/>)

for thresholding networks and calculating their degree samples. In-house made scripts were used for additional network properties and for calculating entropy as described in the experimental chapters.

## 2.5 Electroencephalography

Electroencephalography (EEG) has been instrumental in assessing brain activity with high temporal resolution. In chapter 6, I used EEG-based network to study the complexity in anaesthetized infants. In this section I will describe the premises of obtaining and processing EEG data.

### 2.5.1 EEG signals

#### *Origin*

Neurons communicate information by means of electrochemical signals passing through the synapses (Hormuzdi et al., 2004). At the synapses, neuronal activity is transferred chemically from one neuron to another via neurotransmitters that are released from the presynaptic cell and attached to specific receptors located on the postsynaptic cells. This process induces excitatory and inhibitory postsynaptic potentials and associated currents around the synapses (Kirschstein and Köhling, 2009). Currents derived from synapses move through the dendrites and cell body to a trigger zone in the axon base and pass through the membrane to the extracellular space along the way. These extracellular currents generated by populations of neurons generate an electric potential (and electric field), a vector whose amplitude is measured in volts per distance. It is worth noting that the neurons that contribute to the EEG signal are mostly the pyramidal neurons of the cortex. These are arranged in palisades with the apical dendrites aligned perpendicularly to the cortical surface. In that regard, the cumulative summation of their electrical potential has a specific direction and, thus, is more easily measured in the scalp (Spruston, 2008). Taken together, due to their unique orientation and their increased number, the aggregate electrical field of pyramidal neurons in deep cortical layers reaches the scalp as a measurable signal (Buzsáki et al., 2012; Cohen, 2017).

### ***Recording***

Properly placed electrodes can record the EEG signal with submillisecond time resolution. EEG measurements employ recording systems consisting of electrodes with conductive media, amplifiers with filters, analog to digital (A/D) converters and recording devices. Electrodes read the signal from the head surface. When neuronal activity takes place, there is a potential change in the scalp that can be measured by using a basic electric circuit between the signal (active) electrode and a reference electrode. An extra electrode, called ground electrode, is needed for getting differential voltage by subtracting the voltages showing at active and reference points. In turn, amplifiers allow the microvolt signals to be accurately, digitized by enhancing the signal. Finally, a converter changes signals from analog to digital form, and a computer device stores and displays obtained data. Electrodes usually follow a conventional placement and cover frontal, temporal, central, parietal, and occipital parts of the scalp. Their number varies: from devices with few electrodes ( $< 10$ ) to high-density EEG devices ( $> 100$  electrodes).

### ***Preprocessing***

Raw EEG data is obtained with a high sampling rate (usually at 1024Hz). Due to inherit artefacts (eye blinks, line noise artefacts, etc.) and artefacts induced during recording (movement), a preprocessing pipeline is usually applied. Although there is no standard preprocessing pipeline, steps include downsampling (usually at 256Hz), filtering, line noise removal and noisy channel rejection (Puce and Hämäläinen, 2017). These methods can be complemented by visual inspection in order to assist with noisy data detection.

When the length of EEG data is sufficient, an additional method is used for artefact detection called independent component analysis (ICA). This method takes into consideration that the recorded EEG data comes from a number of linearly mixed independent components attributed to neurophysiological and non-neurophysiological sources. In that regard, ICA separates the EEG data by finding components that are maximally independent (Puce and Hämäläinen, 2017). If the

EEG data is noisy, then some of the components will correspond to non-neurophysiological signals such as eye blinks. Thus by visually inspecting each component, one can reject those that are related to non-neurophysiological signals (Puce and Hämäläinen, 2017).

### 2.5.2 Frequency analysis

The high temporal resolution of the EEG data allows decomposition of the EEG signal into different frequencies (spectral decomposition). Canonical frequencies, including gamma (>30Hz), beta, (12-30Hz), alpha (8-12Hz), delta (1-4Hz) and theta (4-8Hz), have been linked to distinct roles in brain function (Buzsáki and Draguhn, 2004; Uhlhaas et al., 2008). From Fourier analysis, a decomposition of a discrete-time signal- $x(t)$  (such as the EEG signal from one electrode) into different frequencies  $f$  can be obtained using its Fourier transform:

$$X(f) = \sum_{-\infty}^{+\infty} x(t)e^{-2i\pi ft}$$

Power at each frequency can be defined as  $S_{XX}(f) = X(f)X^*(f)$  where  $X^*(f)$  is the complex conjugate of  $X(f)$ .

Because data is recorded only for finite segments (i.e. for limited duration of EEG recording), only parts of the above transform can be calculated thus introducing bias in the calculation. Specifically, for a segment of limited duration  $T$  only the truncated transform can be computed:

$$X(f) = \sum_{t=1}^T x(t)e^{-2i\pi ft}$$

To this end, it has been shown that breaking the data into different parts (or “tapering” the data) and calculating the frequency decomposition at different chunks of time can reduce part of this bias (van Vugt et al., 2007). More refined approaches have been introduced that find the “optimal” tapering of the data such as the bias in estimating frequency decomposition is minimal (Bokil et al., 2010). These methods

find a certain number  $K$  of taper functions (Slepian functions) that maximize the frequency decomposition at each frequency i.e. they find a number of functions  $w$  such that their frequency decomposition  $U(f)$

$$U(f) = \sum_{t=1}^T w(t)e^{-2i\pi ft}$$

is maximal at each frequency  $f$ .

Then these functions are used instead of the previously mentioned Fourier decomposition to obtain a multi-taper frequency decomposition as follows

$$X_{k=1,\dots,K}(f) = \sum_{t=1}^T w_t(k)x(t)e^{-2i\pi ft}$$

where  $w_t(k)$ ,  $k = 1, \dots, K$ , are all the taper functions.

I will use this method to obtain a frequency decomposition of EEG signals in chapter 6.

### 2.5.3 EEG functional connectivity

Synchronization between the oscillatory activities of different brain regions is an important feature of brain communication (Fries, 2005). This is usually defined in terms of coherence (or connectivity) between EEG signals and it is measured by relating their spectral decompositions at each frequency  $f$  as

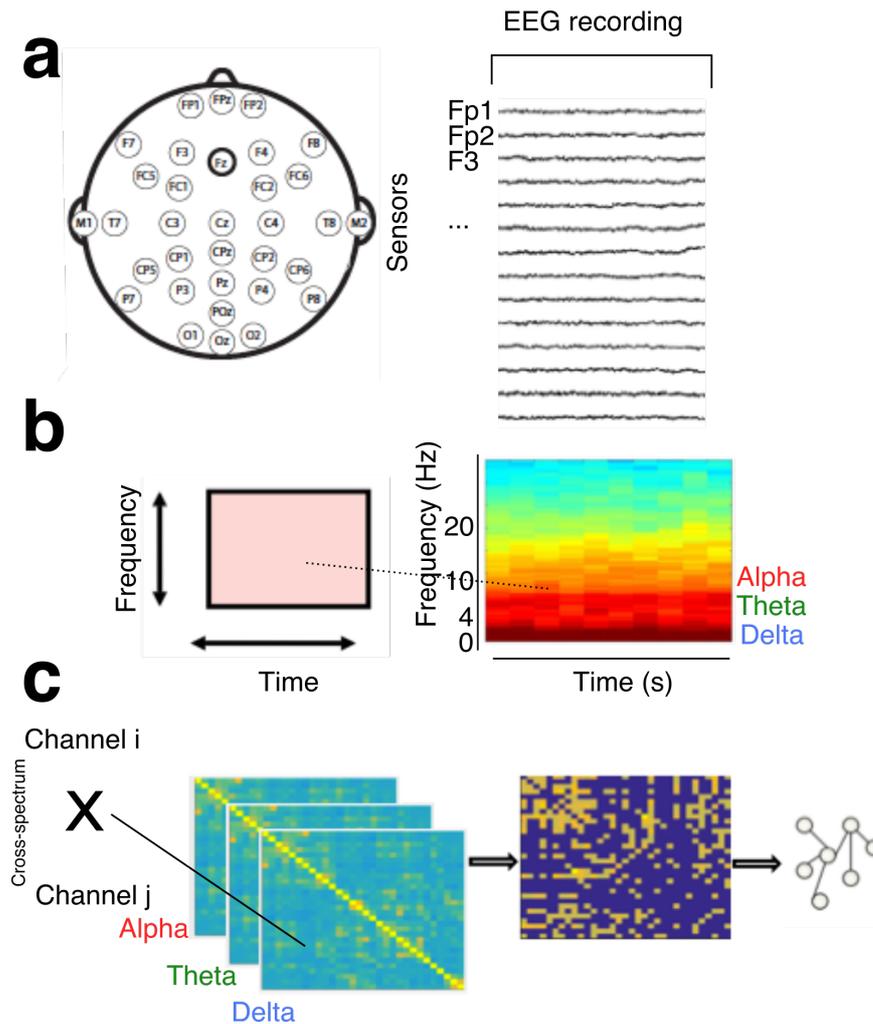
$$C_{XY}(f) = \frac{S_{XY}(f)}{\sqrt{S_{XX}(f)S_{YY}(f)}}$$

where  $S_{XX}(f) = \langle X(f)X^*(f) \rangle$ ,  $S_{YY}(f) = \langle Y(f)Y^*(f) \rangle$  represent the signal power (averaged over all the tapers of the data) of signals  $X$  and  $Y$  for all frequencies  $f$  and  $S_{XY}(f) = \langle X(f)Y^*(f) \rangle$  represents the cross-spectral power of signals  $X$  and  $Y$  (averaged over all the tapers of the data).  $C_{XY}(f)$  is a complex number (real and imaginary part) between 0 and 1 and can be interpreted as the frequency-equivalent of

correlation between the two signals  $X$  and  $Y$ . One issue for using this quantity as a measure of connectivity in EEG data relates to the so-called volume conduction problem. This refers to the case where two electrodes/sensors that measure the activity of the *same* neuronal source will show artificially high coherence (Stam et al., 2007). Measures such as the imaginary part of coherence have been used to address this issue (Nolte et al., 2004). Whole-brain EEG connectivity can be assessed by calculating the coherence between each pair of electrodes at each frequency. There are two particular disadvantages with respect to using the imaginary part of coherence that should be addressed here. First, it can be strongly affected by the phase of the coherence, which can lead to very small values when the sources of interest are nearly in phase (or in phase opposition)(Stam et al., 2007). Second, additional noise sources might change the sign of imaginary coherence. To overcome these, additional measures such as the weighted phase lag index have been proposed (Vinck et al., 2011).

However as addressed by the authors the use of either of these measures is pertinent to the quantity, quality, and amount of the data. For example, weighted phase lag index should be utilized when multiple sessions and trials are considered-something that did not apply to my resting-state EEG analysis; this is why EEG analysis was conducted using only the imaginary part of coherence. An important point made by the authors is that additional validation of results needs to take place at the source level-something that I took into consideration when examining EEG connectivity results at both the sensor and source levels.

Network analysis methods mentioned previously can be applied here. An overview of the methods used for EEG analysis and connectivity is given in Fig. 2.6.



**Figure 2.6 - From EEG data to EEG networks.** (a) EEG data is acquired at a high sampling rate from set of electrodes placed on the scalp. After preprocessing, different techniques are applied to obtain relevant information about the oscillatory activity and power of each signal. (b) Multi-taper time-frequency decomposition is one such method that extracts the oscillatory information at different tiles of time and frequency. This provides more accuracy in estimating oscillatory activity and power compared to traditional Fourier analysis methods. (c) The oscillatory activity of each signal can be correlated with that of another one at each particular frequency (cross-spectrum). Repeating this process results in EEG connectivity networks corresponding to each frequency. In turn, standard network analysis techniques (thresholding, binarizing) can be applied to study the organization of these EEG networks.

#### 2.5.4 Source reconstruction

Connectivity at the sensor level includes a number of electrodes and provides an indirect proxy for the connectivity of the underlying neuronal sources. To overcome the spatial limitations, source reconstruction methods have been applied in order to reconstruct activity at the neuronal source using EEG data. Quantitatively the

neuronal sources are represented as a 3-dimensional grid of dipoles derived from the individual's anatomical data (Jatoi et al., 2014). To reconstruct the sources' time-courses a mathematical relationship between the EEG data and the dipoles is required (Henderson et al., 1975). This is called the forward problem. To address the forward problem one needs to represent how electrical activity moves from the source to the scalp. The mathematics of this process takes into consideration the inhomogeneous and anisotropic conductivity profile of the head volume (head model). Towards this direction, obtaining accurate representation of the head model is important as different volume conductor tissues (e.g. scalp, skull, brain) have different conductivities and different levels of anisotropy. Standard head models for EEG forward modelling use image segmentations of the scalp, the skull and the brain and employ refined geometric descriptions (for example using a tessellation) in order to provide realistic bases of how electrical activity diffuses along these tissues (Kybic et al., 2005). To obtain these geometric descriptions, mathematical methods, such as the boundary element method (BEM), are applied on the segmented tissues (scalp, skull, brain) coming from each individual's T1 anatomical images (Akalin-Acar and Gençer, 2004). In addition to the head model, a source model is also obtained from each individual's T1 anatomical data. Specifically, the source model quantifies the position and orientation of the dipoles. Each dipole  $i$  is associated with two parameters  $\mathbf{s} = \{\mathbf{r}_i, \boldsymbol{\theta}_i\}$ : a three-dimensional vector  $\mathbf{r}_i$  representing its location and  $\boldsymbol{\theta}_i$ , a three-dimensional vector representing its orientation. These will be later inferred when source activity will be derived from the EEG data.

Based on the information coming from the head and source models, a mathematical model is constructed in order to relate EEG and dipole activity. One such model uses a collection of vectors called the leadfield matrix, of which each vector's strength and direction quantify how each dipole contributes to each sensor's activity. This can be written as

$$\mathbf{B}_s = \mathbf{L}_s \mathbf{J}_s + \mathbf{Y}$$

where  $\mathbf{B}_s$  is a  $e \times n$  matrix representing the EEG data with  $e$  being the number of electrodes and  $n$  being the number of time points. Each row in this matrix represents

each electrode's time course.  $\mathbf{J}_s$  is a  $d \times n$  matrix that represents the dipole data that need to be inferred with  $d$  being the number of dipoles. Each row in this matrix represents each dipole's time course.  $\mathbf{L}_s$  is the leadfield matrix with dimensions  $e \times d$  and it is used to relate the electrode and dipole time courses. Finally  $\mathbf{Y}$  is a  $e \times n$  noise matrix reflecting the uncertainty of the measurements.

Formulating the forward problem and obtaining the leadfield matrix allows us to *inversely* solve the problem for the time courses of the dipoles (Pascual-Marqui, 1999). In mathematical terms this can be written as an optimization problem:

$$\min_{\mathbf{s}} \|\mathbf{B}_s - \widehat{\mathbf{B}}\|$$

where  $\mathbf{B}_s$  is the EEG data as modelled using the leadfield,  $\widehat{\mathbf{B}}$  is the real EEG data obtained and  $\mathbf{s}$  represents the dipole parameters. One popular technique is beamforming where a single dipole's activity is obtained by looking at how it contributes to the measured EEG activity compared to other dipoles (Van Veen et al., 1997). The goal is to estimate the activity at a dipole of interest while avoiding the cross-talk from other dipoles i.e. avoiding the effect of other dipoles in the estimation of activity at the dipole of interest (Van Veen et al., 1997). To do so, one can use a spatial filter  $W$  that minimizes the contributions of other dipoles except the one from the dipole of interest. Formally a spatial filter  $\mathbf{W}_s$  is computed by

$$\mathbf{W}_s^T = \frac{\mathbf{L}_s^T}{\|\mathbf{L}_s\|}$$

where  $\mathbf{L}_s^T$  is the transpose matrix of the leadfield matrix  $\mathbf{L}_s$  and  $\|\mathbf{L}_s\|$  is its norm. Given the spatial filter and the solution to the optimization problem  $\mathbf{B}_s$ , the activity of each source can be calculated as  $\mathbf{W}_s^T \mathbf{B}_s$ . This approach guarantees that when only one dipole is active, the solution of the inverse model will be related only to that particular dipole by "spatially" filtering out the contributions of other dipoles (Darvas et al., 2004). One limitation of this approach is that the spatial filter has limited spatial resolution meaning that it can only block the contribution of a certain number of dipoles. Other approaches that can partially deal with the spatial filtering problem include the Linearly constrained minimum variance (LCMV) beamforming (Van

Veen, et al., 1997) or the Backus and Gilbert method (Grave de Peralta Menendez and Gonzalez Andino, 1999).

After obtaining time courses from all the dipoles, their cross-spectrum can be calculated as previously resulting in a source-based connectivity network. Furthermore, in order to obtain a more interpretable representation of cortical connectivity, source connectivity can be grouped into ROIs using a specific parcellation.

### 2.5.5 Software

The bulk of EEG analysis and source reconstruction was conducted using the following software packages (alongside in-house made scripts):

a) the *EEGLAB* toolbox (Delorme and Makeig, 2004)

(<http://sccn.ucsd.edu/eeglab/index.php>)

b) the *Fieldtrip* toolbox (Oostenveld et al., 2011)

(<http://www.fieldtriptoolbox.org/>)

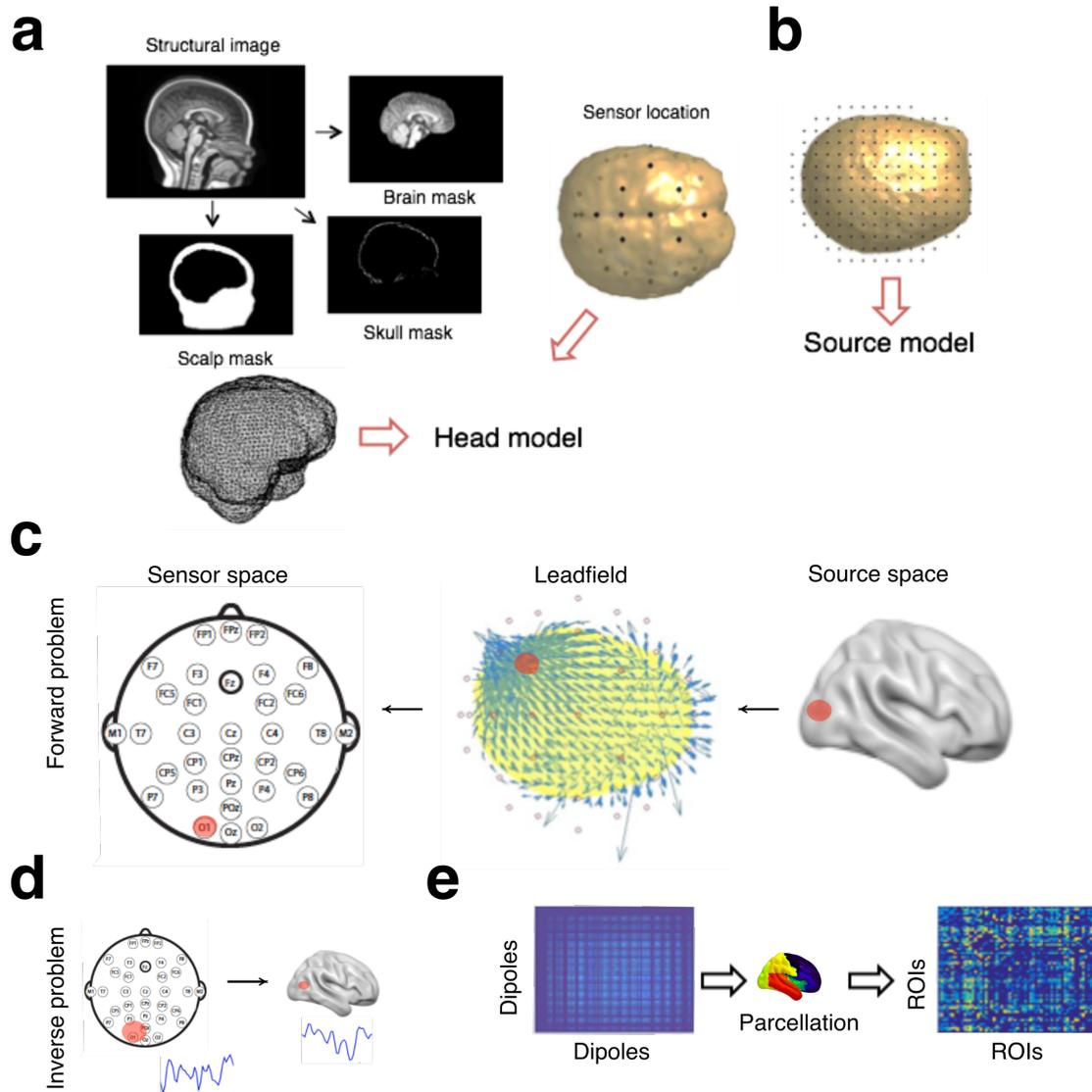
d) the *chronux* toolbox (Mitra and Bokil, 2008)

(<http://chronux.org/>)

### 2.6 Code availability

Code for in-house made scripts (as mentioned in this chapter and mentioned later in the experimental chapters) will be made available in my *github* website

(<http://github.com/ioannispappas322>)



**Figure 2.7 - Source reconstruction: from EEG data to cortical networks.** (a) The head model consists of scalp, skull and brain compartments obtained from each subject's anatomical data. The head model defines how electrical activity from the sources (or dipoles) is diffused along the head tissue and reaches the sensors. (b) The source model assumes that these dipoles are placed in a 3-dimensional grid in the brain. (c) The forward problem pertains to how source activity translates into sensor activity. This is done usually in the form of vectors (leadfield matrix) that use the head model, the source model and the sensors' positions. The vectors' locations and orientations relate how dipole activity explains sensor activity. For example, a dipole located in the occipital cortex will have leadfield vectors as shown in panel c and will mostly explain the contribution of occipital electrodes. (d) The inverse problem uses the leadfield matrix and the EEG data to obtain the dipole time courses. (e) After obtaining these, connectivity can be quantified as in the EEG sensor data by taking the dipoles' cross-spectrum at different frequencies. In turn, dipole networks can be transformed to ROI-based networks with the use of a specific parcellation. Due to the assumptions about how activity is measured from pyramidal cells usually only cortical dipoles (and regions) are considered.

## **Chapter 3: Complexity of functional connectivity emerges from optimal structure**

Since the early measurements by Biswal and colleagues (Biswal et al., 1995), it has been shown that functional connectivity (obtained using resting-state fMRI) is not random but it shows consistent spatio-temporal patterns. As I discussed in chapter 1, one of the major challenges is to understand the complexity of functional connectivity as it emerges from a balance between segregation and integration (Tononi et al., 1998). Under this framework, system criticality provides the theoretical foundation upon which one can understand how the opposite forces of segregation and integration co-exist in the brain (Tagliazucchi and Chialvo, 2012). These theories pertain to the *collective* behaviour of the system's constituent elements and how this macroscopically shapes the complexity of the system.

Towards this direction, studies have examined criticality in the brain on the basis of its relationship to how brain regions are structurally connected to each other (Park and Friston, 2013). In that regard, simulations have shown that structural connectivity in combination with certain dynamic models signifying a “bistable” state of segregation and integration can reproduce functional connectivity as the one obtained from real data (Rubinov et al., 2011; Haimovici et al., 2013).

Another avenue towards understanding the complexity of functional connectivity involves exploring the underlying structural communication paths and how these are intertwined to produce complex brain dynamics (Avena-Koenigsberger et al., 2017). In light of this, I propose the use of a model that links the complexity of functional connectivity network with a set of efficient communication paths in the structural connectivity network. Specifically, by using a game-theoretic model, I quantify the collective communication dynamics in terms of *Nash equilibrium* between efficient structural paths. In that regard, I show that the structural connectome contains a set of “optimal” structural connections that emerge from the equilibrium between each region's tendencies to efficiently communicate information to the rest of the brain. In turn, I show that optimal connections can predict the complexity of functional connectivity as it is obtained from real data. Thus I provide evidence for the criticality

of complexity of functional connectivity: it can only be derived upon the equilibrium of structural connections that support communication in an optimal way.

### 3.1 Introduction

Unlike other biological networks, brains are unique in the sense that their underlying static structural connections give rise to highly flexible and diverse functional connectivity patterns (Park and Friston, 2013). Understanding the complexity of functional connectivity, as a solution that stems from its underlying structural substrate, requires decoding the brain's communication model i.e. how structural pathways support efficient communication (Laughlin and Sejnowski, 2003). Recent advances in diffusion imaging techniques have assisted with precise quantification of the dense network of white matter pathways (Hagmann et al., 2008) while network analysis has shown that structural networks exhibit a topology characterized by reduced mean shortest path length and high clustering, features which potentially enhances efficient communication of information (Bullmore and Sporns, 2012). At the same time, functional connectivity shows systematic patterns that diverge from their structural counterparts, suggesting that its information can be communicated globally without necessarily following topologically short paths (Cole et al., 2014b). Thus the complexity of functional connectivity cannot be trivially derived based on the premises of structural underpinnings.

To this end, approaches to reconcile the discrepancy between functional connectivity and the underlying structural connectivity have focused on a search for the brain's underlying communication models. It is possible that communication models that allow for information exchange at an optimal level might provide the basis for explaining the versatility of functional connectivity (Avena-Koenigsberger et al., 2017).

Thus a central question arises: Can an optimal communication model explain the complexity of functional connectivity? Optimal models can be derived via optimization techniques or, when multiple elements/nodes are involved as in

biological networks, via game-theoretic approaches such as Nash equilibria between each element's competing strategies (Fabrikant et al., 2003). In that regard, optimal solutions are produced from the critically balanced (equilibrium) strategies between different players trying to maximize their payoff.

When brain regions are the metaphorical players, such strategies might be translated into how each region is going to connect to the rest of the brain in an optimal way in order to maximize communication efficiency. Previous studies have argued that wiring in the brain is optimal in the sense that it maximizes communication efficiency within certain anatomical constraints (Koulakov and Chklovskii, 2001; Chklovskii et al., 2002). In addition, efficient communication is essential in the sense that it allows the brain to adapt to varying functional demands and allocate resources between different brain regions (Avena-Koenigsberger et al., 2017). Thus, an optimal communication model in the brain might consist of a set of optimal connections, or an “optimal network”, reflecting the equilibrium between all regions' efficient communication strategies.

This provides a fruitful framework for studying functional connectivity and its complexity. An optimal communication model of structural connections might provide a suitable scaffold for explaining how this complexity is derived. Optimal connections reflect a balance between every region's efficient communication strategies suggesting that information upon these connections can be efficiently communicated from the local/segregated level to the integrated/whole-brain level. Notably, the notion that of an equilibrium implies that this complexity is critical i.e. it can result only from a certain configuration of optimal connections and other configurations fail to do so.

This work provides evidence towards the emergence of complexity of functional connectivity upon an optimal communication model using a multi-step approach. First, I identified optimal connections in the structural connectivity-connectomes of healthy individuals using a game-theoretic model (Gulyás et al., 2015). In turn, I verified that information communicated upon optimal structural connections explained functional connectivity and its complexity. I employed a suitable algorithm

that predicts functional connectivity by utilizing paths of the structural network (Becker et al., 2018) to show that the optimal structural networks contain core features essential for the more accurate prediction of functional connectivity compared to using the features of the entire structural network or other non-optimal connections. Then I showed that the complexity of functional connectivity of the predicted matrices using optimal structural networks is similar to the one obtained from the real data, indicating that it can be derived only upon these optimal connections.

Parts of this chapter have been presented as a poster at the Connectomics, Keystone Symposia conference (2017) in Santa Fe, USA.

### **3.2 Materials and Methods**

The Nash model was applied to structural connectivity data in order to see whether optimal structural connections existed in the structural connectomes of healthy individuals. In turn, these were used to predict functional connectivity data and its complexity in order to provide evidence for the emergence of complexity from the Nash equilibrium of optimal structural connections. Prior to describing the Nash model and the prediction methodologies, in sections 3.2.1, 3.2.2, and 3.2.3 I describe the structural and functional connectivity data used in this experimental chapter.

#### **3.2.1 Participants**

Data was provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Uğurbil; 1U54MH091657)(Van Essen et al., 2013) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Data was downloaded from the Human Connectome Project (HCP) website:

*(<http://www.humanconnectome.org/>)*

Structural and resting-state functional magnetic resonance images from a total of 50 subjects were used in the analysis, all of whom were between the ages of 22 and 35.

### 3.2.2 Structural connectivity

HARDI data was used for mapping diffusion in the brain (Sotiropoulos et al., 2013). Using HARDI data is advantageous for identifying complex axonal pathways because diffusion-weighted signal is acquired in more directions than usual DTI sequences, thus providing increased accuracy in determining white-matter pathways (Tuch, 2004). Data was acquired using a single shot single refocusing spin-echo, EPI sequence with a multishell diffusion scheme of 90 diffusion weighted directions with b-values of 990, 1985, and 2985  $sec/mm^2$  and spatial resolution of  $1.25 \times 1.25 \times 1.25mm^3$ . Preprocessing was conducted as part of the HCP pipeline and included eddy current and motion corrections, gradient non-linearity correction, and transformation to native structural space determined by the T1. The diffusion tensors were reconstructed in *dsi-studio* using generalized q-sampling imaging (GQI) (Chen et al., 2015). GQI is a model-free reconstruction procedure that quantifies the density of diffusion in different orientations, thus providing directional information regarding crossing fibres. *DSI-studio* uses the spin distribution function (SDF) that has greater sensitivity and specificity in fibre orientation (Yeh et al., 2010). A deterministic fibre-tracking algorithm was used to obtain Generalized Fractional Anisotropy (GFA) (which is a HARDI anisotropy measure similar to the popular DTI fractional anisotropy/FA) values between brain regions (Yeh et al., 2013). This method uses spherical harmonics and orientation distribution functions (ODF) to accurately characterize estimates of anisotropy (Cohen-Adad et al., 2008; Smith et al., 2013). Parameters of the fibre-tracking algorithm include the anisotropy threshold (how far tracking lines will be traced in each voxel), the angular ratio (the *a priori* knowledge of the tracking line curvature), and the step size (that represents in which consecutive voxels anisotropy is going to be estimated). The anisotropy ratio threshold was set to 0.0518329, the angular threshold was set to 60 degrees, and the step size was 0.625mm (half of the voxel size in one dimension).

To obtain ROI-based structural connectivity using the GFA values, I parcellated each subject's T1 using the Lausanne parcellation consisting of 234 or 129 cortical ROIs depending on the resolution (Hagmann et al., 2008; Daducci et al., 2012). This analysis was done using the *easy\_lausanne* software ([http://github.com/mattcieslak/easy\\_lausanne](http://github.com/mattcieslak/easy_lausanne)). For each ROI I also obtained the centre of mass coordinates in each volunteer's native space and the resulting Cartesian distances between ROIs (Fischl et al., 1999).

In turn, using the GFA values and the Lausanne parcellation, I created a  $234 \times 234$  or  $129 \times 129$  connectivity matrix depending on the resolution of the parcellation used. Each entry in the matrix contained either a positive number corresponding to the GFA (when a non-zero GFA between each pair of brain regions was obtained) or a 0 when a zero GFA value was obtained (Cammoun et al., 2012).

The reason that two different parcellation resolutions were chosen was to show the reliability of the proposed network model in capturing structural connections irrespective of the number of ROIs chosen. It has recently been suggested that the number of ROIs plays an important role in how network features are dispersed across the brain. Higher resolution connectome maps can potentially reveal characteristics that are more specific to functionally specialized regions whereas coarser resolution connectome maps can reveal a more abstract network organization (Reus and van den Heuvel, 2013). The goal was to verify that the network model could capture structural connectivity at both of these levels.

### **3.2.3 Resting-state fMRI and functional connectivity**

Whole-brain EPI data was acquired with a 32-channel head coil using a 3T Siemens Skyra scanner, modified for use in the HCP. The acquisition parameters were as follows: repetition time (TR) = 720ms, echo time (TE)=33.2ms, flip angle=52 degrees, bandwidth = 2290Hz/pixel, field of view =  $208 \times 180$ mm, slice number=72, 2mm isotropic voxels and multiband acceleration factor = 8. T1-weighted images were obtained using an MPRAGE sequence using a TR=2530ms, TE =1.15ms, field-of-view  $256 \times 256$ mm, flip angle=7 degrees, 1mm isotropic voxels (Chai et al., 2012; Uğurbil et

al., 2013). The first 13 volumes (corresponding to 10 seconds) were removed to eliminate saturation effects and achieve steady state magnetization. The HCP minimal preprocessing pipeline was used to preprocess functional data (Glasser et al., 2013). This included artefact removal, motion correction, registration to structural T1-weighted scan, and non-linear registration into MNI space (Smith et al., 2013).

Connectivity analysis was performed using the CONN functional Connectivity Toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012). Functional images were highpass filtered below 0.009Hz to remove low frequency drifts due to scanner noise. Physiological and motion related noise were dealt with by using the anatomical *CompCor* technique, which removes the first 5 principal components of the signal from white matter and cerebrospinal fluid masks, as well as the motion parameters and their first-order temporal derivatives (Behzadi et al., 2007). In detail, to minimise partial voluming, the white matter and CSF masks were eroded by one voxel, which resulted in substantially smaller masks than in the original segmentations (Chai et al., 2012). I then used the eroded white matter and CSF masks as noise regions of interest. Specifically I used the *CompCor* method to remove the top 5 principal components from white matter and CSF masks alongside the motion parameters and their temporal derivatives. A temporal band-pass filter of 0.008–0.09Hz was also applied on the time series to restrict the analysis to low-frequency fluctuations (Fox et al., 2005).

Following preprocessing, I computed temporal correlations between each region in the parcellation scheme used, resulting in a  $234 \times 234$  or  $129 \times 129$  functional connectivity matrices for each subject.

### **3.2.4 Assignment of ROIs to large-scale networks**

In order to assist with interpretation of the results I further assigned each ROI to LSNs of interest. To do so I used the 7 cortical network masks from Yeo et al. (Yeo et al., 2011) and I calculated each ROI's overlap with each LSN. The maximum overlap served as the criterion for assigning each ROI to an LSN. The number of connections within and between the different networks was defined using the real structural matrices from the HARDI data.

### 3.2.5 Nash equilibrium Network model

In this section I describe the Nash model used to obtain optimal connections in the structural connectome. The Nash equilibrium Network Game model (NNG) is a network construction game derived from the Nash equilibrium of a non-cooperative game, where each player or node follows a strategy independently of others in order to maximize a certain payoff (Nisan, 2007). In the context of this work, each node's strategy was to maximize efficient communication with the rest of the brain by making efficient "navigable" paths while maintaining a minimum number of edges. I will provide definitions for each node's strategy using graph-theoretic terms following closely the formulation used by Gulyás and colleagues (Gulyás et al., 2015). A triplet of coordinates  $x, y, z$  represented the position in the Euclidean space of the nodes of interest  $V$  (e.g. the ROI centre of mass of the 234 or 129 parcellation resolutions). Distance between two nodes  $p_1$  and  $p_2$  with coordinates  $x_{p_1}, y_{p_1}, z_{p_1}$  and  $x_{p_2}, y_{p_2}, z_{p_2}$  respectively was defined as the Euclidean distance and was calculated as

$$dist(p_1, p_2) = \sqrt{(x_{p_1} - x_{p_2})^2 + (y_{p_1} - y_{p_2})^2 + (z_{p_1} - z_{p_2})^2}$$

For each pair of nodes  $u \in V$  and  $v \in V$ , let  $S_v^u = \{w \mid dist(v, w) < dist(u, w)\}$ . This represents the set of nodes that one can efficiently navigate to from  $u$ , using  $v$  as an intermediate hop.

The problem with finding a strategy for a node  $u \in V$  is formulated as follows. First, node  $u \in V$  is associated with a collection of sets  $S_v^u$ , each one corresponding to the rest of the nodes  $v \in V \setminus \{u\}$ . In turn, the optimal strategy of node  $u$  consists of constructing edges to those nodes  $v'$  such that their  $S_{v'}^u$  sets belong to the minimum cover set of the sets  $S_v^u$ . For a collection of sets, the minimum cover set problem refers to selecting the minimum number of sets such that their union includes all the elements appearing in the collection of sets. Formally, for each  $v \in V \setminus \{u\}$  the problem is to find the binary values of the decision variables  $d_v$  indicating whether the  $S_v^u$  is going to be selected or not as part of the minimum cover set. To assist with the problem formulation I assign a variable  $a_{v',v} = 1$  based on whether a node  $v'$  belongs

to a set  $S_v^u$ . The minimum cover set problem for node  $u$  can then be written as the following optimization problem.

$$\begin{aligned} & \min \sum d_v \\ & \text{subject to } \sum a_{v'v} d_v \geq 1 \quad v, v' \in V \setminus \{u\} \\ & d_v = \{0,1\} \end{aligned}$$

Instead of solving the integer form of the minimum set cover problem, I considered its linear relaxation by allowing the decision variables to be real non-negative numbers rather than integers, i.e.  $d_v \geq 0$ . This relaxation transforms a hard linear integer programming problem into a related problem that can be solved in polynomial time (from a computational complexity perspective) while, at the same time, the solution of the latter is at least as good as the original problem (Schrijver, 1987).

The described procedure results in a number of connections for node  $u$  that represent its optimal strategy. If this is repeated independently for all nodes  $u \in V$ , it can be proved that the resulting network is the Nash equilibrium of all the nodes' strategies and is characterized by maximum navigability with minimum number of edges (Gulyás et al., 2015). To obtain NNG models/networks for each subject's T1 parcellation I used the methodology that follows. For each subject's ROIs' coordinates, I calculated the minimum cover set for each ROI  $u$  by formulating and solving the corresponding problem using the *glpk* library (<http://www.gnu.org/software/glpk>). Edges between the ROI  $u$  and the rest of the ROIs  $v \in V \setminus \{u\}$  were constructed when the software's output for  $d_v$  was 1. This was repeated for all ROIs of each subject resulting in a subject-specific NNG network. Structural optimality results were obtained by comparing the NNG edges with the edges in the structural matrix from the HARDI data and described in the *Structural Connectivity* section. Specifically, I defined a true positive/optimal connection (T) between two ROIs as one that existed in both the real, structural network and the NNG network. Non-optimal connections were those that existed in the real, structural network but were not present in the NNG network. False positives were defined as those connections that existed in the

NNG network but not in the structural connectivity network (F). Optimality was defined as the ratio  $|T|/|M|$  where  $|T|$  was the total number of optimal connections and  $|M|$  was the total number of connections in the NNG network. False positives were defined as  $|F| = |M| - |T|$ .

### 3.2.6 Predicting functional from structural connectivity

After obtaining the optimal connections for each individual using the NNG network, I utilized these to predict functional connectivity data and its complexity. In the next section I describe the algorithm for conducting this analysis.

#### *Formulation of the problem*

For each pair of functional and structural connectivity matrices, I used a transformation of structural to functional connectivity matrices in order to predict functional connectivity (Becker et al., 2018). Suppose that for an individual subject  $j$  I obtained structural and functional connectivity matrices of dimensions  $n \times n$ :  $\mathbf{S}_j$  and  $\mathbf{F}_j$  respectively. The first step consists of writing the predicted functional connectivity matrix as

$$\tilde{\mathbf{F}}_j = \mathbf{R} \left( \sum_{r=0}^k a_r \mathbf{S}_j^r \right) \mathbf{R}^T$$

The term  $\sum_{r=0}^k (a_r \mathbf{S}_j^r)$  is a weighted sum of powers of  $\mathbf{S}_j$  up to order  $k$  (polynomial transformation of order  $k$ ). Spectral graph theory states that powers of  $k$  of the structural connectivity matrix are related to the walks of length  $k$  in the underlying structural graph (Chung, 1997). These are traversals from one vertex of the graph to another (with a potential repetition of the same vertices) using the edges of the graph. Therefore, higher-order transformations show how information can be communicated upon long traversals in the graph, something that is potentially important for predicting functional connectivity. After the polynomial transformation, the rotation matrix  $\mathbf{R}$  is used to transform the eigenvectors of the

matrix  $\mathbf{S}_j$  in order to align to those of the real functional connectivity matrix  $\mathbf{F}_j$ . In the second step, I solved the optimization problem that finds the best approximation  $\tilde{\mathbf{F}}_j$ , i.e. I solved the following problem

$$\min_{\{a_r\}_{r=0}^k, \mathbf{R}} \left\| \tilde{\mathbf{F}}_j - \mathbf{F}_j \right\| = \left\| \mathbf{R} \left( \sum_{r=0}^k a_r \mathbf{S}_j^r \right) \mathbf{R}^T - \mathbf{F}_j \right\|, \mathbf{R}^T \mathbf{R} = \mathbf{R} \mathbf{R}^T = \mathbf{I}_n, \det \mathbf{R} = 1$$

where  $\| \cdot \|$  stands for the Frobenius norm and  $\mathbf{I}_n$  is the all-ones diagonal matrix of dimension  $n$ . The two constrains guarantee that the matrix  $\mathbf{R}$  is a rotation matrix.

Prior to solving this minimization problem, I will provide notation. I write  $\mathbf{v}_{i=1}^n, \lambda_{i=1}^n$  and  $\mathbf{u}_{i=1}^n, \phi_{i=1}^n$  for the eigenvectors and eigenvalues of  $\mathbf{S}_j$  and  $\mathbf{F}_j$  respectively. First I considered the vectors  $\boldsymbol{\phi} = [\phi_1, \dots, \phi_n]^T$ ,  $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T$  and the matrices  $\mathbf{V} = [v_1 | \dots | v_n]$  and  $\mathbf{U} = [u_1 | \dots | u_n]$  for  $\mathbf{S}$  and  $\mathbf{F}$  respectively. I also considered the Vandermonde matrix

$$\mathbf{L} = \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^k \\ 1 & \lambda_2 & \dots & \lambda_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \dots & \lambda_n^k \end{bmatrix}$$

where  $k$  is the order of the polynomial transformation. The minimization problem can be decomposed into two problems that I solved in their generic form.

The first problem is that of finding the optimal coefficients to fit the polynomial transformation of order  $k$  to the matrix  $\mathbf{S}$ . Generally, for a matrix  $\mathbf{g}$  and coefficient matrix  $\mathbf{A}$ , this problem can be written as

$$\min_x \|\mathbf{A}\mathbf{x} - \mathbf{g}\|$$

where  $A_{ij} = t_i^{j-1}$  can be written as the Vandermonde matrix

$$\begin{bmatrix} 1 & t_1 & \dots & t_1^k \\ 1 & t_2 & \dots & t_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \dots & t_n^k \end{bmatrix}$$

The solution to this problem is based on optimization theory (Boyd and Vandenberghe, 2007) and can be obtained as

$$\mathbf{x}^* = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{g}$$

The second problem rotates the eigenvectors of the weighted sum of powers of  $\mathbf{S}_j$  to fit the eigenvectors of  $\mathbf{F}_j$ . In general, for matrices  $\mathbf{A}$  and  $\mathbf{B}$  the problem pertains to finding a rotation matrix  $\mathbf{\Omega}$  as follows.

$$\min_{\mathbf{\Omega}} \|\mathbf{\Omega} \mathbf{A} \mathbf{\Omega}^T - \mathbf{B}\|$$

$$\mathbf{\Omega}^T \mathbf{\Omega} = \mathbf{\Omega} \mathbf{\Omega}^T = \mathbf{I}_n, \det \mathbf{\Omega} = 1$$

For real symmetric matrices, this problem can be solved using the eigenvalue decomposition of the matrix  $\mathbf{A} = \mathbf{U}_A \mathbf{\Sigma}_A \mathbf{U}_A^{-1}$  and the eigenvalue decomposition of  $\mathbf{B} = \mathbf{U}_B \mathbf{\Sigma}_B \mathbf{U}_B^{-1}$  with the solution being the matrix  $\mathbf{\Omega}^* = \mathbf{U}_A \mathbf{U}_B^T$  (Schönemann, 1968).

By utilizing these formulas for the prediction problem, I obtained the pair of solutions

$$(a_0^*, \dots, a_k^*)^T = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \phi, \quad \mathbf{R}^* = \mathbf{U} \mathbf{V}^T$$

and thus  $\hat{\mathbf{F}}_j$  was written as

$$\hat{\mathbf{F}}_j = \mathbf{R}^* \left( \sum_{r=0}^k a_r^* \mathbf{S}_j^r \right) (\mathbf{R}^*)^T$$

The algorithm for predicting  $\hat{\mathbf{F}}_j$  was used in three scenarios that differed according to their structural connectivity matrix inputs 1) using the whole-brain structural connectivity matrix  $\mathbf{S}_j$  2) using the (sub) network of  $\mathbf{S}_j$  consisting only of optimal connections 3) using the (sub) network of  $\mathbf{S}_j$  consisting only of non-optimal connections. To assess the importance of structural connections in each scenario, I used two different goodness-of-fit measures between the predicted  $\hat{\mathbf{F}}_j$  and the real  $\mathbf{F}_j$  functional connectivity matrices as described in the following section.

***Evaluation of prediction***

The predictive ability of optimal connections was compared against other structural connections in order to provide evidence for their important role in driving functional connectivity. To so, I quantified the goodness-of-fit between the predicted  $\widehat{\mathbf{F}}_j$  and the real  $\mathbf{F}_j$  functional connectivity matrices by employing two methods: First I used correlation of the upper triangular entries obtained from the real matrix with those entries obtained from the predicted matrix. Second I used the proposed homology-based evaluation. This evaluation was based on comparing the number of connected components of the predicted and real functional connectivity matrices at different edge density levels  $\lambda$  (Betti numbers  $\beta_0(\lambda)$  and  $\widehat{\beta_0(\lambda)}$  respectively) (Liang and Wang, 2017). Here edge density refers to the percentage of correlation entries in the functional connectivity matrix. Specifically, zero density refers to an empty matrix whereas density of 1 indicates the presence of all values in the matrix.  $SSE_\beta$  was used to evaluate the goodness-of-fit for the prediction and can be formulated as

$$SSE_\beta = \frac{1}{n^2} \int_0^1 (\beta_0(\lambda) - \widehat{\beta_0(\lambda)})^2 d\lambda$$

where the integral spans all edge densities from 0 (no correlation entries existing) to 1 (all correlation entries present) and  $n$  is the dimension of the matrix. Intuitively, the smaller the score the better the fit of the predicted matrix to the real matrix as the number of different connected components is smaller.

**3.2.7 Complexity of functional connectivity networks**

In turn, the predictive ability of optimal connections was evaluated in the context of how well they could predict the complexity of functional connectivity data. First, to obtain complexity of functional connectivity, functional connectivity matrices were thresholded using the BCT toolbox at a range of edge densities  $\tau$ . This allowed for individuals' networks to have the same number of edges. I chose a range of densities from  $\tau = 7\%$  to  $\tau = 19\%$  in steps of 2%. The lower threshold was chosen as a more conservative threshold compared to the 4% threshold chosen by Power et al. (Power et

al., 2011) with the purpose of keeping the networks from being severely fragmented. The upper threshold was chosen so as to prevent the networks from being random by introducing unnecessary edges (Bassett and Bullmore, 2006). As introduced in the Methods chapter (2), complexity of functional connectivity was calculated by using the entropy of the degree distribution  $p_i$  across all regions of the parcellation (Zhao et al., 2011; Zamora-López et al., 2016). A standard entropy estimator calculates entropy by utilising the probability distribution of the degree sample  $p_i$ . This was done by calculating the probability of each degree value appearing once or twice and so on (naïve estimation). To formalize these, for a degree sample of length  $k$  obtained from the functional connectivity network, I calculated its fingerprint ( $F(1), F(2), \dots$ ) where  $F(1)$  was the number of nodes appearing once,  $F(2)$  was the number of nodes appearing twice and so on (Valiant and Valiant, 2013). Then entropy was defined as  $H = - \sum_{i \in p_i} p_i \log p_i = - \sum_i \frac{i}{k} F(i) \log i/k$ . Results presented here were for an average over the aforementioned range of thresholds  $\tau$  in order to show that the results were not anchored to some specific threshold.

The complexity of predicted functional connectivity networks obtained were compared to the ones obtained from real functional connectivity networks for cases of optimal, non-optimal, and whole-brain structural connections.

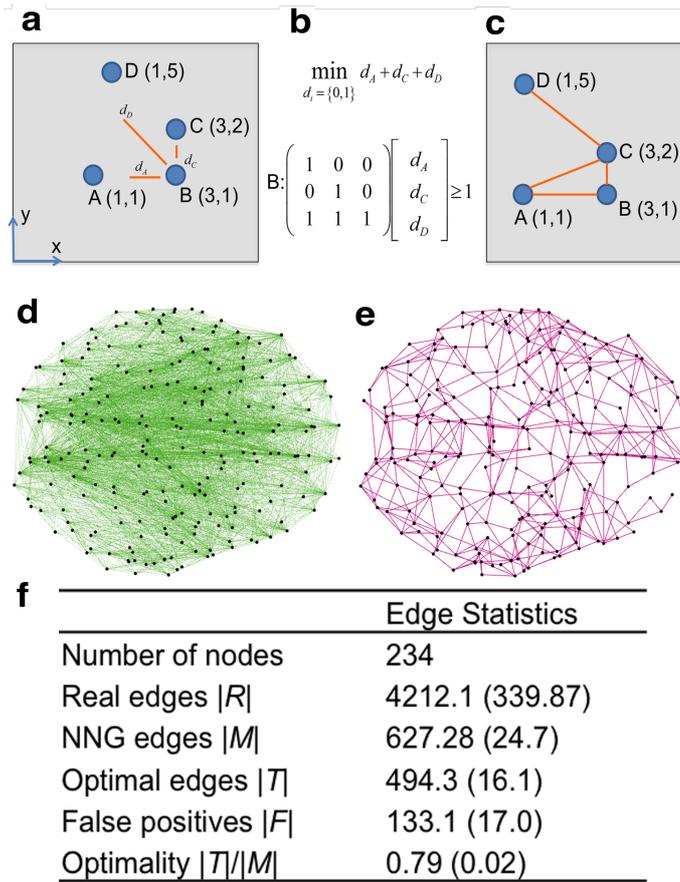
### 3.2.8 Statistical analysis

Data was tested for normality using the Kolmogorov-Smirnov test. Statistical analysis was conducted using a one-way, repeated-measures analysis of variance (ANOVA) with the use of optimal, non-optimal, and whole-brain connections as a factor. The  $p$  values reported are from (two-tailed) *post hoc comparisons* after Bonferroni correction. I also report effect sizes for ANOVA using the bayes factor and the eta squared quantities.

### 3.3 Results

#### 3.3.1 Identifying optimal connections in structural connectivity networks

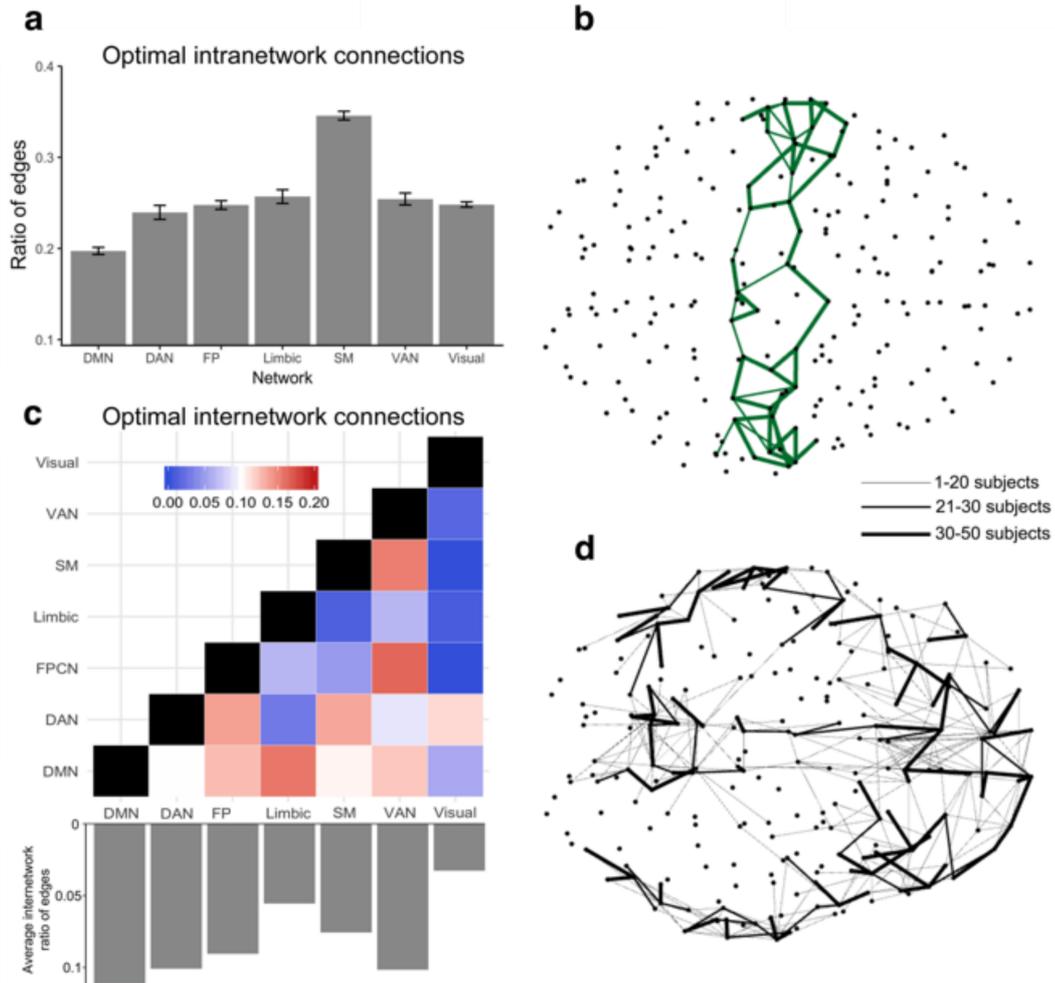
The Nash equilibrium Network Game model (NNG) takes as input the brain's topographical organization, as defined by the spatial location of brain regions or nodes (derived from T1 scans), and creates a synthetic network by maximizing the objectives of each node. Specifically, the resulting NNG network achieves Nash equilibrium by simultaneously maximizing efficient information routing between all nodes in the network (navigability) while minimizing the number of connections (Gulyás et al., 2015). Given these data, the NNG model solves an optimization problem (for each ROI independently) and produces a set of connections between the ROIs (Fig. 3.1a, 3.1b, and 3.1c). Initially, I determined whether the connections predicted by the NNG model were present in the real structural connectivity networks obtained from the HARDI data. I first report results for the 234-ROI parcellation. The mean optimality score across subjects (mean=0.79, stdev=0.02) suggested that all brain networks displayed a consistent pattern of optimal connectivity (Fig. 3.1f). The number of false positives was low across subjects (mean=133.1, stdev=17.0) compared to the number of connections that the NNG produced (mean=627.28, stdev=24.7) indicating statistical robustness of our model (Gulyás et al., 2015). To show that the results were consistent across different parcellation resolutions I applied the same model to the 129-ROI parcellation scheme. Structural networks using this parcellation also contained a consistent number of optimal connections (mean=0.7161, stdev=0.0430) while false positives remained low across subjects (mean=93.8, stdev=15.8) (Appendix Table 3.1).



**Figure 3.1 - Identifying optimal connections in structural connectivity networks using the NNG model.** (a) Example of the game-theoretic model (NNG) applied to a four-node graph on the  $xy$  plane taken from Gulyás et al. (2015). Each node is associated with three binary decision variables indicating whether it will connect to other nodes. To obtain the decision variable values, I compute three sets  $S_A^B$ ,  $S_C^B$ , and  $S_D^B$  corresponding to three decision variables  $d_A$ ,  $d_C$ , and  $d_D$ . For example,  $S_A^B$  is defined as the set of node's  $B$  neighbors  $w$  that are closer to the destination  $A$  i.e. nodes such that  $dist(w, A) < dist(w, B)$  or else  $S_A^B = \{w | dist(w, A) < dist(w, B)\}$  ( $dist$  is the Euclidean distance) (b) Decisions are based on the minimum cover set of these three sets; here I find that  $d_A = d_C = 1$  and  $d_D = 0$ . Therefore, BC and BA connections will be created. (c) By repeating calculations for the remaining nodes, it can be shown that the resulting network has maximum navigability (as a result of the “greedy” selection of neighbours close to target) and a minimum number of connections (as a result of the minimum cover set). (d) Example of whole-brain structural network. Black dots represent the location of each ROI/node and green lines show existing connections (edges) between ROIs. (e) Optimal connections in magenta colour obtained using the locations of ROIs shown in  $d$  and the NNG model. (f) Optimality statistics for the structural networks at the 234-ROI resolution. Results are presented in the form of mean (stdev) over  $n=50$  subjects. For similar results at the 129-ROI resolution-see Appendix Table 3.1.

### 3.3.2 Optimal structural connections in large-scale networks

I next investigated the extent to which optimal connections were distributed across established cortical networks. To assess this, I assigned each ROI from the Lausanne 234-ROI parcellation to a previously defined cortical network (Yeo et al., 2011). I then computed the number of optimal connections between ROIs within networks (intranetwork optimal connections) and between ROIs of different networks (internetwork optimal connections). To account for differential network sizes I divided the number of optimal connections by the total number of intra- or internetwork connections. This revealed that the SM network contained the highest proportion of optimal intranetwork connections (Fig. 3.2a, 3.2b), suggesting that connections within the SM form a comparatively self-integrated structure. This is in agreement with earlier studies showing that the somatomotor network has a low participation coefficient and high local efficiency (Power et al., 2011). The DMN showed the lowest optimal intranetwork connectivity, possibly reflecting the metabolically expensive fronto-parietal connections linking the MPFC and the PCC (Greicius et al., 2009). In contrast, I found that the DMN had substantially higher internetwork optimal connectivity than any other network (Fig. 3.2c, 3.2d), in support of the idea that the DMN may act in part, as a global workspace, integrating information from other LSNs (Dehaene and Naccache, 2001; Vatansever et al., 2015b).



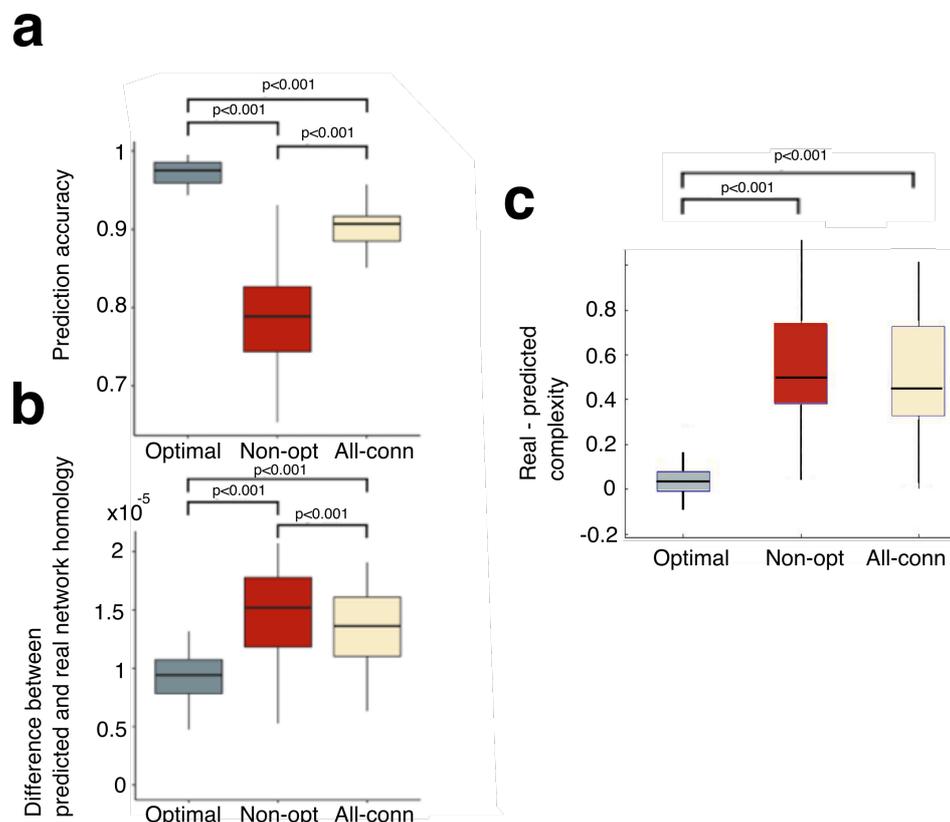
**Figure 3.2 - Intra- and internetwork optimal connections in structural brain networks** (a) Optimal intranetwork connections in each of 7 cortical brain networks. Bars show mean and standard error of the mean. (b) Intranetwork optimal connections in the somatomotor network, the network that showed the greatest number of intranetwork connections. (c) Optimal internetwork connections between each cortical network averaged over subjects. Bars below show the mean of optimal connections from each network to the other 6 networks. (d) Internetwork optimal connections from the DMN, the network that showed the greatest number of internetwork connections. For (b) and (d) line emphasis corresponds to connections that exist in more than 30 (out of the 50 subject sample) subjects. Network definitions are from Yeo et al. (2011). All numbers reflect ratios of the number of optimal connections over the total number of intra- or internetwork connections respectively. *Abbreviations:* DMN = Default Mode Network; DAN = Dorsal Attention Network; FP = Fronto-Parietal Network; SM = Somato-Motor Network; VAN = Ventral Attention Network; VIS = Visual Network.

### 3.3.3 Predicting functional connectivity and its complexity using optimal structural connections

Next, I probed the capacity of the optimal structural network to predict whole-brain functional connectivity and its complexity. I employed a recently developed algorithm that predicts resting-state functional connectivity using the Eigen-structure of a structural connectivity matrix (Becker et al., 2018). By using only the optimal connections in the structural connection matrix as input, this method allowed me to assess how well can optimal connections predict functional connectivity and its complexity. Then I compared the predicted with the actual functional connectivity matrices obtained from the CONN toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012). Comparison between the predicted and real matrices was conducted using Pearson correlation between their entries. For the 234-ROI parcellation and a sufficiently large order polynomial transformation ( $k=5$ ), I found that predictive ability was superior when using only the optimal connections (correlation coefficient between predicted and real functional matrices mean  $r=0.965$ ,  $stdev=0.032$ ) compared to using all the structural connections (correlation coefficient between predicted and real functional network mean  $r=0.902$ ,  $stdev=0.023$ ) (Fig. 3.3a). I also found that the predictive ability dropped considerably when using only the non-optimal structural connections (correlation coefficient between predicted and real functional network mean  $r=0.784$ ,  $stdev=0.061$ ). A one-way repeated-measures ANOVA confirmed that each prediction was significantly different from the others ( $F(2,98)=206.32$ ,  $p<0.0001$  Bonferroni corrected multiple comparisons test, bayes factor=91.7949, eta squared=0.8085). For prediction results across different polynomial transformation orders, see Appendix Fig. 3.1. These results were further verified using a newly introduced method from homology theory (Liang and Wang, 2017) that calculates prediction accuracy by quantifying the difference between the number of connected components in the predicted and real functional matrices, such that a higher score implies poorer prediction. Here I observed that the differences in connected components were smaller using only optimal connections thus showing their high predictive accuracy ( $F(2,98)=28.07$ ,  $p<0.0001$  Bonferroni corrected multiple comparisons test, bayes factor=19.0641, eta squared=0.3642) (Fig. 3.3b). Thus, I argue

that this set of optimal connections is the most important driver of whole-brain functional connectivity.

In turn, I used these optimal connections to show that complexity of functional connectivity is related to the underlying optimal structure. As a measure of complexity of functional connectivity networks, I used the entropy of their degree distribution (Zhao et al., 2011; Zamora-López et al., 2016). In that regard, I calculated the entropy of the degree distribution of functional connectivity networks from the predicted functional connectivity matrices obtained above (corresponding to optimal, the non-optimal and all the structural connections). Then I compared the complexity of these networks with those obtained from real functional connectivity data and I found an effect on retrieving complexity depending on which set of structural connections was used ( $F(2,98)=73.96$ ,  $p<0.0001$ , bayes factor=45.5461, eta squared=0.6015). I found that there was a high correspondence between predicted and real complexity only in the case of optimal connections implying that the complexity of data could be retrieved successfully only when optimal connections were considered (Fig. 3.3c) ( $p<0.001$  for both cases Bonferroni corrected). Similar results were obtained for the 129-ROI parcellation-Appendix Fig. 3.2)



**Figure 3.3 - Predicting functional connectivity and its complexity from structural connectivity.** (a) The prediction algorithm uses the Eigen-structure of a polynomial transformation of order  $k$  (here only for  $k = 5$ ) of the structural connectivity matrix to predict the functional connectivity matrix. Each box plot shows the variation of the prediction accuracy over  $n=50$  subjects in terms of cross-matrix correlation from each set of structural connections. Higher scores show higher predictability. See Appendix Fig. 3.1 for different polynomial orders  $k$ . (b) Here I used methods from homology theory to correlate predicted and real functional matrices. Lower scores reflect a smaller difference between the real and predicted functional networks. Both prediction scores show that using only optimal (Opt) connections was significantly more predictive of whole-brain functional connectivity than using non-optimal (Non-opt) or all structural connections (All-conn). (c) Differences between predicted complexity and complexity obtained from real functional connectivity matrices using optimal, non-optimal and whole-brain connections. Complexity was obtained as the entropy of degree distribution of functional connectivity networks. See Appendix Fig. 3.2 for the 129-ROI parcellation results. The closer the value is to 0, the better the prediction of complexity is. Boxplots' thick lines show the median value while whiskers reflect the maximum and minimum values of the data.

### 3.4 Discussion

The goal of this chapter was to find out how brain complexity emerges from structural connectivity. First, by using diffusion imaging data from 50 healthy volunteers, I identified a consistent proportion of optimal connections in structural connectivity networks. Earlier work has shown that structural networks have small-world architecture i.e. their organization is characterized by a high efficiency in communication usually facilitated by the presence of short paths (Bullmore and Sporns, 2012; Betzel and Bassett, 2016). How does the optimal communication model concur with the small-world hypothesis? Optimal connections were derived on the premise of each node's strategy to construct paths that can promote efficient communication. Thus structural networks efficiently communicate information in a pattern that takes into consideration how each node differentially routes information to the rest of the network, expanding the small-world framework where one would expect information to be routed using only shortest paths (Avena-Koenigsberger et al., 2017).

In addition, application of the NNG model to structural brain networks by using the brain's unique layout in three-dimensional space illustrates the significance of the underlying cortical organization to structural connectivity as derived by the diffusion

data. The notion that an optimal communication model of structural connectivity is intertwined with the geometrical placement of the regions in the cortex has been the focus of research from early descriptions of cortical organization (Cherniak, 1994) to recent studies utilizing advances in mapping axonal interconnections (Ercsey-Ravasz et al., 2013). Studies suggest that the positioning of cortical regions is optimal in the sense that it minimizes axonal wiring cost and enhances communication under the various physical constraints relative to brain size (Koulakov and Chklovskii, 2001; Klyachko and Stephens 2003; Cherniak et al., 2004) in contrast to any sub-optimal layout which would compromise efficient connectivity (Chklovskii et al., 2002). An additional explanation comes from certain morphogenetic accounts suggesting that, although increasing tensions (associated with expanding brain volume and size) affect local connections, the aggregate/whole-brain wiring length is minimal (Van Essen, 1997). Taken together, the existence of an optimal communication model might be associated with how the structural connectome has evolved in order to attain minimal wiring length and promote efficient communication at all levels.

I then asked how optimal structural connections are differentially distributed across LSNs. First, a high number of optimal connections were found within the SM network. This is consistent with previous studies suggesting that the SM network has high local communication efficiency possibly reflective of continuous and comparatively rigid processing demands (Power et al., 2011). In contrast, the DMN had a low number of intra-network optimal connections potentially attributed to the fact that the NNG is not best suited for the capture long-range connections linking the MPFC with the PCC (Greicius et al., 2009; Betzel et al., 2017).

Despite having a low number of optimal intra-network connections, the DMN had the highest number of optimal inter-network connections suggesting it is efficiently communicating information to other LSNs. Along these lines, studies on structural connectivity have singled out the DMN as a network extensively connected to regions of other LSNs (Parvizi et al., 2006; Hagmann et al., 2008; Greicius et al., 2009; Ongur and Price, 2010). These results show that this connectivity is supported by principles of efficient communication and minimum wiring cost reflecting flexible information exchange of the DMN with other LSNs (Vatansever et al., 2015a). This finding is also

in line with research demonstrating a central role for the DMN in cognitive processing (Vatansever et al., 2017). In a recent study by Margulies and colleagues (Margulies et al., 2016) the authors showed that DMN regions occupy ideal positions along a principal gradient, a property related to the topographic organization of large-scale connectivity, thus promoting efficient information processing. The NNG model provides a complementary approach in understanding DMN's unique multi-functional organization. I showed that the topographic locations of the DMN regions are ideal in the sense that they allow the emergence of a high number of optimal connections towards the rest of the brain. When conceptualizing the proportion of internetwork optimal connectivity in a continuous spectrum, our results place the DMN at the one end, distant from the contributions of other networks, thus providing evidence for its functional heterogeneity and flexibility.

I then used a structural-to-functional prediction algorithm in order to identify the extent to which optimal connections can predict functional connectivity and its complexity. The prediction algorithm utilized path information embedded into the structural connectivity matrix by expressing functional connectivity as a weighted combination of powers of the structural connectivity matrix, each one associated with walks of different lengths (Becker et al., 2018). I found that optimal connections were more predictive of whole-brain functional connectivity than the entire structural connectome. This finding suggests that paths consisting of optimal edges serve as better predictive features for explaining functional connectivity. Interregional communication using optimal paths may provide a framework for explaining flexible brain function in order to satisfy different cognitive demands. Beyond this, I used these optimal connections to retrieve the complexity of functional connectivity networks. Complexity addresses the segregation and integration in functional connectivity networks (Zamora-López et al., 2016). The existence of an optimal communication model supports the idea that information can be transmitted efficiently from the local/regional to the global/whole-brain level. It is therefore possible that an optimal communication model can provide a suitable scaffold for how segregated information can become integrated across the whole brain thus supporting the emergence of complex functional connectivity. It is worth noting that optimal connections were differentially distributed across networks. This potentially

reflects that the integration “part” of complexity could be anchored more to networks with high number of internetwork optimal connections such as the DMN, while the segregation “part” could be anchored more to sensory systems with high number of intranetwork optimal connections such as the SM.

Because complexity derived upon the equilibrium of optimal connections resembled the one obtained from the real data, I argue that it is critical, meaning that only the balance/equilibrium of an efficient structural connectivity can explain this complexity and other, non-equilibrium configurations would fail to do so. Corroborating the premises of criticality in the brain as a system, I showed that complexity of functional connectivity emerges from the collective outcome (in a game-theoretic fashion) of each node’s tendency to efficiently communicate with the rest of the brain.

Critical complexity of functional connectivity is an important feature of the healthy brain as it has been associated with the capacity of information that can be transmitted across the brain (Shew et al., 2011) or the tendency of the brain to adapt to different cognitive demands (Beggs and Plenz, 2003; Beggs, 2008). More importantly, as I shall show in the next chapter, changes in complexity, as aberrations from this criticality, might be a marker of loss of consciousness.

### **3.5 Conclusion**

In sum, I provided evidence that the complexity of functional connectivity can be derived upon a balance/equilibrium of optimal structural connections. Can an unconscious brain signify departure from this balance? To answer this question, in the next chapter I will use alterations of complexity to show how they can discriminate between different states of consciousness.

## **Chapter 4: Reduction of complexity during loss of consciousness**

Criticality refers to the state of a system within a range of “critical” transitions where even small perturbations can change its behaviour (Tagliazucchi and Chialvo, 2012). At that range, a system displays specific properties that can only be understood as the collective behaviour of its constituent parts and not as the outcome of a single part alone. I focused on one such property, the complexity of functional connectivity networks as a means of investigating the diversity/non-uniformity of connectivity across the brain. In chapter 3, I showed that the complexity of functional connectivity in the healthy brain is critical as it emerges from the collective configuration (Nash equilibrium) of the optimal structural connections in the human brain.

Disruption of consciousness due to pharmacological causes can drastically perturb the (critical) complexity of the healthy brain. Alluding to the second major objective of this work, I claim that deviations from complexity of the awake, conscious brain could be used to signify loss of consciousness. What is the motivation behind this hypothesis? Complexity of functional connectivity has been linked to the brain’s ability to critically balance segregation and integration of information (Zamora-López et al., 2016). It is this balance that generates integrated yet diversified information, a feature that has been deemed tantamount to the richness of conscious experience (Tononi, 2004). Aberrations of complexity in the healthy brain would imply changes in this balance where functional connectivity networks would be biased towards a more integrated or segregated configuration (Tononi et al., 1998). In that regard, theoretical models suggest that loss of consciousness is associated with *reduction* in complexity i.e. a shift towards a more segregated configuration where information cannot be broadcasted globally in the network (Carhart-Harris et al., 2014). For example, GNW states that processes at individual/segregated units cannot attain a conscious level unless they are integrated in the global workspace (Dehaene and Changeux, 2011). Tononi and Edelman also suggest that neuronal information becomes conscious only if it is sufficiently integrated across the whole brain (Tononi and Edelman, 1998). Thus the goal of this chapter is to show that reduction in

complexity of functional connectivity can capture consciousness-dependent changes by using data from anaesthetic-induced unconsciousness and patients with disorders of consciousness.

#### **4.1 Introduction**

Functional connectivity is believed to measure neural synchronization by quantifying relationships between BOLD signals from different brain regions (Dosenbach et al., 2007). Under this context, a standard measure of a region's level of synchronization to the rest of the brain is its number of edges (degree) with evidence showing that the degrees of functional connectivity conform to specific distributions following a non-trivial heavy-tailed pattern (Vértes et al. 2012). The complexity of this distribution emphasizes the co-existence of highly connected regions (usually forming densely connected modules) and sparsely connected regions that together support efficient communication in the brain (Zamora-López et al., 2016).

As I showed in chapter 3, the complexity of the degree distribution is a critical trait of the healthy awake brain. If a critical complexity characterizes the awake brain's functional connectivity, then fundamental questions pertain to how this complexity changes during altered states of consciousness. One account suggests that network complexity reflects the capacity to process local and global information in the conscious brain (Tononi and Edelman, 1998). The consciousness spectrum is therefore assumed to be echoed by a complexity spectrum where pharmacologically-induced or pathologically driven decreases in consciousness correspond to decreased complexity and a shift towards a more segregated brain (entropic brain hypothesis) (Alkire et al., 2008; Carhart-Harris et al., 2014; Tagliazucchi et al., 2014). Studies focusing on the precise quantification of alterations in complexity are methodologically inconclusive. Limited work has shown reorganization of degrees in comatose patients (Achard et al., 2012) however with no precise characterization of where these changes take place (for example hubs or non-hubs).

Here, I used the entropy of the degree distribution similar to chapter 3 as a measure of functional connectivity complexity. However, in an effort to see whether the effects on complexity were more evident in different parts of the brain, I analysed different parts of the degree distribution entropy separately. This allowed me to precisely characterize reductions in complexity by showing changes in the entropy of regions with important functional role in network integration.

I applied this methodology to data from propofol-induced sedation, anaesthesia and patients with disorders of consciousness. My goal was twofold: first to characterize changes in complexity during loss of consciousness. Second, my goal was to delineate a commensurate spectrum of complexity; from high levels of complexity corresponding to the awake state to reduced complexity in sedation and anaesthesia and even more reduced complexity in patients with disorders of consciousness.

The rationale, results and conclusions for these parts are presented in the following sections.

A part of this chapter has been published as an article titled “Brain network disintegration during sedation is mediated by the complexity of sparsely connected regions” in *Neuroimage*, 2018.

## **4.2 Changes in complexity under propofol-induced sedation and anaesthesia**

### **4.2.1 Overview**

A working hypothesis in this thesis is that consciousness is lost is because the complexity of functional connectivity is impaired and the brain shifts towards a less integrated/more segregated configuration (Alkire et al., 2008). As I argued in previous chapters, one approach for testing this hypothesis is to use the entropy of functional connectivity as a measure of complexity. In addition to this approach, a more refined quantification of entropy has been adopted here by focusing on different parts of the degree distribution that will provide information regarding a) where the complexity changes take place and b) how these affect network integration. Specifically, this approach allowed me to consider the following questions. First, I asked whether

reduced consciousness is characterized by decreased global/whole-brain entropy. Second, I asked whether a decrease in global/whole-brain entropy would be disproportionately evident in a specific part of the degree distribution, motivated by evidence that loss of consciousness is characterized by connectivity changes that do not span the entire distribution (Alkire et al., 2008; Schiff, 2008; Guldenmund et al., 2013). A possible explanation could be the differential effect of the anaesthetic agents on the pattern of specific connections. Sparsely connected regions are primary candidates as they are potentially more susceptible to anaesthetic agents compared to highly connected ones-hubs (Alkire et al., 2008). In addition, limited work on functional connectivity data (Gallos et al., 2012) has led to the hypothesis that perturbation of sparse connectivity might produce disconnected network components at a regional level causing loss of network integration; however, evidence for this during states of decreased consciousness has not yet been shown.

Moreover, I asked whether specific network-level connectivity changes disproportionately contribute to sedation-induced changes in entropy. Previous work has discussed the importance of two specific LSNs in this context. On the one hand, the FP network has been implicated in the loss of information integration during decreased consciousness (Guldenmund et al., 2016) while others have shown that the DMN is re-organized during sedation and anaesthesia (Boveroux et al., 2010; Stamatakis et al., 2010).

Lastly, obtaining a connectivity-based marker predictive of an individual's behaviour during decreased consciousness remains a fundamental challenge, mainly due to the variety of connectivity changes in different regions. Here I investigated the relationship between complexity and behaviour by asking whether the level of degree entropy during the awake state could predict an increase in the reaction times observed during a task executed under sedation.

I addressed these questions using fMRI imaging data obtained during different levels of propofol sedation. Administration of propofol was used to achieve light and moderate sedation followed by a recovery state. Resting-state functional MRI scans during these four (including baseline-awake) sedation levels allowed us to construct

networks representative of each individual's functional connectivity and to quantify entropy at different levels of sedation. I complemented these results by showing similar trends in a deep anaesthesia-related dataset. In this experiment individuals transitioned from an awake state to light anaesthesia and, in turn, a fully anaesthetised state.

#### **4.2.2 Materials and Methods**

##### ***Sedation data***

Part of the data used in this experiment was previously used in a work by Stamatakis and colleagues (Stamatakis et al., 2010).

##### ***Participants***

25 healthy, right-handed volunteers with no history of neurological disorders (16 male, 9 female, range=19-52 years) were included in the propofol experiment. All procedures were according to the local ethics permission from the Cambridgeshire 2 Regional ethics committee.

##### ***Experimental setup and procedure***

In order to test the effect of sedation on functional connectivity networks and their complexity, fMRI data was obtained during different propofol-induced sedation levels. The experimental design (leaving out task data that was collected during the same experiment) consisted of four scanning runs of resting-state fMRI corresponding to four levels of sedation (awake, light sedation, moderate sedation, recovery). The order of the administration of light and moderate sedation was randomized. Data was obtained during a resting-state period where propofol was administered intravenously. A computer controlled intravenous infusion was used aiming to achieve three target plasma levels - no drug (awake), 0.6  $\mu\text{g/ml}$  (light sedation), and 1.2  $\mu\text{g/ml}$  (moderate sedation-corresponding to Ramsay score 3)

(Ramsay et al., 1974; Marsh et al., 1991; Absalom and Menon, 2009). fMRI data for each resting-state condition (awake, light sedation, moderate sedation, recovery) was collected for 5 minutes. During scanning we instructed volunteers to close their eyes and think about nothing in particular throughout the acquisition of the resting-state BOLD data. Volunteers were informed of the risks of propofol administration and were also informed about more minor effects of propofol such as pain on injection, sedation, and amnesia. In addition, standard information about intravenous cannulation, blood sampling, and MRI scanning was provided. During data collection there were always two trained anaesthesiologists present, and observed the volunteer from the MRI control room and on a video link that showed the volunteer in the scanner. Task data during a semantic judgement task was also collected for all the sedation levels after obtaining the resting-state data. The experimental design is described in detail in Adapa et al. (2014). In short, the task experiment consisted of four scanning runs (corresponding to the sedation levels described previously) where each scanning run lasted 5.5 minutes and comprised alternating 30 seconds blocks of words and acoustically matched non-speech (buzz/noise) stimuli. Stimuli were presented with a stimulus onset asynchrony (SOA) of three seconds in silent intervals between scans. Blocks of 8 stimuli were presented with an additional 6 seconds of silence between blocks to allow estimation of baseline activity. Participants were asked to respond with a button press to indicate whether presented words referred to living or non-living items and whether non-speech stimuli were buzz-type or noise-type items. Response times acquired from this experiment were analysed (as described in the *Behavioural data* section) and correlated with results from the complexity analysis as described in the *Behavioural data correlations* section.

### ***Behavioural data***

There were two types of stimuli used: a) words that were pseudo-randomly drawn from a set of 280 items (140 living items, e.g. tiger, birch, and 140 nonliving items, e.g., table, stone) in subsets of 40 items (20 living, 20 nonliving) and were matched for relevant psycholinguistic variables (word frequency, length, imageability, acoustic amplitude, and familiarity). Participants heard 4 of the 7 groups of spoken words in

the four scanning runs with assignment of items to sedation levels counter-balanced over participants. b) Buzz/noise stimuli that were generated by extracting the amplitude envelope of a spoken word and using that envelope to modulate either a broad band noise (noise) or a harmonic complex with a 150Hz fundamental frequency (buzz). These sounds were filtered to match the average spectral profile of the source word. This process was implemented using a custom script implemented with Praat software (<http://www.praat.org>). Blocks of 8 stimuli were presented with an additional 6s of silence between blocks to allow estimation of resting activity. Participants were asked to respond with a button press to indicate whether presented words referred to living or nonliving items and whether nonspeech stimuli were buzz-type or noise-type items. Behavioural responses were categorized into correct responses, incorrect responses, and time-outs (no response to stimuli within 3 s of presentation). Here I used reaction times for responses corresponding to correctly classified stimuli (measured from word onset).

### ***Image acquisition and preprocessing***

Functional data was acquired using a Siemens (Erlangen, Germany) 3T scanner at the Wolfson Imaging Center, Cambridge. Functional data was obtained using an EPI sequence and each functional BOLD volume consisted of 32 interleaved, descending, oblique axial slices, 3mm thick with interslice gap of 0.75mm, in-plane resolution of 3mm, field of view =  $192 \times 192$ mm, TR = 2000ms, TE = 30ms, and flip angle=78 degrees. T1-weighted structural images were acquired at 1mm isotropic resolution in the sagittal plane, using an MPRAGE sequence with 1mm isotropic resolution, with TR= 2250ms, TE=2.99ms, TI=900ms, field-of-view= $256 \times 256$ mm, and flip angle=9 degrees.

Images were preprocessed and modelled using SPM version 12.0 and MATLAB Version 12a platforms. The first six volumes were discarded to allow for MR signal equilibration. All imaging data was preprocessed following a standard pipeline of slice-timing and motion correction and normalization to the MNI space. Functional images were smoothed with an 8 mm FWHM Gaussian kernel.

### ***Regions of interest***

The main objective of this work was to quantify the complexity of functional connectivity networks and how it changes at different levels of sedation. I adopted a whole-brain approach in which correlation matrices were built based on regions of interest (ROIs) spanning the entire brain. Details are presented below. I used a set of predefined regions of interest (ROIs) as follows: a) Cortical ROIs from the FSL Harvard-Oxford maximum likelihood cortical atlas divided bilaterally into areas of the left/right hemisphere (91 ROIs). b) Subcortical ROIs from the FSL Harvard-Oxford maximum likelihood subcortical atlas (15 ROIs) (Makris et al., 2005). c) Cerebellar parcellation from the Automated Anatomical Labelling (AAL) atlas (26 ROIs) (Tzourio-Mazoyer et al., 2002). The combined atlas of 132 ROIs (eAAL atlas) was provided with the CONN toolbox described below. The main motivation behind using this atlas (compared to the Lausanne parcellation used in chapter 3) was to compensate for the absence of finely defined subcortical and cerebellar regions that potentially play an important role in loss of consciousness (Alkire et al., 2008).

### ***Assignment of ROIs to large-scale networks***

To help with interpretation of the results, I further classified regions based on their LSN membership. To do so, I defined cortical LSNs using the 7 masks from Yeo et al. (2011) that are available online

*([http://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation\\_Yeo2011](http://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation_Yeo2011))*

In order to assign each ROI to one of the networks, I used an in-house script to calculate the overlap of each ROI with each network. Each ROI was assigned with an overlap number for each network by counting the number of voxels that were both in the ROI and the network. The maximum overlap (across networks) determined network membership. Subcortical and cerebellar ROIs were grouped together manually to become “subcortical” and “cerebellar” networks respectively. For confirmation of our results on residual entropy, I also used the 10 Smith et al. (2009) canonical network masks that are available online

(<http://www.fmrib.ox.ac.uk/datasets/brainmap+rsns>)

I followed an identical methodology for assigning each ROI to the 10 networks.

### ***Functional connectivity networks***

I used the CONN functional connectivity toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2011) to obtain the time series from the ROIs described above. A strict temporal preprocessing pipeline of nuisance regression was applied. Linear regression confounds included *CompCor* components attributable to white matter and CSF signals (Behzadi et al., 2007), subject-specific six realignment parameters alongside their first-order derivatives and the effects of the drug level (Fair et al., 2007a). In addition, a high pass filter of 0.009Hz was applied to remove low frequency fluctuations due to scanner noise.

Z-transformed linear correlation coefficients were calculated between the 132 time series corresponding to each ROI and used to construct correlation matrices for the awake state, mild sedation, moderate sedation, and recovery using inbuilt MATLAB functions. For each individual, I constructed four correlation matrices corresponding to four different conditions. Each matrix was then proportionally thresholded using the BCT toolbox (<http://sites.google.com/site/bctnet/>) at a range of densities  $\tau$  in order to construct networks having the same number of edges across individuals and conditions. Here I chose a range of densities from  $\tau = 7\%$  to  $19\%$  in steps of  $2\%$ . The lower bound was chosen empirically as a more conservative value compared to the  $4\%$  chosen by Power and colleagues to construct network with same parcellation (Power et al., 2011) with the purpose of keeping networks from being severely fragmented. The upper bound was chosen so as to prevent the networks from being random by introducing unnecessary edges (Bassett and Bullmore, 2006). For each  $\tau$  in that range, degrees for all the regions were calculated using the BCT toolbox. The degree for each region was defined as its number of edges regardless of the strength of the correlation (binarized degree sample). Results presented in this work were averaged over this range of thresholds in order to show robustness of the results with respect to different thresholds.

***Complexity of functional connectivity networks: global, rare and frequent entropy***

Complexity was calculated for the whole-brain functional connectivity networks using the entropy of the degree distribution  $p_i$  defined as

$$H = -\sum_i p_i \log p_i.$$

Because I wanted to investigate how different parts of entropy change, instead of calculating entropy in the same way for all regions as in chapter 3, I calculated entropy in two parts; one for regions that their degree value appeared frequently in the sample (frequent degrees) and one for regions that their degree value appeared infrequently (rare degrees). This was based on the assumption that propofol would have a differential rather than a uniform effect on entropy. From a methodological perspective, different approaches were used for calculating rare and frequent entropies. The rationale behind this was that standard (naïve) entropy estimators cause bias in the estimation of entropy when sample elements do not appear often (as in rare degrees) (Valiant and Valiant, 2013). I present the details in the following parts.

To use notation, for a degree sample of size  $k$  obtained from a brain network, I calculated its fingerprint ( $F(1), F(2), \dots$ ) where  $F(1)$  was the number of nodes appearing once,  $F(2)$  was the number of nodes appearing twice and so on. Given a frequency level  $B \in \{1, \text{length of fingerprint}\}$  indicating an index in the fingerprint, the part of the fingerprint from  $B$  to the end of the fingerprint ( $F(B + 1), \dots, \text{end of fingerprint}$ ) corresponded to degrees appearing frequently in the sample whereas the part of the fingerprint ( $1, \dots, F(B)$ ) corresponded to the rare degrees.

The histogram of the degree distribution  $p$ , denoted by  $h_p(x)$ , is equal to the number of degrees that occur with probability  $x$ . Using this notation, entropy is defined as

$$H = -\sum_{x:h_p(x) \neq 0} h_p(x) x \log x$$

The histogram of the distribution  $h_p(x)$  and probabilities  $x$  were calculated separately for frequent and rare degrees. Specifically, for the frequent degrees a histogram

$h_{frequent}(x) = F(i)$  with probability  $x = \frac{i}{k}$ , was assigned for each  $i \in (F(B + 1), \dots, \text{end of fingerprint})$ . Thus the entropy of frequent degrees was calculated as  $H_{frequent} = - \sum_{i>B} F(i) \frac{i}{k} \log i/k$ . This is also called the naïve estimation of entropy and it is identical to the calculation of entropy used in chapter 3. For the rare degrees, a histogram of probability distribution was derived by solving an optimization problem. First, an initial, non-informative mesh of probabilities was considered  $x_1, \dots, x_l$ , with

$$x_1 = \frac{1}{k^2}, x_i = 1.1x_{i-1}, 0 < x_i \leq 1$$

In turn, the goal was to find a histogram  $h_{rare} = h_1(x_1), \dots, h_l(x_l)$  corresponding to these probabilities such that it would minimize the quantity

$$\sum_{i \leq B} \frac{1}{\sqrt{F(i) + 1}} \left| F(i) - \sum_{j=1}^l h_j(x_j) \text{poi}(x_j k, i) \right|$$

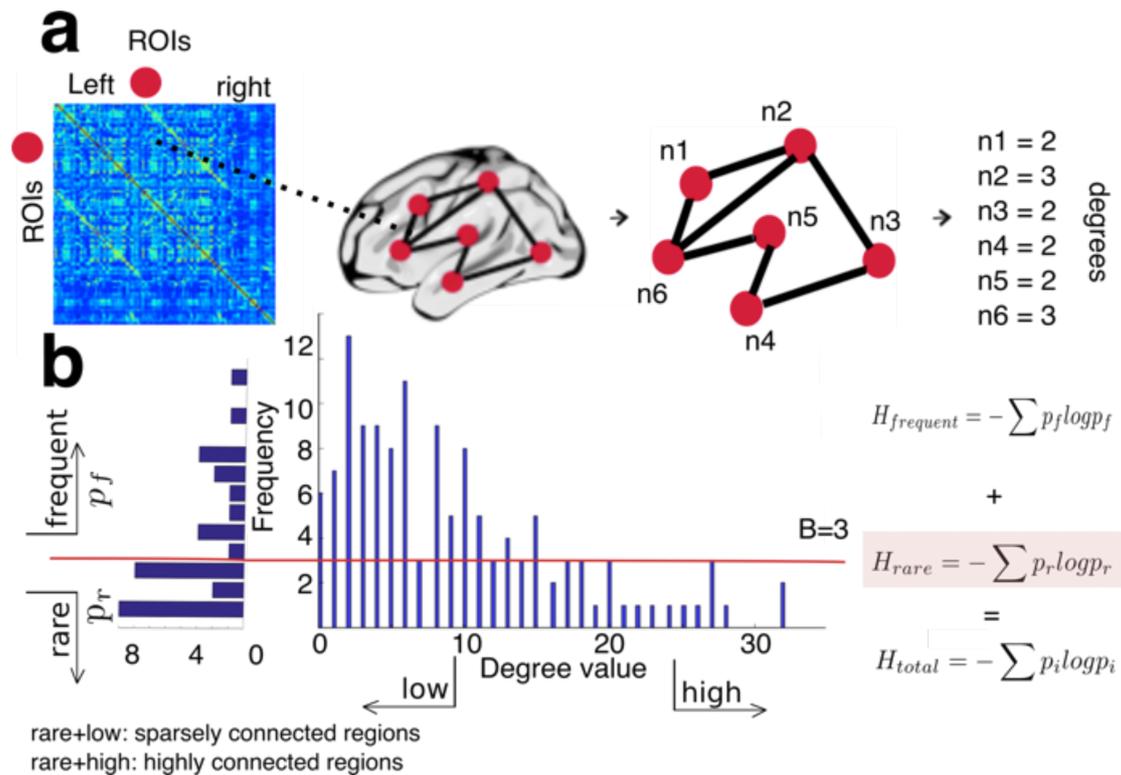
under the constraints that  $\sum_{j=1}^l x_j h_j(x_j) = \sum_{i \leq B} \frac{F(i)}{k}$  and  $h_j(x_j) \geq 0 \forall j$ .

The notation  $\text{poi}(x_j k, i)$  for a degree associated with probability  $x_j$ , represents the probability of that degree to appear  $i$  times in the sample of size  $k$  (this probability follows a Poisson distribution). Practically speaking, the goal was to find a probability histogram  $h$  that minimizes the distance between the expected fingerprint  $\sum_{j=1}^l h_j(x_j) \text{poi}(x_j k, i)$  and the observed fingerprint  $F(i)$  of the rare degrees. In this way the bias in estimating entropy for the rare degrees would be minimized. The distance between the expected and observed fingerprints was penalized by the inverse of standard deviation of  $F(i)$  calculated as  $1/\sqrt{F(i) + 1}$  as this would minimize their statistical discrepancy (Valiant and Valiant, 2013). The constraint  $\sum_{j=1}^l x_j h_j(x_j) = \sum_{i \leq B} \frac{F(i)}{k}$  guaranteed that the total probability would sum to 1. The constraint  $h_j(x_j) \geq 0, \forall j$  ensured that all probability values were nonnegative.

The solution of this problem resulted in a histogram  $h_{rare}$  and the rare entropy was then calculated as  $H_{rare} = -\sum_{j=1}^l h_j(x_j)x_j \log(x_j)$ . The details of how to solve the optimization problem are described in Valiant and Valiant (2013).

There is an advantage of using an optimization algorithm to estimate the entropy in the rare degree “regime”. When sample elements are seen only a handful of times (such as the rare degrees), the naïve estimation underestimates their underlying distribution and entropy i.e. it causes bias in the entropy estimation (Valiant and Valiant, 2013). To this end, optimisation techniques such as the one mentioned before, have proven to be more successful in retrieving the entropy compared to the naïve estimator (Valiant and Valiant, 2013). In the case when sample elements are seen often (such as the frequent degrees), the previous optimization technique and the naïve estimator have virtually indistinguishable results in recovering the underlying entropy (Valiant and Valiant, 2013). Thus I chose to use the naïve estimator for calculating the entropy of the frequent degrees.

The code used to obtain the rare degree distribution can be found here (<http://theory.stanford.edu/~valiant/code.html>). The definition of rare degrees was associated with degrees appearing in the fingerprint with frequencies from 1% (i.e.,  $B = 1$ ) to 4% (i.e.  $B = 5$ ) of the degree sample size (132). The upper bound was chosen empirically so as to include approximately half of the obtained average fingerprint length over all individuals and conditions. Global/whole-brain entropy was calculated as the sum of the frequent ( $H_{frequent}$ ) and rare ( $H_{rare}$ ) entropies. All entropy results shown were calculated as an average over different  $B$  levels (and over different network densities  $\tau$  as I mentioned previously).



**Figure 4.1 - Degree entropy of functional connectivity networks as a measure of complexity: a more refined approach.** Part (a) shows how to obtain a network representation from functional connectivity data. Each entry in the matrix is a Pearson correlation between the BOLD signals of two brain regions. A functional connectivity network can be obtained by considering brain regions as nodes and their signal correlations as edges. The degree of each region is defined as its number of edges towards the rest of the network. Part (b) shows the associated degree histogram for a typical functional connectivity network. The number of nodes having degrees appearing once, twice, ... can be quantified using the fingerprint of the sample; that is the histogram of the values in the y-axis of the histogram (“histogram of the histogram”). The red line splits the degrees to i) those that occur below a specific frequency  $B$  (here  $B=3$ ) in the sample (associated with rare degrees entropy,  $H_{rare}$ ) ii) the rest of the degrees (associated with frequent degrees entropy,  $H_{frequent}$ ). The former includes degrees of regions that are rare but have a high degree value (highly connected regions) and degrees of regions that are rare but have a low degree value (sparsely connected regions). The global entropy of the sample is calculated by adding the rare and frequent entropies.

### **Normalizing rare and frequent entropy**

When calculating the effect of sedation on rare and frequent entropies, I normalized these by the lengths of the rare and frequent degree distributions respectively. This was done in order to account for sample size bias in the sense that when summing the probabilities for calculating entropy one wants this to be independent of the number

of probabilities being summed. As the number  $B$  was fixed for calculating rare entropy for each individual at each sedation level the length of the rare distribution (calculated from the part of the fingerprint  $\leq B$ ) was fixed. Thus normalization for rare entropy was not necessary. However, this was not the case for the frequent distribution. For example, it can be the case that the part corresponding to the frequent degrees ( $F(B + 1), \dots, \text{end of fingerprint}$ ) has variable length due to the variable length of the fingerprint thus leading to false estimates. For consistency, I normalized both rare and frequent degree entropies by the lengths of their corresponding distributions  $H_{rare} = -\sum_{i \in p_{rare}} (p_i \log p_i) / \text{length}(p_{rare})$  and  $H_{frequent} = -\sum_{i \in p_{frequent}} (p_i \log p_i) / \text{length}(p_{frequent})$ .

***Characterizing regions with rare-low (sparsely connected) and rare-high (highly connected) degrees***

Splitting the degree sample into two types of degrees (regions with rare, and frequent degrees) allowed me to look into regional effects in entropy. As it will be evident in the results, regions with rare degrees had decreased entropy during sedation showing a similar trend as in whole-brain entropy. Thus I shifted my focus to looking into more specific effects within these regions alluding to their different functional roles and, specifically, their role in network integration. Regions with rare degrees could include regions that have high degrees (rare-high or highly connected regions) or regions with low degrees (rare-low or sparsely connected regions). A hypothesis would be that the effect of propofol would be more evident in the sparsely connected regions (rare-low) and that these would cause the brain networks to become more segregated.

Towards this direction, sparsely connected regions were obtained as follows. I first identified the rare degrees in the original degree sample alongside the top 20 high degrees in the sample. I then took the intersection of these two and I subtracted it from the original set of rare degrees leaving me only with the rare-low degrees. Regions with rare-high degrees were computed similarly by removing the bottom 20 low degrees from the original set of rare degrees. After identifying sparsely connected regions, I used these in several ways:

1) First I identified their impact on network integration by looking at how they affect the largest connected component. To do so, I calculated their impact on the large connected component after virtually lesioning them (Gallos et al., 2012; Del Ferraro et al., 2018). The largest connected component in a network is the largest number of nodes that are connected with each other. In other words, if one starts from any node within the connected component, one can always reach (using a path consisting of edges) every other node within the connected component. To establish the size of the largest connected component after lesioning certain nodes, I used a quantity defined

$$G_{or} - G_{les}$$

where  $G_{or}$  is the size of the largest connected component in the original network,  $G_{les}$  is the size of the largest connected component in the network after lesioning the node of interest. This quantity is usually normalized by  $N_{or}$ , the number of nodes in the original network to address network size bias.

Thus I had a normalized positive measure that quantified the size of the connected component in the lesioned network with respect to the size of the connected component in the original network. The higher the measure the more difference in the connected component and thus the more impactful the lesion is in network integration. To compute the size of the connected component I used a publicly available code found here

*(<http://uk.mathworks.com/matlabcentral/fileexchange/30926-largest-component>)*

2) I used their entropy (rare-low entropy) instead of all rare entropy to see if it can better discriminate between sedation states, the assumption behind this being that sparsely connected regions would be disproportionately affected by propofol. Rare-low entropy was further normalized by the length of their corresponding distributions to account for sample size bias for the reasons mentioned previously.

3) As it will become apparent in the results section, sparsely connected regions were shown to be the strongest contributors to effects observed from earlier analyses. To gather more information on this, I also examined the effect of sedation on the location of sparsely connected regions. For consistency I will not discuss individual regions in

this set of results but each of the 9 LSNs they belong to. I classified the sparsely connected regions as belonging to different canonical networks obtained from Yeo and colleagues (Yeo et al., 2011).

4) I calculated their spatial extent of connectivity with rest of the brain.

### ***Calculating the residual entropy of each large-scale network***

In turn, I shifted my focus from regions to LSNs and their role in whole-brain entropy. To do so, I conducted a network lesion analysis by calculating residual entropy after removing the nodes corresponding to each LSN based on either the Yeo et al. (2011) or Smith et al. (2009) network definitions.

### ***Behavioural data correlations***

As it will become evident in the results section, rare-low entropy and residual entropy could successfully classify between different sedation levels. In turn, I examined the relationship between reaction times and these entropies in order to see if there is a link between entropy and level of responsiveness that we use as a rough proxy of the state of consciousness in the sedated participants. Behavioural data for 20 of the 25 participants were obtained after the resting-state fMRI sessions corresponding to the different states of sedation. The experimental paradigm used to obtain behavioural data is described in detail in a study by Adapa and colleagues (Adapa et al., 2014) and in the *Behavioural data* section presented earlier. It is worth noting that I did not correlate brain metrics with propofol concentration in plasma, as I was looking for effects relevant to the maintenance of consciousness (measured by the level of behavioural responsiveness in this case) and not to the action of the drug itself (e.g. generalized synaptic depressant action). This comes in line with previous literature suggesting that brain connectivity changes that correlate with propofol in the blood might not be identical to those that perturbate the level of behavioural responsiveness (Barttfeld et al., 2015b).

***Statistical analysis***

Statistical comparisons for global, rare, and frequent entropy were performed using ANOVA for repeated measures with the one factor being the four conditions (awake, mild sedation, moderate sedation, recovery). Data was checked for normality using the Kolmogorov-Smirnov test. The p values reported are from (two-tailed) *post hoc comparisons* after Bonferroni correction. I also report effect sizes for ANOVA using the bayes factor and the eta squared quantities.

To compare the effect of sparsely connected regions on the length of the largest connected component during different sedation levels, I used one-way ANOVA for repeated measures with Bonferroni correction for the *post hoc* tests.

To compare the discriminatory ability of sparsely connected regions compared to previous analysis that used only rare entropy, I used an equivalent of Cohen's d-score in the multivariate setting called the Mahalanobis distance (Del Giudice et al., 2012). P values to establish whether there was a difference in the ANOVA effects were obtained by constructing appropriate confidence intervals with bootstrapping using the code provided here

(<http://sites.google.com/site/mvlombardo/matlab-tutorials/computeeffectsizes>).

Statistical comparisons regarding the location and connectivity of sparsely connected regions were carried out after grouping these in LSNs. I performed pairwise t-tests between the location and connectivity values between the two extreme conditions: awake and moderate sedation. The p values obtained were further Bonferroni corrected for 9 networks meaning that a p value was multiplied by 9 times before considered statistically significant.

To examine the effect of sedation on the residual entropy of 9 networks, one-way ANOVA was used. P values coming from *post hoc* comparisons after Bonferroni correction. The p values were further corrected for multiple comparisons taking into account the 9 networks under consideration. When the Smith et al. (2009) network definitions were considered, corrections were carried out for 10 networks.

Finally, I examined the relationship between rare-low entropy and behavioural data. The correlation and p values reported were obtained using Pearson's linear correlations between reactions time and (rare-low or residual) entropy.

### ***Anaesthesia data***

To verify the results on the sedation dataset a similar analysis was conducted on dataset where healthy individuals transitioned to a fully anaesthetized state.

### ***Participants***

19 healthy, right-handed volunteers with no history of neurological disorders (18-40 years; 13 males) healthy volunteers participated in this study. Four volunteers (1 male) were excluded from data analyses of the study due to technical difficulties in data acquisition resulting in a total of 15 volunteers. A resting-state scan (8 minutes) was acquired, during which volunteers were asked to relax with their eyes closed and not fall asleep. Ethical approval was obtained from the Health Sciences Research Ethics Board and Psychology Research Ethics Board of Western University, Canada. All experiments were performed in accordance with the relevant guidelines and regulations set out by the research ethics boards.

### ***Experimental setup and procedure***

Resting-state fMRI data was acquired while participants were awake (non-sedated), lightly anaesthetized (Ramsay score 3) and deeply anesthetized (Ramsay score 5) with propofol induction (Ramsay et al., 1974). 8-minute resting-state data was obtained for awake, light anaesthesia, deep anaesthesia and recovery conditions. Intravenous propofol was administered with a Baxter AS 50 (Singapore). An effect-site/plasma steering algorithm was used in combination with the computer-controlled infusion pump to achieve step-wise increments in the sedative effect of propofol. The pharmacokinetic model provided target-controlled infusion by adjusting infusion

rates of propofol over time to achieve target blood concentrations (Marsh et al., 1991). Propofol infusion commenced with a target effect-site concentration of 0.6 µg/ml (state of light anaesthesia). If the Ramsay level was lower than 5, the concentration was slowly increased by increments of 0.3 µg/ml with repeated assessments of responsiveness between increments to obtain a Ramsay score of 5 (state of deep anaesthesia). Throughout the scanning, the participant's behavioural profile was monitored inside the scanner room by the anaesthesia nurse and one of the anaesthesiologists and outside from the scanner control room, with an infrared camera that displayed the participant's face. No movement, fluctuations of sedation, or any other state change, was observed during the deep sedation scanning for any of the participants included in the study.

### ***MRI acquisition and preprocessing***

Scanning was performed using a 3 Tesla Siemens Tim Trio system with a 32-channel head coil, at the Robarts Research Institute in London, Ontario, Canada. Functional images were acquired using an EPI sequence (33 slices, voxel size:  $3 \times 3 \times 3$ mm, inter-slice gap of 25%, TR=2000ms, TE=30ms, matrix size=  $64 \times 64$ mm flip angle=75 degrees). An anatomical volume was obtained using a T1-weighted MPAGE sequence (32 channel coil, voxel size:  $1 \times 1 \times 1$ mm, TA=5 min, TE=4.25ms, matrix size=  $240 \times 256$ mm, flip angle=9 degrees). Data was preprocessed using the same SPM pipeline as in the sedation experiment.

### ***Regions of interest***

I used the eAAL atlas with the same ROIs as described in the sedation paradigm section.

***Assignment of ROIs to large-scale networks***

I used the same methodology as described in the previous section to associate each ROI with an LSN.

***Functional connectivity networks***

I used the CONN functional connectivity toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2011) to obtain the time series from the previous ROIs. A strict temporal preprocessing pipeline of nuisance regression was applied. Linear regression confounds included *CompCor* components attributable to white matter and CSF signals (Behzadi et al., 2007), subject-specific six realignment parameters alongside their first-order temporal derivatives and the effects of the drug (Fair et al., 2007a). In addition, a highpass filter of 0.009Hz was applied to remove scanner noise. Functional connectivity matrices were obtained by taking the pairwise Pearson correlations between each pair of ROIs. Functional connectivity matrices were thresholded using the same range of thresholds  $\tau$  as in the sedation experiment.

***Complexity of functional connectivity networks***

Global entropy, rare entropy, frequent entropy and residual entropy were calculated as previously.

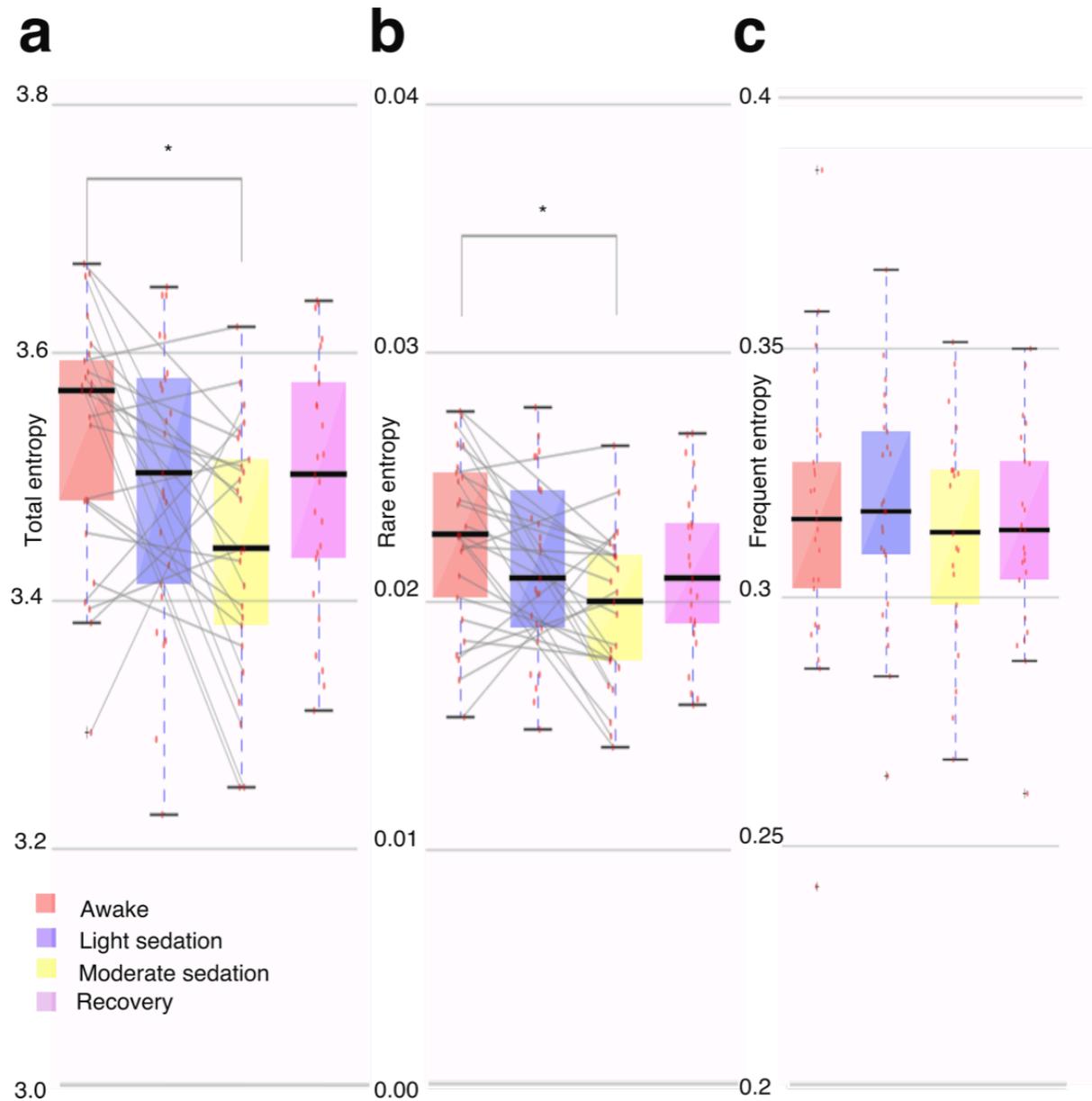
***Statistical analysis***

Data was checked for normality using the Kolmogorov-Smirnov test. Statistical analysis included ANOVA with the anaesthesia level as a single factor. The p values reported were from *post hoc* comparisons after Bonferroni correction. I also report effect sizes for ANOVA using the bayes factor and the eta squared quantities.

## 4.2.2 Results

### *Global and rare entropy in the whole-brain*

I first asked whether the global entropy of the degree distribution of functional connectivity networks changes with increasing sedation. A repeated-measures ANOVA showed a significant effect of sedation on global entropy ( $F(3,72)=3.39$ ,  $p=0.0226$ , bayes factor=0.7874, eta squared=0.1585) (Fig. 4.2a) and *post hoc* t-tests showed that entropy was decreased during moderate sedation compared to the awake state ( $p=0.0137$ ). When I considered different parts of the degree distribution, I found a significant effect of sedation on rare entropy ( $F(3,72)=3.46$ ,  $p=0.0208$ , bayes factor=0.8635, eta squared=0.1612) (Fig. 4.2b) with *post hoc* t-tests revealing a significant decrease during moderate sedation compared to the awake state ( $p=0.0120$ ). I conducted a similar analysis for the frequent entropy where I observed no alterations across sedation conditions ( $F(3,72)=0.29$ ,  $p=0.8352$ ) (Fig. 4.2c). In other words, the pattern I observed with rare degrees alone mirrored the pattern of global entropy. *Post hoc* tests revealed that moderate sedation had significantly less rare entropy than the awake state ( $p=0.0066$ ). Collectively these results suggested that the effect of sedation was more evident in regions with rare degrees potentially reflecting the selective action of the anaesthetic.



**Figure 4.2 - Complexity is decreased during moderate sedation.** Entropy of the degree distribution was calculated as a proxy of complexity for functional connectivity networks. Part (a) boxplot shows changes in the global\whole-brain entropy from the awake state to moderate sedation. Part (b) boxplot shows changes in rare entropy alone (part of entropy corresponding to regions with rare degrees) with increased sedation. Part (c) boxplot shows that the frequent entropy (part of entropy corresponding to regions with frequent degrees) remains unchanged across sedation levels. Single asterisk (\*) indicates *post hoc* significance  $p < 0.05$  Bonferroni corrected. Grey lines between boxplots indicate changes in individual volunteers from awake to moderate sedation. Boxplots' thick lines show median values and whiskers represent 1.5 times the inter-fourth range. Since rare and frequent entropy are parts of the total entropy, boxplots of (b) and (c) are on a different scale. Entropy of rare and frequent degrees was normalized by the length of their respective distribution to account for size bias; thus plots b+c do not sum to a.

***Rare entropy in the whole-brain: the role of sparsely connected regions***

Regions with rare degrees may be highly connected (high degrees) or sparsely connected (low degrees) with each having different functional roles in network communication. Information produced locally usually within groups of highly connected regions (modules or components) propagates to the distant parts of the brain through sparsely connected regions; thus the importance of individual regions for whole-brain network integration varies (Bullmore and Sporns, 2012). Motivated by the hypothesis that sparsely connected regions might have a profound impact on network integration, I identified these within each individual's rare degree sample and I assessed their role in network integration. Mathematically one can quantify this by measuring the size of the largest connected component in the network after these regions have been virtually lesioned/removed (Gallos et al., 2012; Del Ferraro et al., 2018-Fig. 4.3a). The larger the connected component in the lesioned network the less the impact in network integration of the nodes that were lesioned. Using this methodology I found two interesting results regarding the role of the sparsely connected regions in network integration: 1) During the awake state the sparsely connected regions were associated with a smaller connected component compared to an equally sized sample of randomly chosen (rare and frequent) nodes ( $F(3,72)=4.58$ ,  $p=0.0054$ , bayes factor=2.0778, eta squared=0.2028 and Fig. 4.3b) 2) This effect was abolished with increasing sedation implying a loss of their role in network integration (*post hoc* test awake vs. moderate sedation  $p=0.0046$ ). Combining the last result with the fact that I observed a reduced complexity in the rare degrees implies that sparsely connected regions reorganized in a pattern that prevented network integration when participants were moderately sedated.

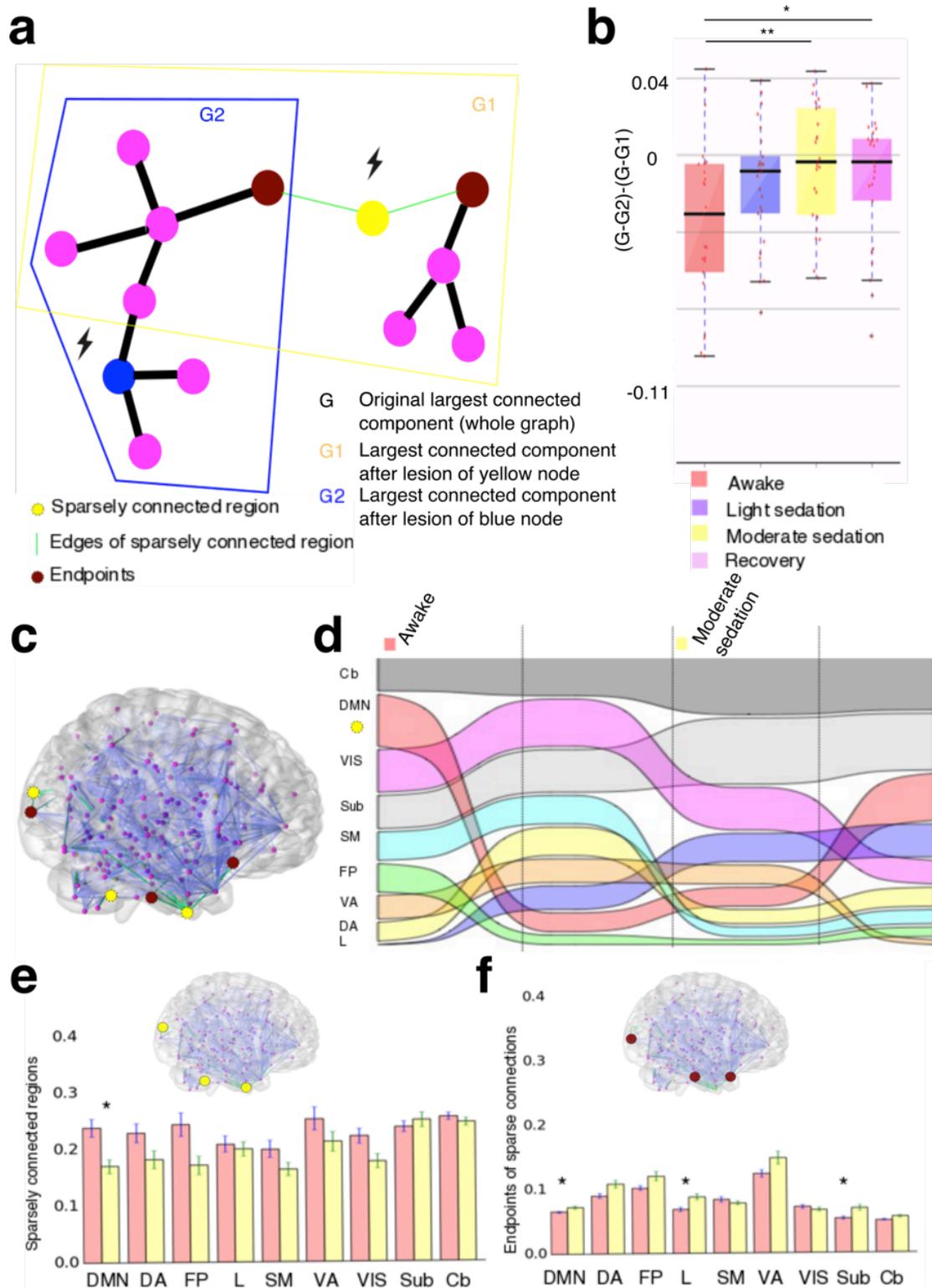
***Using only sparsely connected regions to discriminate between sedation levels***

The previous results collectively suggested that the complexity of sparsely connected regions might be more predictive of anaesthetic depth than other regions. Here I sought to confirm this by first removing the rare-high degrees from each participant's sample and, using the Mahalanobis distance (Del Giudice et al., 2012) to compare

effect sizes. I found a stronger effect size compared to the effect size when using the unaltered rare entropy (Mahalanobis  $p=2.9e-04$ ). When rare-low degrees were removed, the effect size did not change (Mahalanobis  $p=0.944$ ) thus signifying a greater contribution from sparsely connected regions to the ANOVA effect size. Collectively, these findings demonstrated that sparsely connected regions could be used to discriminate between different consciousness states with alterations in their connectivity linked with their role in network integration.

### ***The location and connectivity of sparsely connected regions***

To facilitate the characterization and localisation of sparsely connected regions, I investigated canonical network membership utilising a set of predefined functional networks widely used in resting-state fMRI literature (Yeo et al., 2011) (alongside 2 manually defined networks spanning subcortical and cerebellar regions). The first observation was that network membership of sparsely connected regions changed as volunteers transitioned across different sedation levels (Fig. 4.3c, 4.3d, and 4.3e). When comparing the awake state to moderate sedation I found decreases in the DMN's sparse connections (paired two-tailed t-tests were used across 9 networks,  $t(24)=3.1175$ ,  $p=0.0423$ , Bonferroni corrected for 9 networks), a finding supported by previous work demonstrating a reorganization of key regions in the default mode network during moderate sedation (Stamatakis et al., 2010). Additionally, I found that sparsely connected regions changed their connectivity endpoints (regions they were connected to) with increasing sedation (Appendix Fig. 4.1). Specifically, when comparing awake to moderate sedation, there was an increase in the number of endpoints belonging to the DMN, limbic, and subcortical networks (paired two-tailed t-tests  $t(24)=3.1781$ ,  $p=0.0364$ ,  $t(24)=3.3516$ ,  $p=0.023$ ,  $p=0.032$ , Bonferroni corrected for 9 networks) (Fig. 4.3c and 4.3f) suggesting an increased presence of sparse connectivity towards these regions as has been reported elsewhere (Mhuirheartaigh et al., 2010; Schröter et al., 2012).



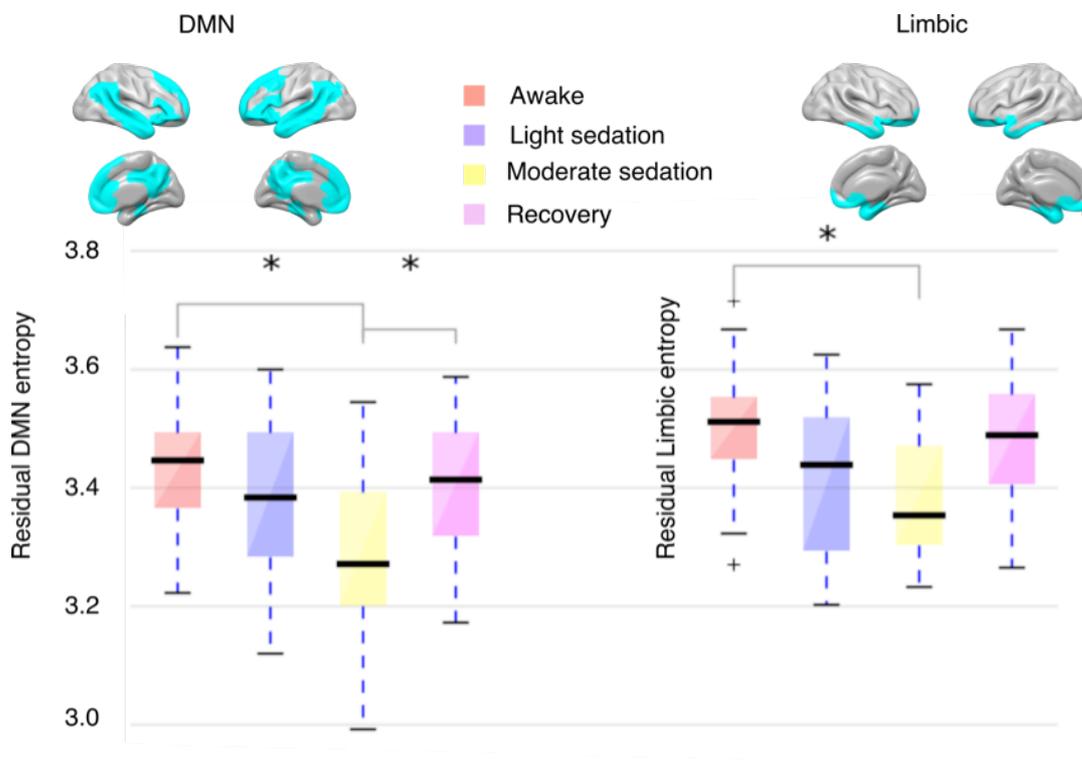
**Figure 4.3 - The functional role, location, and connectivity of sparsely connected regions during different levels of sedation.** Part (a) shows one can define the impact of certain nodes of interest to network integration. This can be quantified using  $G - G_i$  where  $G$  is the largest connected component in the original network and  $G_i$  is the size of the largest connected component in the lesioned network

after removing nodes of interest (for example random degree nodes  $G_2$  or sparsely connected nodes  $G_1$ ). Part (b) shows the impact of sparsely connected regions on network integration. Removing sparsely connected regions had a bigger impact than removing randomly chosen (rare or frequent) nodes. The difference is getting smaller as individuals become anaesthetized showing that the reduced complexity of sparsely connected regions might be linked to their losing ability to enhance network integration. Part (c) shows example locations, in the brain of one healthy volunteer, of sparsely connected regions (yellow)-i.e. regions with degrees that their value was infrequent and low. Their connections to other regions are shown in light green and their connectivity endpoints are shown in brown. In part (d) the alluvial diagram shows changes in the median of the 25 subjects (width of colour band) of the amount of sparsely connected regions in different networks with increasing sedation. In part (e) I quantified these changes by calculating the percentage of sparsely connected regions and their network membership during the awake state and moderate sedation. In part (f), I show where sparsely connected regions connect to (i.e. connectivity endpoints). For an alluvial representation of (f) I refer the reader to Appendix Fig. 4.1. In parts (e) and (f) the y-axes indicate percentages that I calculated by dividing by the number of regions in each network. For the ANOVA boxplots' thick lines show median values and whiskers represent 1.5 times the inter-fourth range for  $n=25$  subjects and single and double asterisks (\*), (\*\*) imply *post hoc* significance  $p<0.05$  and  $p<0.01$  Bonferroni corrected for multiple comparisons respectively. For parts (e), (f) barplots show mean values and errors bars show standard error of the mean. Single asterisk (\*) implies significance  $p<0.05$  Bonferroni corrected for 9 paired t-tests corresponding to the number of different networks. *Abbreviations:* DMN: Default Mode Network. DA: Dorsal Attention Network. FP: Fronto-Parietal Network. L: Limbic network. SM: Somato-Motor Network. VA: Ventral Attention Network. VIS: Visual Network. Sub: Subcortical. Cb: Cerebellum.

### ***Global and rare entropy after lesions: The importance of canonical networks***

The previous section provided evidence as to where sparse connectivity changes occurred during sedation with the DMN encompassing most of these alterations. To directly assess the importance of changes in LSN connectivity during sedation I conducted a lesion simulation by removing single networks and calculating the residual global entropy across the four sedation levels (residual entropy) (Kaiser et al., 2007; Hagmann et al., 2008; Váša et al., 2015) (Fig. 4.4a). I found an effect of sedation on residual entropy only when I removed the DMN or limbic networks ( $F(3,72)=7.84$ ,  $p=0.0002$ , bayes factor=5.5433, eta squared=0.3034,  $F(3,72)=5.26$ ,  $p=0.0016$ , bayes factor=2.8102, eta squared=0.2261 respectively). *Post hoc* t-tests for comparing moderate sedation to awake state revealed a decrease only in the DMN and limbic residual entropies ( $p=0.0048$ ,  $p=0.0273$ , Bonferroni corrected for 9 networks) (Fig. 4.4

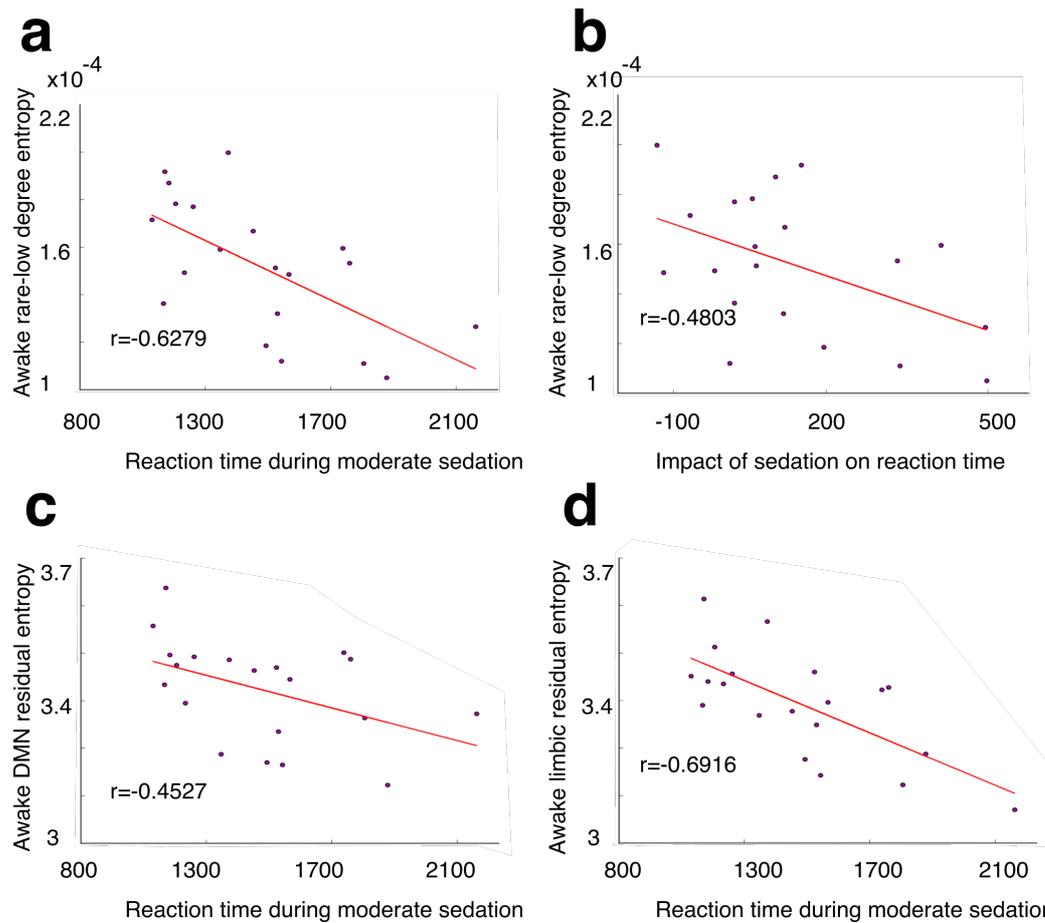
and Appendix Fig. 4.2a). I also observed an increase in DMN residual entropy (entropy following DMN removal) from moderate sedation to recovery ( $p=0.0172$ , Bonferroni corrected for 9 networks). The same effect held when considering rare entropy of the residual network alone (Appendix Fig. 4.3). These results shift the focus of the findings to the connectivity of the DMN suggesting its vital role in maintaining whole-brain entropy during sedation. I performed a further validation study by changing our network definitions using those from the Smith et al. (2009) parcellation and, again, I observed a decrease in the DMN residual entropy (Appendix Fig. 4.2b).



**Figure 4.4 - Complexity of functional connectivity after removing networks.** Boxplots display a decrease in residual global entropy during sedation after removing the DMN and limbic networks. Single asterisk (\*) implies *post hoc* significance  $p < 0.05$  after Bonferroni correction within sedation levels and for 9 networks. Boxplots' thick lines show median values and whiskers display 1.5 times the inter-fourth range. For changes in residual entropy for all networks I refer the reader to Appendix Fig. 4.3. *Abbreviations:* DMN: Default Mode Network, L: Limbic Network.

***Using entropy of sparsely connected regions to predict behaviour under sedation***

Finally, I attempted to establish a relationship between network entropy and behaviour. EEG results have shown that participants who had relatively reduced connectivity during the awake state (a close to 'drowsy' EEG signature) showed slower reaction times during moderate sedation (Chennu et al., 2016) i.e. were more affected by sedation. Since changes in connectivity will likely result in entropy (as it is computed here) changes, I investigated the relationship between rare-low degree entropy during the awake state and reaction times (from a semantic decision task) during moderate sedation (Adapa et al., 2014). I observed a significant anti-correlation ( $r=-0.6279$ ,  $p=0.0030$ ) of reaction times during moderate sedation with rare-low degree entropy during the awake state i.e. the longer the reaction time the less entropy at baseline (Fig. 4.5a). I also asked whether rare-low degree entropy during the awake state could predict differences (i.e. amount of change) in the reaction times between the awake state and moderate sedation. Here I also found a significant anti-correlation between the two ( $r=-0.4803$ ,  $p=0.0321$ ) (Fig. 4.5b) reinforcing the earlier suggestion that entropy during the awake state can successfully relate to cognitive ability during moderate sedation. Given the earlier findings, I attempted to establish whether these relationships were driven by the DMN and limbic networks. I found again an anti-correlation of DMN residual entropy with reaction times during sedation ( $r=-0.4527$ ,  $p=0.0450$ ) (Fig. 4.5c) as well as an anti-correlation of limbic residual entropy with reaction times during moderate sedation ( $r=-0.6916$ ,  $p=0.0007$ ) (Fig. 4.5d) suggesting that DMN and limbic residual entropy were most affected by sedation.



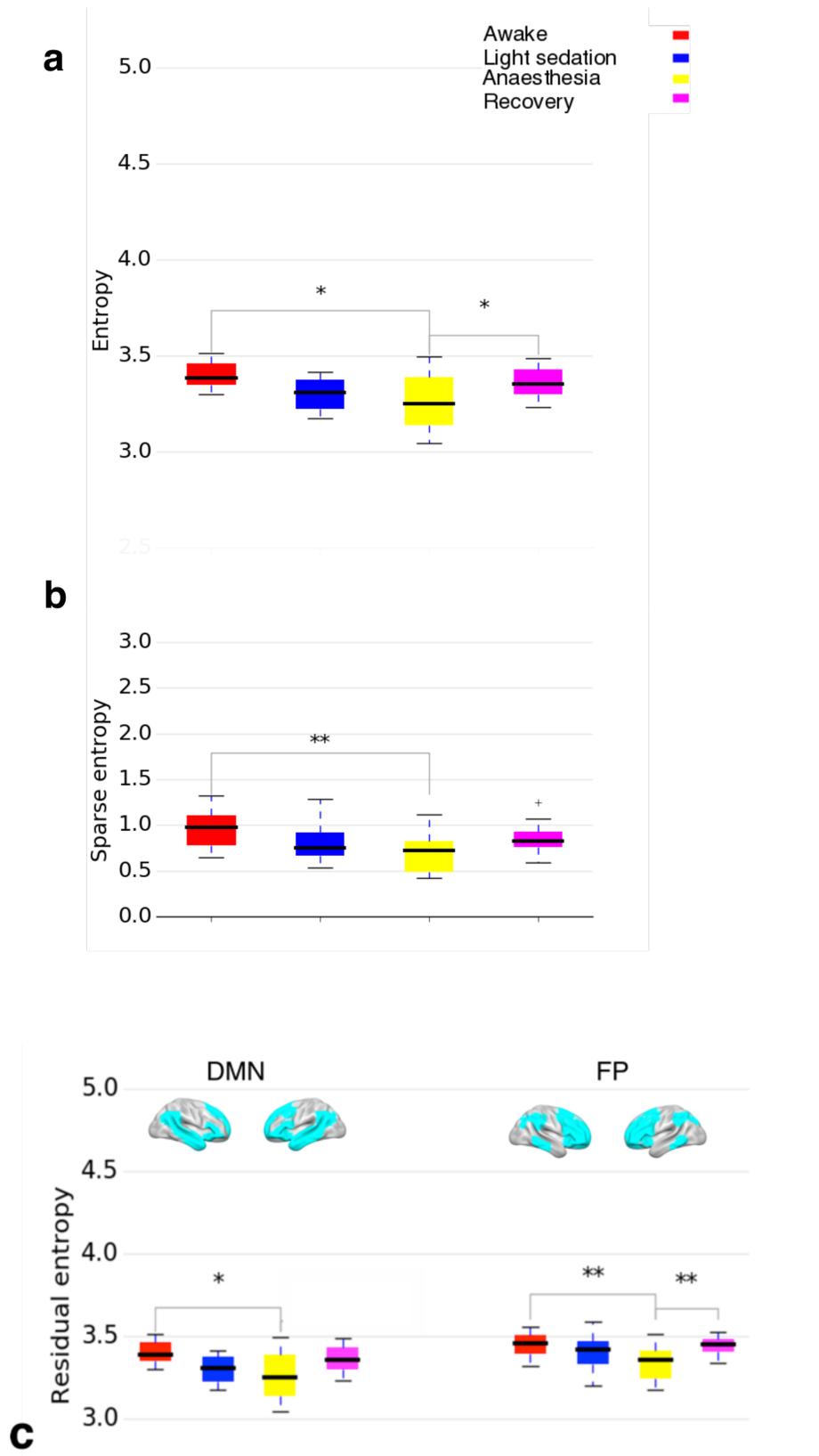
**Figure 4.5 - Individual differences in reaction times during moderate sedation correlate with rare-low and residual entropies during the awake state.** Part (a) shows that higher rare-low entropy (entropy of sparsely connected regions) during the awake state indicates faster reaction times during moderate sedation (negative correlation). Part (b) shows that rare-low entropy during the awake state correlates with the impact of sedation on reaction times i.e. the reaction times during moderate sedation minus the reaction times during the awake state. Parts (c) and (d) show that the residual entropies of the DMN and limbic networks also negatively correlate with reaction time during moderate sedation. Rare-low entropy was normalized by the length of the degree sample after removal of the high-degree nodes and is therefore on a different scale compared to the residual entropies.  $r$  values indicate Pearson correlations.

### ***Results from anaesthesia data***

I conducted a similar analysis for entropy in the anaesthesia data. As in the sedation data, I discovered a decrease in complexity that was reflected in the entropy of the rare degrees. Specifically, I observed an effect of anaesthesia level in global/whole-brain entropy ( $F(3,45)=4.96$ ,  $p=0.0038$ , bayes factor=2.4892, eta squared=0.3060). Global entropy was significantly reduced during deep anaesthesia compared to the awake

condition ( $p=0.0112$ ) (Fig. 4.6a). On the contrary, going from deep anaesthesia to recovery, there was an increase in global entropy ( $p=0.01$ ). In turn, I investigated the effect of the anaesthesia level to rare entropy. Similarly to the sedation experiment, rare entropy was affected by anaesthesia ( $F(3,45)=4.17$ ,  $p=0.0095$ , bayes factor=1.6286, eta squared=0.2704). When going from awake to deep anaesthesia, I observed a decrease in rare entropy ( $t(14)=3.1$ ,  $p=0.0073$ ), however no significant result was observed when going from deep anaesthesia to recovery (Fig. 4.6b). As previously, the frequent entropy was not significantly affected by anaesthesia ( $F(3,45)=2.23$ ,  $p=0.12$ ). In addition, rare-low entropy was more predictive of the anaesthesia level compared to the unaltered rare entropy (Mahalanobis  $p=0.03$ ) while this was not the case for the rare-high entropy (Mahalanobis  $p=0.911$ ).

For the residual entropy of different LSNs I observed that, as in the sedation experiment, DMN's residual entropy was decreased when going from awake to anaesthesia ( $t(14)=2.8891$ ,  $p=0.0112$ ) (Fig. 4.6c). However, I also observed that the FP's residual entropy was significantly decreased when going from awake to deep anaesthesia ( $t(14)=3.0844$ ,  $p=0.0076$ ) and increased when going from deep anaesthesia to recovery ( $t(14)=3.6065$ ,  $p=0.0026$ ) (Fig. 4.6c).



**Figure 4.6 - Complexity is decreased during anaesthesia.** Part (a) boxplot shows changes in the global entropy from the awake state to anaesthesia. Single asterisk (\*) and double asterisks (\*\*) indicates *post hoc* significance  $p < 0.05$  and  $p < 0.01$  respectively Bonferroni corrected. Boxplots' thick lines show

median values and whiskers represent 1.5 times the inter-fourth range from the median. Since rare is a part of the total entropy, boxplots of (a) and (b) are on a different scale. Part (b) boxplot shows changes in rare entropy alone with increasing anaesthesia. Part (c) the decreases in residual entropies of the DMN and fronto-parietal networks. Single asterisk (\*) and double asterisks (\*\*) implies *post hoc* significance  $p < 0.05$  and  $p < 0.01$  respectively Bonferroni corrected within anaesthesia levels and for 9 networks. Boxplots' thick lines show median values and whiskers represent 1.5 times the inter-fourth range. *Abbreviations*: DMN: Default Mode Network, FP: Fronto-Parietal Network.

### **4.3 Changes in complexity: from sedation to disorders of consciousness**

#### **4.3.1 Overview**

In the previous sections I quantified the complexity of sparsely connected regions and confirmed its ability to discriminate different levels of sedation and anaesthesia.

Unlike sedation/anaesthesia, disorders of consciousness are cases of loss of consciousness usually caused by severe brain injury (Laureys, 2005). Patients with disorders of consciousness usually have severely impaired awareness while their wakefulness might be spared (Di Perri et al., 2013). These cases have been broadly categorized into cases of minimally conscious state/MCS, vegetative state/VS (now also called the unresponsive wakefulness syndrome) and coma state depending on the levels of awareness and arousal (Laureys et al., 2010; Gosseries et al., 2011). Stratification of patients into these categories is usually done using clinical behavioural markers such as the CRS-R score (Gerrard et al., 2014). Despite their usefulness, it is now accepted that the gathering of elementary behavioural signs has to be complemented by brain activity measures in order to unveil brain responses not necessarily reflective in the patients' behavioural responses (Naccache, 2017). In that regard, brain-response-related clinical diagnosis of loss of consciousness in these patients can be broadly summarized in three approaches (Naccache, 2017). The first approach is to engage the patient in a paradigm that they can perform only consciously. If a patient's brain shows patterns of activity observed in conscious controls performing this task, one may infer he/she is conscious. Towards this direction, there has been important work on identifying brain responses in patients asked to "imagine play tennis/navigating in your home" (Naccache, 2006; Owen et al., 2006). A second approach is to conduct region-specific brain stimulation and look at

the brain activity in response to the stimulation (Casarotto et al., 2016). Under this framework, work by Casali and colleagues have shown that the pattern of Transcranial Magnetic Stimulation/TMS-induced EEG activity can discriminate between patients with different categories of disorders of consciousness (Casali et al., 2013). The third approach relates to investigating the brain at rest using MRI imaging techniques. The usefulness of studying resting-state activity and connectivity has been discussed in the introduction; it is a method that provides a wealth of information of spontaneous brain dynamics that does not require task-related designs (and subtractive techniques) or brain stimulations. In that regard, I focused only on resting-state fMRI data.

Conventional brain imaging studies have shown heterogeneous results, suggesting that no damage on specific brain region can be unequivocally related to the states of consciousness. For example, in some cases, brain metabolism has shown similar levels at healthy controls and patients in vegetative state (Schiff et al., 2002). On the contrary, functional connectivity analysis on resting-state fMRI data has shown some promise in the pursuit of an objective marker of the state of consciousness irrespective of brain damage (Naccache, 2017). This has been proposed as the disconnection syndrome in the sense that loss of consciousness in these patients is associated with the level of impairment in their functional connectivity (Monti, 2012). Towards this direction it has been shown that VS patients are characterised by “isolated” cognitive modules that, in the absence of global integration, do not generate conscious experience (Schiff et al., 2002; Kotchoubey et al., 2005). Further evidence suggest that VS patients exhibit reduced resting-state connectivity, as compared to MCS patients and healthy volunteers especially in key DMN regions believed to integrate information (Vanhaudenhuyse et al., 2010). Thus, decreased information integration could underlie the absence of consciousness in these patients (Tononi, 2008).

If loss of integration is a marker of absence of consciousness, how can the previous sedation/anaesthesia results discussed in this chapter inform us on loss of consciousness in these patients? Evidence from past fMRI studies suggest similarities in how functional connectivity changes during anaesthesia and injury-induced loss of consciousness. For example, functional connectivity especially in the FP and DMN networks seem to be reduced (Heine et al., 2012). Another common theme is the loss

of long-range connections in both of anaesthesia and disorders of consciousness cases, especially between frontal and posteromedial cortices that have been deemed important for network integration (Laureys et al., 2000a). In addition, functional connectivity of the thalamus, a region that its functional connectivity is believed to support whole-brain integration (Hwang et al., 2017), appears to be affected both during anaesthesia and disorders of consciousness (Laureys et al., 2000a; Guldenmund et al., 2013).

Given these similarities on the behaviour of regions responsible for network integration in both anaesthesia and disorders of consciousness, I asked whether the complexity of sparsely connected regions could discriminate between different states of consciousness. To do so I used data from patients with disorders of consciousness. The goal was to provide evidence for consciousness-specific alterations in the complexity of regions responsible for network integration. Based on previous evidence regarding the role of sparsely connected regions in network integration, I restricted my analysis on the rare-low entropy.

### **4.3.2 Materials and Methods**

#### ***Disorders of consciousness data***

##### ***Participants***

A total of 32 adults meeting diagnostic criteria of unresponsive wakefulness syndrome (also known as a vegetative syndrome) or minimally consciousness state due to severe anoxic or traumatic brain injury were include in this study. MCS patients were further subcategorized into MCS+ and MCS-. MCS+ describes high-level behavioural responses (i.e. command following, intelligible verbalizations or non-functional communication) and MCS- describes low-level behavioural responses (i.e. visual pursuit, smiling or crying to emotional stimuli etc.)(Bruno et al., 2011). Participants were in a similar age range as in the sedation experiment (range= 19-52 years). Scanning occurred at the Wolfson Brain Imaging Centre, Addenbrooke's hospital written informed consent was obtained by the individuals legally responsible for making decision on the patients' behalf. CRS-R score were also obtained for these

patients in order to clinically distinguish VS, from MCS and conscious (exit-MCS) patients (Gerrard et al., 2014). The CRS-R score assesses auditory, visual, motor, and verbal functioning, as well as communication and arousal. Each of its subscales is assigned a maximum score between two and six points, which results in a maximum total score of 23 points-corresponding to the awake state (Giacino et al., 2004).

### ***MRI acquisition and preprocessing***

Resting-state fMRI was acquired for 10 minutes using a Siemens Trio 3T scanner (Erlangen, Germany). Functional images were acquired using an echo planar sequence with the following parameter (32 slices of  $3 \times 3 \times 3.75$ mm resolution, TR=2000ms, TE=30ms, flip angle=75 degrees). T1-weighted images were also obtained with an MPRAGE sequence using the following parameters TR=2300ms, TE=2.47ms, 150 slices of resolution  $1 \times 1 \times 1.2$ mm. The preprocessing pipeline consisted of similar steps as in the sedation and anaesthesia experiments. However due to the brain distortion caused by injury, I observed unacceptable normalisation in 6 subjects. Normalisation checking was conducted by visually contrasting the MNI template with the normalised brains (using the *checkreg* button in SPM) and unacceptable normalisation was considered when I observed abnormal skewing in a particular dimension. By utilizing a two-step process 1) creating a template out of the distorted brains and mapping the template to MNI space 2) wrapping individual brains to the template and, in turn, to MNI space. Creating the template was done by using the *buildtemplateparallel.sh* function implemented in the package ANTS (<http://stnava.github.io/ANTs/>). Normalisation of the template was conducted using the non-linear transformation *SyN* also implemented in ANTS. To wrap individual brains to MNI space I wrapped them first to the template and then, using the deformation field obtained from the template normalization, I mapped them to MNI space. Because of the presence of severe deformations caused by brain injury, 2 out of the 6 subjects were excluded despite efforts in assisting with the normalisation.

***Regions of interest***

I used the eAAL atlas as described in the sedation/anaesthesia paradigm section.

***Functional connectivity networks and complexity of sparsely connected regions***

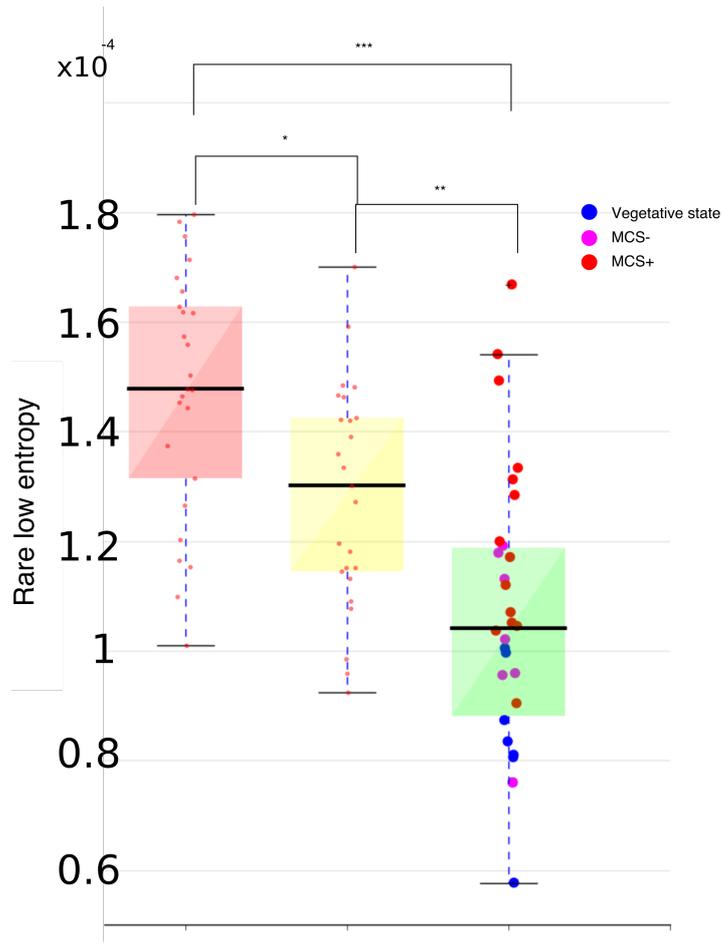
I used the CONN functional connectivity toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2011) to obtain the time series from the previous ROIs. A strict temporal preprocessing pipeline of nuisance regression was applied. Linear regression confounds included *CompCor* components attributable to white matter and CSF signals (Behzadi et al., 2007), subject-specific six realignment parameters and their first-order temporal derivatives. In addition a highpass filter of 0.009Hz was applied to remove scanner noise. Connectivity matrices were obtained by taking the pairwise Pearson correlations between each pair of ROIs. Functional connectivity matrices were obtained as previously and thresholded using the same range of thresholds  $\tau$  as in the sedation and anaesthesia experiments. Rare-low entropy was obtained as previously by identifying sparsely connected regions in the functional connectivity networks. Rare-low entropy was averaged over the range of thresholds  $\tau$ .

It is worth noting that the similar age range and acquisition parameters (both datasets acquired in WBIC, Cambridge) of the sedation and DOC data allowed me to compare results from these two different experiments. In that regard, in this section I did not include data from the anaesthesia experiment as this was obtained with a different setup, something that could potentially introduce bias in the comparisons between different experimental data.

**4.3.3 Results**

I observed a monotonic decrease of rare-low degree complexity from disorders of consciousness to the awake state acquired during the sedation experiment (Fig. 4.7) ( $F(2,77)=20.97$ ,  $p<0.0001$ , bayes factor=12.7364, eta squared=0.4496). Specifically, I observed a statistically significant decrease between the awake state and moderate

sedation ( $p=0.028$ ), patients with disorders of consciousness and awake state ( $p<0.0001$ ) and patients with disorders of consciousness and moderate sedation ( $p=0.0015$ ). This result suggested that the complexity of sparsely connected regions was significantly reduced in patients with disorders of consciousness compared to sedation and the awake state showing that their brains might be further segregated.



**Figure 4.7 - Complexity of sparsely connected regions from sedation to disorders of consciousness.**

A monotonic decrease in the complexity of sparsely connected regions was observed when going from awake to moderate sedation and disorders of consciousness. We also show complexity across different categories of disorders of consciousness (vegetative state, MCS+, MCS-). Boxplots' thick lines show median values and whiskers represent 1.5 times the inter-fourth range. Single asterisk (\*), double asterisk (\*\*) and triple asterisks (\*\*\*) imply *post hoc*  $p<0.05$ ,  $p<0.01$ , and  $p<0.001$  significance respectively Bonferroni corrected. *Abbreviations:* DOC: Disorders of Consciousness.

#### 4.4 Discussion

Network models suggest that functional connectivity is far from random but that it is characterized by specific organization (Bullmore and Sporns, 2012). The complexity of functional connectivity networks is considered to promote local and global efficiency in communication through the existence of heterogeneously connected regions (Zamora-López et al., 2016). In other words, functional connectivity networks of the awake brain might be described by a “sweet spot” of complexity critically balancing segregation and integration. Evidence for this critical complexity was shown in chapter 3. In this chapter, I hypothesized that alterations in complexity could discriminate between different states of consciousness, as these would mark a departure from this criticality towards a less complex and more segregated configuration.

In that regard, previous work has hypothesized that loss of consciousness is associated with a less complex and more stereotypic pattern of brain interactions (Tononi and Edelman, 1998; Alkire et al., 2008). I observed a decrease in whole-brain complexity during sedation suggesting that brain connectivity was characterized by a diminished repertoire of degrees. This result was further confirmed with the anaesthesia dataset. Diminished degree repertoire implies a more segregated network organization where information can still be communicated efficiently at the local/regional level but not at the global/whole-brain level (Power et al., 2010). I attempted to provide more direct evidence about how integration is lost by looking at local changes in complexity.

I showed that the observed decrease in whole-brain complexity during sedation and anaesthesia was driven by regions that have a rare number of connections. I believe that this result has important implications for how propofol affects brain connectivity. First, the diminished repertoire in the degrees of few brain regions suggests that the anaesthetic has more severe effects on a limited number of regions. However, it is not clear why this selective action can be associated with loss consciousness. Contemporary theories suggest that the selective action of propofol on areas crucial for network integration may be enough to cause loss of consciousness (Alkire et al., 2008). Alluding to this hypothesis, I showed that sparsely connected regions identified

within the set of regions with rare degrees were important for network integration. Expanding on this argument I used only the complexity of these regions to show that it can discriminate between sedation or anaesthesia levels better than using the degree of highly connected regions. Building on the work of Granovetter regarding the importance of weak ties in large-scale social structures (Granovetter, 1973), sparsely connected regions have been deemed important for bridging different network components thus providing efficient information exchange between distant parts of the brain (Markov et al., 2011; Gallos et al., 2012; Park and Friston, 2013; Goulas et al., 2015). Furthermore, a different study reported a decrease in rare, long-range connections accompanied by loss of whole-brain integration during propofol-induced loss of consciousness (Schröter et al., 2012). Contributing to this, I showed that sparsely connected regions had an important role in network integration during the awake state. As sedation increased, this role was diminished suggesting that their decreased complexity might be directly linked with their inability to integrate information. This can be one explanation for why the observed decreased complexity of sparsely connected regions during sedation is associated with loss of consciousness.

I also found that sparsely connected regions were significantly decreased in the DMN during moderate sedation. Studies have shown that the DMN, besides highly connected regions, includes regions with sparse connectivity (such as rare, long-distance connections) (Smallwood et al., 2012; Vatansever et al., 2015b), however their role in loss of consciousness has not been established before. In light of our previous observations regarding the importance of sparse connections, I speculated that the DMN during sedation is deprived from its ability to integrate information from far and diverse sources.

An additional account for why the complexity of sparsely connected regions might be linked to loss of consciousness can be recapitulated in terms of loss of information capacity (Tononi, 2004; Schrouff et al., 2011). For example, observed alterations in subcortico-cortical interactions deemed critical for maintaining wakefulness (Brown et al., 2011), can lead to impaired consciousness because they can reduce the capacity of the brain to process information (Alkire et al., 2008). This work provided additional explanatory power for this approach. First I showed that during moderate sedation,

the connectivity of sparsely connected regions spanned DMN and subcortical regions significantly more than the awake state, aligning with previous evidence of sparse connectivity between DMN and subcortical regions during decreased consciousness (Stamatakis et al., 2010; Guldenmund et al., 2013). At the same time this sparsely connected sub-network of cortico-cortico and subcortico-cortical regions was characterized by decreased in rare-low degree entropy, potentially reflecting its decreased capacity to process information (Brunel, 2016). Further research can focus on the importance of sparsely connected regions especially in thalamocortical loops that have been deemed critical for regulating consciousness (Ching et al., 2010). Interestingly the thalamus has been shown to have an enhanced role in network integration (Hwang et al., 2017). A relationship of this with our data (in the context of reduced complexity of sparsely connected regions) is something that can be investigated in the future.

In addition, I quantified the importance, from a complexity perspective, of canonical networks during sedation. Here I observed that the DMN residual entropy was significantly decreased during sedation. An association of impaired connectivity in the DMN and loss of consciousness has been previously reported (Guldenmund et al., 2016). I showed that virtually removing the DMN was associated with a significant decrease in entropy during sedation suggesting its importance in whole-brain complexity during loss of consciousness. This result, in conjunction with the fact that the number of sparsely connected regions within the DMN were significantly decreased, comes in line with the previous observations about the diminished role of DMN in whole-brain integration during loss of consciousness.

Further results revealed a significant decrease in the limbic residual entropy. This network-as defined by Yeo et al. (2011)-consists of ventral frontomedial and orbital regions as well as inferior temporal and parahippocampal regions. Orbitofrontal regions have been associated with engaging visceromotor control (Lindquist and Barrett, 2012) while parahippocampal regions have been linked to conscious sensory perception (Li et al., 2012) both features of conscious experience. Interestingly, the residual entropies of motor and visual networks were not affected during moderate sedation in alignment with evidence showing that primary sensory cortical activity is

not modulated by sedation (Plourde et al., 2006; Boveroux et al., 2010). It is worth noting that the residual entropy following the virtual removal of the DMN was the only one that significantly increased when comparing recovery to moderate sedation. A previous study has shown that the connectivity of DMN-specific regions is restored during recovery (Långsjö et al., 2012). Accordingly, I expected the DMN to re-obtain its connectivity pattern during recovery.

During full anaesthesia, I observed that the DMN residual entropy was also decreased. However, the limbic network was not implicated in this case and instead was replaced by the FP possibly indicating that its connectivity is significantly affected during deep anaesthetic levels. Studies from propofol-induced anaesthesia have shown that the connectivity of the FP is reduced (Boveroux et al., 2010). One reason for why deep levels of anaesthesia might be associated with impaired FP connectivity is that anaesthesia affects connectivity between frontal and parietal regions (Hudetz, 2012) and thalamo-frontal connectivity (Alkire et al., 2008) thus impairing network integration. Can these findings be reconciled with the residual entropy results shown in sedation? One working hypothesis is that of graded suppression of connectivity: the FP is relatively intact in cases of sedation but affected during deep anaesthesia (Martuzzi et al., 2010). However, further evidence is required to confirm this hypothesis.

In addition, by tracking individual behavioural data during moderate sedation, I discovered that entropy during the awake state was a potential marker of impaired reaction times during sedation. This result adds to a recent EEG study showing the potential of alpha band network connectivity as a predictor of unresponsiveness during propofol-induced sedation (Chennu et al., 2016). This result might signify that behaviour during sedation or anaesthesia can be predicted by the complexity of our brains during the awake state. I believe that this can have important implications for further understanding individual responses under anaesthesia and anaesthesia monitoring in general. The correlations, however, presented here will require subsequent replication given their exploratory nature.

Finally, I used the complexity of sparsely connected regions to differentiate between sedation and disorders of consciousness. Specifically, I showed that when going from awake to sedation and disorders of consciousness complexity is reduced. A particular question is how comparable these states are. For example vegetative state impairs only awareness while anesthesia appears to affect both awareness and wakefulness.

One approach for this is to consider that complexity gauges information about integration of specialized modules in the brain; thus what we believe that complexity is capturing aggregate information that combines features of different systems responsible for awareness and wakefulness. Towards this direction, previous studies using complexity of EEG signals across the whole brain (a measure that they called perturbational index) have suggested that complexity can classify between consciousness states in a way that goes above and beyond the dichotomy of awareness and wakefulness (Casali et al., 2013). In this study Casali and colleagues show that patients at the vegetative state have low perturbational index despite their high behavioural arousal. At the same time, the perturbational index of these was comparable or lower to the one observed in healthy individuals under anaesthesia suggesting different states of consciousness impairment. Together with the results presented here, such findings might provide additional directions as to how these different states can be compared against each other, possibly delineating a “continuum of consciousness” that cannot be captured by the standard wakefulness/awareness quantification.

Additional speculation as to why the complexity measure presented in this thesis could discriminate between different cases of loss consciousness in a “continuum” fashion pertains to the fact that reduced complexity of sparsely connected regions reflects the effect of loss of consciousness on long-range connectivity, something that has been observed both in anaesthesia and disorders of consciousness (MacDonald et al., 2015). Another interesting hypothesis that could complement the results presented here is that of the impairment in thalamocortical connectivity. Thalamocortical connectivity has been shown to be affected in disorders of consciousness (Monti et al., 2015) while its restoration has been associated with emergence from vegetative state (Laureys et al., 2000a). Thus it is possible that alterations in the complexity of these

regions might affect whole-brain complexity. By combining the findings from sedation and disorders of consciousness, I showed that the complexity of sparsely connected regions was able to classify between anaesthetic-induced and clinical loss of consciousness. This finding vindicates the claim that the complexity of functional connectivity is significantly downgraded in patients with disorders of consciousness compared to the anaesthetic state or the healthy brain; something that, as I discussed previously, might be attributed to severe impairment in network integration (Casali et al., 2013; Sarasso et al., 2015). In addition, I showed that complexity could classify between anaesthetic conditions and states of consciousness in patients based on resting-state functional connectivity. This is particularly important in cases where clinical assessment using traditional behavioural markers is difficult to obtain and usually confounded by other factors such as motion reflexes (Monti, 2012).

#### **4.5 Conclusion**

Collectively, I showed shifts in the complexity of functional distribution under different states of consciousness. Specifically, I showed these alterations might reveal a shift towards less integrated/more segregated networks. These are known to facilitate global communication, thus alterations in their connectivity might reflect loss of communication in the whole brain. One point that still needs to be elucidated relates to whether changes in functional connectivity and its complexity are static or they change concomitant with the changes in brain dynamics. Complexity is calculated on the premises of an average functional connectivity over the scanning interval. But what happens at each point of time? In the next section my focus will be on the brain dynamics during anaesthesia and disorders of consciousness and to see whether complexity at the temporal domain can also inform us about different states of consciousness.

## **Chapter 5: Dynamic functional connectivity and its complexity during loss of consciousness**

In the previous chapter I showed that the complexity of functional connectivity networks could differentiate between states of consciousness in both anaesthetic-induced loss of consciousness and disorders of consciousness. One assumption of this method is the representation of functional connectivity (and its properties) as an average over the scanning time with a secondary assumption that the connectivity values remain almost constant over time (Cole et al., 2010). Although this is a convenient framework for analysing functional connectivity, given the dynamic nature of brain activity, even during rest, it is expected that functional connectivity will change across time (Hutchison et al., 2013). These changes are not just fluctuations of a random process but have been shown to have a neurobiological basis potentially reflective of diverse brain dynamics (de Pasquale et al., 2010). Thus, besides the static methods for studying resting-state functional connectivity introduced earlier in this thesis, recent dynamic approaches have been developed in order to study and quantify the *temporal* variations/changes in functional connectivity (Hutchison et al., 2013).

The static and dynamic approaches should not be adopted in a manner exclusive to each other but as complementary to each other when it comes to studying loss of consciousness. The temporal dynamics of functional connectivity are considered important in the context of consciousness. The repertoire of functional connectivity across time is believed to reflect the number of accessible brain states across time, something that has been linked with the brain's capacity of conscious processing (Carhart-Harris et al., 2014; Tagliazucchi et al., 2014). The temporal dimension of these changes is often regarded as the temporal aspect of complexity (complexity over time) that complements the (spatial) complexity of functional connectivity in order to support the richness of conscious experience (Tononi, 2004).

Under this premise, previous studies have begun investigating the temporal changes of functional connectivity during loss of consciousness. Previous studies have shown abnormal dynamic functional connectivity changes during anaesthesia (Barttfeld et al.

2015a; Hudetz et al., 2015) and disorders of consciousness (Di Perri et al., 2018) showing alterations in brain dynamics otherwise not seen with the static analysis. Thus precise characterization of the temporal dynamics of functional connectivity is a useful way of understanding the temporal dimension of loss of consciousness and of brain function.

As in the (spatial) complexity of static functional connectivity where I computed distribution of connectivity across different regions, complexity of dynamic functional connectivity (or temporal complexity) can be derived from the distribution of brain connectivity across time as the brain fluctuates between discrete states with different connectivity patterns (Baker et al., 2014). Thus the goal of this experimental chapter is to use a dynamic functional connectivity model for extracting relevant states and, in turn, for studying the complexity of dynamic connectivity in cases of anaesthetic-induced unconsciousness and disorders of consciousness.

## 5.1 Introduction

The vast majority of fMRI studies assess differences in functional connectivity by calculating BOLD correlations with the assumption that functional connectivity is static. While this is true in some cases, studies have shown that the brain has richer temporal dynamics that cannot be detected using static functional connectivity methods (Hutchison et al., 2013).

In the context of fMRI, with a sampling resolution not as rich as that of EEG or MEG studies, a plethora of methods, labelled as dynamic functional connectivity (dFC), have been developed to quantify how functional connectivity changes across time. The most common one is the sliding window that results in different functional connectivity matrices per a window of time and assesses their changes over time (Hutchison et al., 2013).

Why is assessing dynamic functional connectivity important in the context of loss of consciousness? Besides alterations in spatial complexity, current theories postulate that the state of consciousness is related to the diversity or repertoire of states that are

available for the brain to access (Tononi, 2004; Carhart-Harris et al., 2014; Tagliazucchi et al., 2014). This has also been framed under the context of “richness” in conscious experience that an unconscious brain fails to attain (Barttfeld et al., 2015a). Under this premise, a decreased repertoire in functional connectivity between different windows, would potentially be associated with a decreased level of accessible states and an impaired state of consciousness (Tagliazucchi et al., 2014).

In that regard, evidence for reduction of the repertoire of functional connectivity has been shown in anaesthetized rats (Ma et al., 2017; Paasonen et al., 2018) and macaques (Barttfeld et al., 2015a). In cases of patients with disorders of consciousness, Di Perri and colleagues found disrupted within and between networks dynamic functional connectivity (Di Perri et al., 2018). Thus initial evidence suggests that dFC is altered in cases of anaesthetic-induced and disorders of consciousness.

However, dFC assessed with the commonly used sliding window method has its disadvantages, the most important being whether functional connectivity obtained at different windows or time reflects distinct brain dynamic patterns (Preti et al., 2017). To overcome this obstacle, new approaches have been introduced that allow the separation of brain dynamics into discrete states at which brain activity recurs during different time points (Baker et al., 2014). A Markov chain is an appropriate mathematical framework for modelling transitions between a number of discrete states (Ross, 1966). In that regard, the Hidden Markov model (HMM) assumes that the brain undergoes transitions between states that can be quantitatively described as transitions between states in a Markov chain (Vidaurre et al., 2017a; Vidaurre et al., 2017b). Brain dynamics and their repertoire can then be evaluated on the premises of visiting and switching between different states.

In light of the advantages in using HMM to explaining brain dynamics, I utilised this technique to address the temporal dynamics of functional connectivity in sedation and disorders of consciousness. The first goal of this chapter was to show that the repertoire of brain dynamics is related to different states of consciousness. In that regard, I identified states that characterized brain dynamics in each population separately. I then related each individual’s brain dynamics with the level of

responsiveness in cases of propofol-induced sedation or the CRS-R score in the case of patients with disorders of consciousness.

In turn, as in the case of complexity reported in the previous chapter, I wanted to see whether the complexity of these brain dynamics mirrored a consciousness- dependent spectrum. To do so, I derived states common to both populations and I considered the switching rate between all the derived states as a proxy for temporal complexity. I then showed that this could discriminate between sedation and disorders of consciousness.

## **5.2 Materials and methods**

### **5.2.1 Participants**

Details for the datasets used for this experimental chapter have been described in chapter 4.

### **5.2.2 Image acquisition and preprocessing**

Image acquisition and preprocessing details have been described in section 4.2 of chapter 4.

### **5.2.3 Hidden Markov Model**

Here I describe the basic premises of the Hidden Markov Model that was used to derive brain dynamics. At each time point  $t$  of brain activity, the observed BOLD time series data was modelled as a mixture of multivariate Gaussian distributions. Each one of the multivariate Gaussian distributions corresponded to a different state  $k$  and was described by first-order and second-order statistics, mean  $\mu_k$  and covariance  $\Sigma_k$  (interpreted as activity and connectivity of each state respectively). Using some notation, if  $x_t$  describes the BOLD data at each time point  $t$ , then the probability of being in state  $k$  is assumed to follow a multivariate Gaussian distribution

$$P(\mathbf{x}_t | s_t = k) \sim \text{Multivariate Gaussian}(\mu_k, \Sigma_k)$$

In turn, I wanted to model how transitions between states take place. The basic Markovian principle that describes the transition between states, assumes that the probability of the data being in state  $k$  at time  $t$  relates only to the probability of being in state  $l$  at time  $t - 1$ . This can be described by the following equation

$$\Pr(s_t = k) = \sum_l \theta_{l,k} \Pr(s_{t-1} = l)$$

where  $\theta_{l,k}$  is the transition probability from state  $l$  to state  $k$ .



**Figure 5.1 - Deriving brain dynamics from resting-state data using the Hidden Markov Model.** (a) ROI time courses were obtained for each participant using a parcellation scheme described earlier (eAAL) and were concatenated across participants. (b) The HMM model assumed that BOLD data is organized into discrete states that the brain occupies (and switches in-between) at each sampled time point. Each time point  $t$  of brain activity is characterized by a mixture of multivariate Gaussian distribution describing the probability of the brain being active in each state. (c) The HMM model that runs on the concatenated ROI data results in a number of states each one characterized by a mean

$\mu_k$  (activity) and covariance  $\Sigma_k$  (connectivity). Fractional occupancy/FO of state  $i$  was defined as the mean probability of that state being active over the duration of the scanning. Switching rate between states  $i$  and  $j$  was defined as the difference in their activation probabilities in consecutive time points i.e. how much a brain that was active on state  $i$  at time  $t$  would switch to being active on state  $j$  at time  $t+1$ .

Together, based on this modelling, HMM infers the  $\Pr(s_t = k)$  probabilities for each state  $k$  and time  $t$  (state time courses) as well as the transition probabilities  $\theta_{l,k}$  and the statistics of each state  $(\mu_k, \Sigma_k)$  that best describe the BOLD data. To make inference tractable, a variational Bayes algorithm was used that works by minimizing the Kullback-Leibler divergence between the real and the modelled data (Wainwright and Jordan, 2008).

After obtaining the time course for each state, additional quantities related to the temporal characteristics of each state could then be obtained for each participant. One important quantity of the HMM model is the average time spent on each state (fractional occupancy-FO) quantified in terms of the average probability of this state being active across time. The higher the FO the more time the brain is active in that state. In turn, the switching rate from state  $i$  to state  $j$  was defined as the number of times a brain goes from activating state  $i$  with probability  $\Pr(s_t = i)$  to activating state  $j$   $\Pr(s_{t+1} = j)$  i.e. the absolute difference  $|\Pr(s_t = i) - \Pr(s_{t+1} = j)|$  for all time points  $t$  (normalized by the total number of switches).

The input data for the HMM model was the following. I obtained a matrix of dimensions: *number of images pooled for all conditions*  $\times$  (*number of subjects*  $\times$  *number of ROIs*) =  $(600 - 6) \times (25 \times 132)$ , by concatenating the awake, light sedation, moderate sedation, and recovery data from all volunteers (after discarding the first six volumes) with the purpose of finding common states characterizing the sedation experiment. Prior to concatenation, the participant-specific time courses were standardized (centered and with standard deviation = 1). To reduce noise in the matrix, I used a principal component analysis (PCA) on the concatenated time courses and I kept the top 25 principal components that explained approximately 70% of the signal variance. The number of states was chosen as  $K = 12$ . Although the number of states is arbitrary, a high number of states might result to overfitting

whereas a low number of states might not be enough to explain the variety of brain dynamics (Vidaurre et al., 2017b). In addition, I used a number of states  $K = 12$  in order to be consistent with previous work using similar models thus enhancing reproducibility and interpretability of the results (Vidaurre et al., 2017b). It is worth noting that although the HMM model works on data at the group level, the state time courses and all consequent metrics (FO, switching rate) were obtained at the subject level. For the disorders of consciousness cohort, the matrix was of dimensions  $(300 - 6) \times (30 \times 132)$  by concatenating the resting-state data across all the patients.

For both experiments, mean activity  $\mu_k$  for each state was converted back to the original 132 regions using the PCA loadings obtained from the PCA decomposition mentioned previously. For connectivity, in order to obtain the spatial maps, I calculated the weighted degree of each region (row-wise sum of correlations for each region) and I converted it back to 132 regions by using the PCA loadings. To help with interpretation of the results, the spatial maps were thresholded by keeping the top 40% of values (activity or connectivity).

Because the HMM states obtained for the sedation and DOC experiments were different, it prevented me for drawing common conclusions about similar brain dynamics that would be able to characterize and discriminate between different states of consciousness. To this end, I wanted to obtain a common HMM model for the two experiments in order to obtain similar states that would characterize *both* of the experiments. In that regard, I concatenated the data from the two experiments along the temporal dimension thus obtaining a matrix of size:

*total number of images*  $\times$  *number of subjects*  $\times$  *number of ROIs*  $(594 \times 25 + 294 \times 30) \times 132$ ).

This was the input to the common HMM model and resulted in a set of common states for both of the experiments that were later analysed.

For the common HMM model, instead of focusing on the FO and switching rate of each state, I quantified the switching rate between all states as a proxy measure of temporal complexity with the hope that it would discriminate between states of

consciousness. This was quantified as the average switching rate over all pairs of states. High switching rate between all states in a Markov chain implies that, at each time point, there is a high probability of switching from the current state to any other state; thus transitions would take place almost randomly between every state. This has also been termed as temporal instability of the Markov chain (Vidaurre et al., 2017b). In that regard, starting from a state  $i$ , one will visit only a certain number states after a number of transitions, while other states would remain unvisited (Ross, 1966). In analogous fashion to the concept of network (spatial) complexity, increases or decreases in switching rate can be regarded as a proxy of temporal complexity in the Markov chain. Viewing the Markov chain as a graph with nodes being states and edges being transitions, high switching rate implies that information is unstable/noisy and thus can be communicated among a restricted number of nodes, analogous to a less complex brain network.

#### **5.2.4 Statistical analysis**

Data was tested for normality using the Kolmogorov-Smirnov test. Statistical analysis for assessing the effect of sedation on fractional occupancy was conducted using a one-way ANOVA. P values were *post hoc* Bonferroni corrected. For correlation-related analyses,  $r$  and  $p$  values indicated Pearson's correlation coefficient and statistical significance. I also report effect sizes for ANOVA using the bayes factor and the eta squared quantities.

### **5.3 Results**

#### **5.3.1 Brain dynamics during sedation**

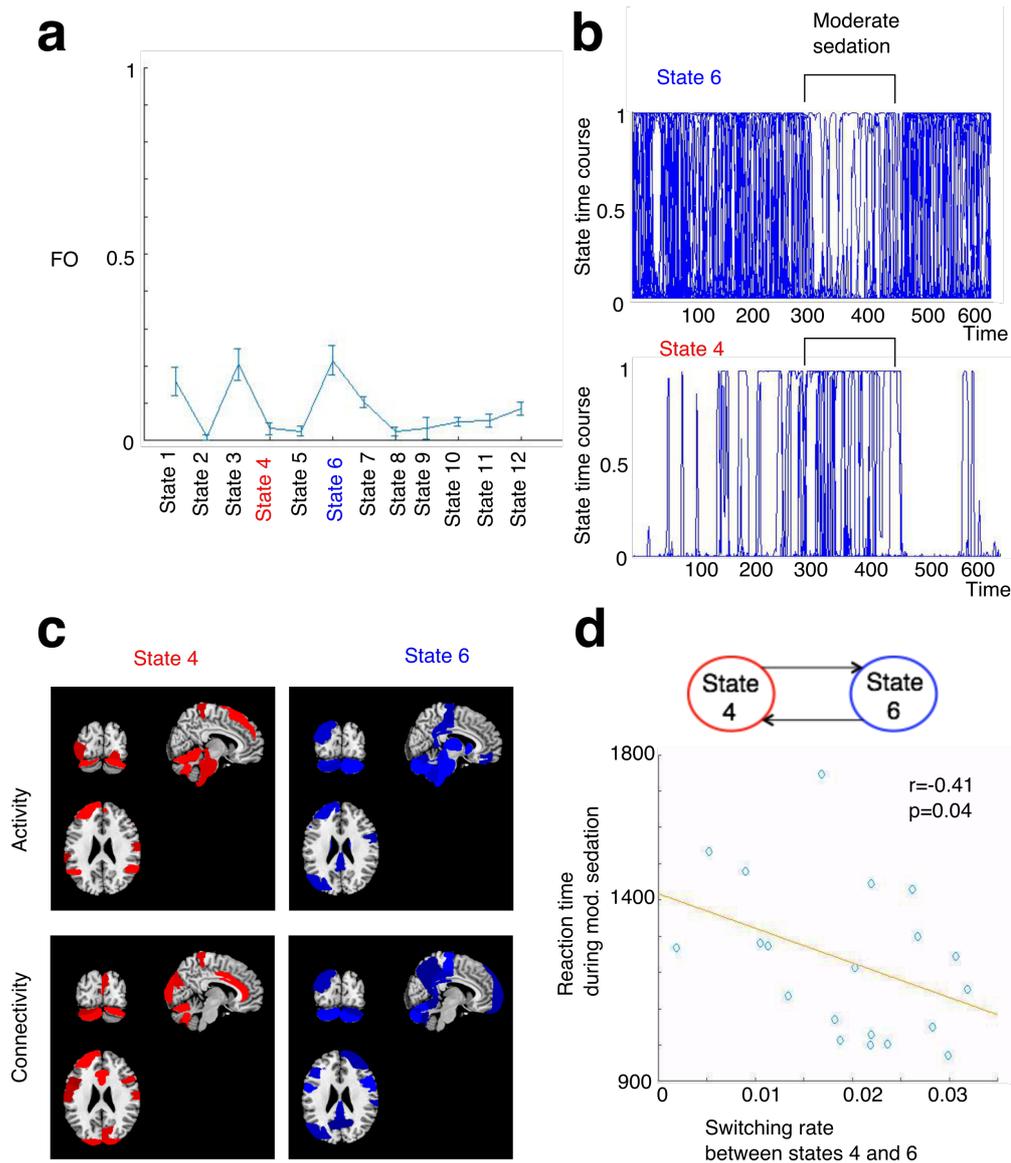
I applied the HMM model with  $K = 12$  states and number of PCA components constrained to 25 to the sedation data presented in chapter 4. This resulted in a number of states each one associated with a vector of probabilities indicating the brain being active at that state.

First I plotted the FO occupancy i.e. the average (over time) probability of each participant's brain being active in that state (for a full list of states see Appendix Fig. 5.1). I observed that two states had the highest mean FO over individuals (state 3 and state 6)(Fig. 5.2a). As FO was derived on the basis of the whole scanning, I wanted to better understand the FO of each state during each experimental condition. To do so, for each individual I plotted the state time courses for each state and I evaluated the FO of each state at each condition separately. By plotting state 6's time course for all individuals, I observed that state 6, despite its high FO, it was rarely visited during sedation (Fig. 5.2b). In other words, the high FO observed for state 6 across the scanning was a trait of the awake, light sedation, and recovery conditions and not the sedated condition. I quantified this effect by measuring state 6's FO across the four conditions separately (awake, light sedation, moderate sedation, recovery). I observed an effect of the sedation level on the FO of state 6 ( $F(3,72)=4.59$ ,  $p=0.0048$ , bayes factor=2.0886, eta squared=0.2032). Specifically, I observed that state 6's FO during moderate sedation was significantly less compared to awake state ( $p=0.0076$ ) and recovery ( $p=0.0242$ ). State 6's activity included parts of the DMN such as the PCC as well as brainstem and cerebellar regions. State 6's connectivity also included parts of the DMN such as the PCC and the precuneus, as well as cerebellar regions (Fig. 5.2c). Similar analysis for state 3 showed non-significant results potentially indicating that it was capturing more subject-specific dynamics rather than sedation-level-dependent dynamics.

It is noteworthy that despite its overall low FO across the scanning session, state 4 was the one mostly visited during sedation, as it was evident by its time course (Fig. 5.2b). Similarly to state 6, I quantified this by measuring the effect of sedation level on the average FO of state 4 for each of the four conditions ( $F(3,72)=3.43$ ,  $p=0.02$ , bayes factor=0.8390, eta squared=0.1601). The FO during moderate sedation was higher when compared to the awake state ( $p=0.042$ ) and recovery ( $p=0.03$ ). State 4's activity pattern included frontal and precentral regions, parts of the cerebellum and the brainstem. State 4's connectivity pattern included the anterior parts of the DMN as well as occipital and cerebellar regions (Fig. 5.2c).

In conclusion states 4 and state 6 showed the highest FO during sedation and non-sedation conditions respectively. Thus I regarded these states as the most relevant states of brain dynamics in this experiment.

If these two states reflect opposite features of brain dynamics during sedation then it is possible that switching between deactivating state 6 and activating state 4 is crucial for how the brain behaves during sedation. Preliminary evidence comes from previous studies suggesting that alterations in the dynamic interaction between states/networks determine how anaesthetized/unresponsive an individual would be (Guldenmund et al., 2016). I thus hypothesised that the switching rate between states 4 and 6 would be reflective of the level of the participants' behavioural performance during sedation. In that regard, I found that their switching rate was negatively correlated with reaction times during sedation ( $r=0.41$ ,  $p=0.04$ ) (Fig. 5.2d).



**Figure 5.2 - Brain dynamics during sedation.** (a) State 6 was the one that showed higher occupancy throughout the duration of the sedation experiment. (b) However, during moderate sedation, state 6 was rarely visited and this was evident by the probability time courses for each individual (blue lines). On the contrary, state 4 was mostly visited during moderate sedation. (c) State 6's activity encompassed region such as the PCC as well as brainstem and cerebellar regions. State 6's connectivity included regions belonging to the DMN such as the PCC and the precuneus. State 4's activity and connectivity mostly included frontal and occipital regions. (d) Switching between state 6 and state 4 correlated with how responsive individuals would be during sedation. For both activity and connectivity spatial maps slices at  $x=98$   $y=41$   $z=97$  are displayed.

Collectively these results showed that the temporal dynamics during sedation were characterized by two opposite effects: low FO in state 6 and high FO in state 4. The switching rate between the two states correlated negatively with reaction times

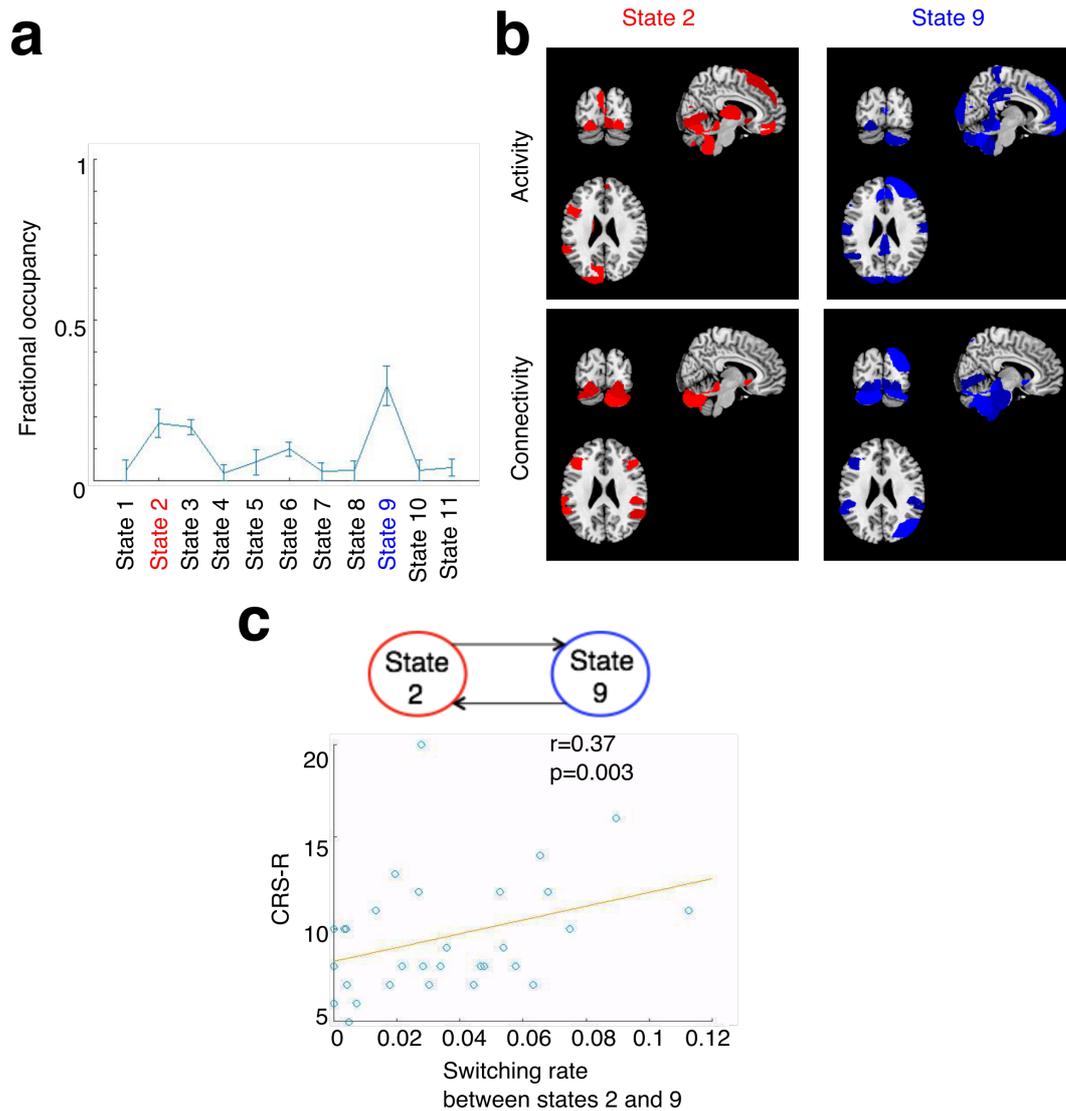
implying that the higher the switching rate the more responsive the individuals were during sedation. This fits with theories arguing that the dynamic toggling between networks can potentially determine the level of responsiveness when consciousness is lost (Di Perri et al., 2016; Lin et al., 2017).

### **5.3.2 Brain dynamics in patients with disorders of consciousness**

Next I investigated the brain dynamics in the disorders of consciousness cohort presented in chapter 4. I used similar parameters ( $K = 12$  states,  $PCA = 25$ ) as previously; however, the HMM model rejected 1 state as it was visited with a rate close to 0. Out of the 11 remaining states, states 2 and 9 showed the highest mean FO (Fig. 5.3a) (for a full list of states see Appendix Fig. 5.2).

State 9's activity pattern included cerebellar and inferior frontal regions as well as parts of the DMN network such as the PCC and the MPFC. State 9's connectivity pattern included brainstem and cerebellar regions. State 2's activity pattern included dorsolateral-prefrontal, and orbitofrontal regions while its connectivity pattern included mostly cerebellar regions (Fig. 5.3b). I considered those to be the most relevant states of brain dynamics during disorders of consciousness.

As it has been previously reported in disorders of consciousness, changes in the connectivity between networks across time might serve as marker of the state of consciousness (Di Perri et al., 2018). I thus hypothesised that the inter-dynamics of states 2 and 9 would be critical for regulating consciousness. To do so, I correlated the switching rate between states 9 and 2 with the CRS-R score. I found that higher switching significantly correlated with higher CRS-R score ( $r=0.37$ ,  $p=0.003$ -Fig. 5.3c) showing that the amount of switching between states was related to their clinical diagnostic score.



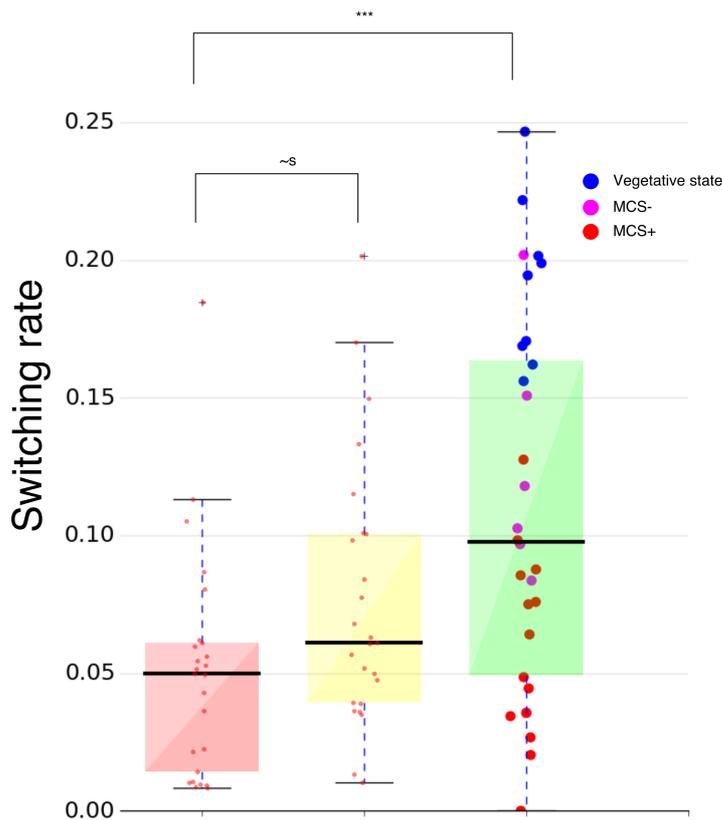
**Figure 5.3 - Brain dynamics in patients with disorders of consciousness.** (a) The Hidden Markov model results in two states, state 2 and state 9 that patients' brains occupied (fractional occupancy/FO) mostly during the data collection. (b) State 9's activity and connectivity patterns included DMN as well as cerebellar and brainstem regions. State 2's activity and connectivity patterns included thalamic, orbitofrontal and cerebellar regions. (c) Switching between states 2 and 9 was positively correlated with the CRS-R score. For both activity and connectivity spatial maps slices at  $x = 98$   $y = 41$   $z = 97$  are displayed.

### 5.3.3 Brain dynamics: from sedation to disorders of consciousness

In the previous sections, I identified certain states describing the brain dynamics during sedation and disorders of consciousness. The switching rate between them related to the level of responsiveness and clinical diagnostic score respectively. A

natural extension would be to compare the states and their switching rates between the two experiments in order to arrive at a conclusion about brain dynamics in different states of consciousness. However, due to the fact these states were derived on the premises of each experiment separately, it was difficult to make any cross-comparisons. Thus the next goal was to obtain a common underlying marker of temporal dynamics like complexity reported in chapter 4. To do so, I pooled the data from the two experiments and I applied the HMM model. This resulted in 12 states common to the new dataset. Because, in this case, there were no states specifically associated with each cohort, I used the switching rate between all states obtained from the model as a proxy of temporal complexity. The switching rate between all states obtained from the model is regarded as a measure of temporal instability (Vidaurre et al., 2017b). From Markov chain theory, the higher the switching rate the more time it takes to visit different states as transitions would be more likely diverged between different states (Ross, 1966). In such a configuration, the chain of transitions is more likely to take place between a restricted number of states and not span the entire set of states, as for some states it would take an increased amount of time to visit. Conceptualizing the transitions as a graph (with nodes being states and edges being state transitions), increased switching rate would be analogous to a less temporally complex brain. Therefore, I considered changes in the switching rate between all states to be a proxy measure of temporal complexity.

In that regard, I quantified the effect of level of consciousness on the switching rate between all states obtained from the common model ( $F(2,77)=8.67$   $p<0.001$ , bayes factor=4.3528, eta squared=0.2525)(Fig. 5.4). I observed that the switching rate in the disorders of consciousness cohort was significantly higher compared to the awake state ( $p<0.001$ ) and approached significance when compared to moderate sedation ( $p=0.056$ ). In light of the previous remarks regarding the time to communicate information between states in a highly unstable/noisy brain, I concluded that temporal complexity decreased from sedation to disorders of consciousness.



**Figure 5.4 - Brain dynamics from sedation to disorders of consciousness.** A number of states common to both cohorts were derived. High switching rate implied more instability and less temporal complexity. I observed that patients with disorders of consciousness exhibited significantly higher switching rate when compared to the awake and approached significantly higher switching rate when compared to moderate sedation. We also show switching rates across different categories of disorders of consciousness (vegetative state, MCS+, MCS-). Boxplots' thick lines show median values and whiskers represent 1.5 times the inter-fourth range. Triple asterisks (\*\*\*) imply *post hoc*  $p < 0.001$  significance and ~s imply approaching significance  $p < 0.05$  Bonferroni corrected. *Abbreviations:* DOC: Disorders of Consciousness.

## 5.4 Discussion

The brain is not static even when consciousness is abolished. Previous studies have shown that anaesthetic states show high temporal variability in functional connectivity that has been related to transitions and recovery from unconsciousness (Hudson et al., 2014). Similar results have been shown for patients with disorders of consciousness (Di Perri et al., 2018). In light of these, in this chapter I used the HMM

model to investigate brain dynamics and their complexity in different cohorts of propofol-induced sedation and disorders of consciousness.

First, I applied the HMM model to the data from the sedation experiment. I found that during sedation brain dynamics were characterized by absence of visits to state 6 that was otherwise often visited during the awake, mild sedation and recovery conditions. State 6's activity pattern included DMN regions and thalamic regions showing that, during sedation, these regions were comparatively less activated. Deactivation of the DMN has been observed in previous paradigms during propofol-induced anaesthesia (Greicius et al., 2008). Propofol-induced anaesthesia has also been shown to deactivate parts of the thalamus (Alkire et al., 2008; Mashour et al., 2014). From a connectivity perspective, state 6 included DMN regions such as PCC and the precuneus. It is thus possible that reduced occupancy of this state might be associated with decreased DMN connectivity. Impaired DMN connectivity has been previously shown in cases of propofol-induced sedation and anaesthesia potentially marking its reduced role in integrating information from other networks (Boveroux et al., 2010; Stamatakis et al., 2010).

While state 6 was dominant before and after sedation, state 4 had high FO when individuals were sedated. This state included activation of regions such as the precentral and prefrontal regions that can be broadly characterized as regions belonging to the FP, "executive" network. Previous work has emphasized the importance of switching between DMN and executive networks as a feature of conscious processing (Guldenmund et al., 2013; Di Perri et al., 2016). Indeed, it has been observed that during anaesthesia and in disorders of consciousness, such networks are pathologically connected (Vincent et al., 2007; Boly et al., 2009). Thus I hypothesised that toggling between the "DMN-like" state 6 and state 4 would suggest more responsive individuals during sedation. In that regard, I correlated the switching rate between these two states and I showed that it was negatively correlated with reaction time during sedation. In other words, the more switching between the two states, the more responsive the participants were. It is worth noting that both of these states included brainstem regions that previous literature has suggested as common

target for the arousal pathways; however further evidence is needed for clarifying why these states include brainstem regions (Brown et al., 2010).

Next, I investigated brain dynamics in patients with disorders of consciousness. I found a state that showed the highest FO among all patients during the scanning (state 9). This state's activity pattern included cerebellar and inferior frontal regions as well as parts of the DMN network such as the PCC. These regions show high overlap with regions previously reported as part of the "DMN's positive connectivity" regions in patients with disorders of consciousness (Di Perri et al., 2016). Why can high activity in these regions be a trait of an unconscious brain? Increased positive DMN connectivity has been observed in anaesthesia and disorders of consciousness and is thought to signify its inability to communicate information to the rest of the brain (Boveroux et al., 2010; Di Perri et al., 2016). This might explain why brain dynamics in some patients would occupy a state with spatial features similar to the ones obtained for DMN's positive connectivity. State 2 also showed high FO among patients. Its activity pattern included dorsolateral–prefrontal, and orbitofrontal cortical regions (regions belonging to the FP, "executive" network) and occipital regions. These regions show high overlap with "DMN's negative connectivity" regions reported in the Di Perri and colleagues work (Di Perri et al., 2016). Anti-correlations or negative connectivity between the DMN and the rest of the brain have been deemed an important feature of conscious cognition through which the DMN integrates information from multiple networks (Dixon et al., 2017). For example, such anti-correlations have been observed in patients that had emerged from a minimally conscious state (Di Perri et al., 2016). Taken together, I conclude that these two reflect opposite states of brain dynamics potentially regulated by how the DMN connects to the rest of the brain.

In turn, based on the previous remarks about the functional role of positive vs. negative DMN connectivity, I hypothesized that transitions between state 2 (related to negative correlations of the DMN) and state 9 (related to positive correlations of the DMN) would relate to loss of consciousness of each patient. In that regard, I found that the switching rate between states 2 and 9 positively correlated with the CRS-R score. Therefore, it might be possible that increased switching from a "DMN positive

connectivity”-related state to a “DMN negative connectivity”-related one can reflect an increased state of consciousness (Di Perri et al., 2018). It is worth noting that the spatial patterns of functional connectivity of these states included mostly cerebellar and brainstem areas. These findings might confirm evidence from previous studies showing increased connectivity between DMN and brainstem and cerebellar regions during loss of consciousness (Stamatakis et al., 2010; Guldenmund et al., 2013).

It is worth noting that there were spatial similarities between states identified in the sedation and DOC cohorts. An intriguing hypothesis for this is that the HMM could be capturing a common core of abnormal dynamic connectivity that encompasses regions of the DMN. Thus it is possible that the extent of alterations in dynamic functional connectivity of the DMN might constitute a common marker for consciousness impairment extending previous reports that have shown such an association from a static functional connectivity perspective (Vanhaudenhuyse et al. 2010). Despite this common state framework between sedation and DOC, there are specific differences; a striking example is the inclusion of the thalamus in DOC state 2. For the DOC it could be the case that abnormal dynamic functional connectivity in this DMN core is also accompanied by further widespread cortico-thalamic structural alterations. Some evidence towards this direction comes from previous studies suggesting that altered DMN functional connectivity in the DOC is underpinned by structural connectivity abnormalities. For example, alterations in the structural connectivity between the PCC and the thalamus have been previously reported in a DOC cohort (Fernández-Espejo et al., 2012). Thus abnormal DOC dynamics could be mediated by abnormalities in both functional and structural DMN connectivity.

As a further research effort for identifying this common core, I also plan to look at how the temporal dynamics of each of these states differ across clinical categories (for example by looking at how fractional occupancy and switching rates of these states differ between the sedation and DOC cohorts).

Collectively, I observed that brain dynamics would be dominated by different states in the two cohorts. In both cohorts, the switching rate between relevant states correlated with the level of responsiveness (in the case of sedation) or CRS-R score (in the case of

disorders of consciousness). Because of the different definitions of each state obtained for each dataset, it is inherently difficult to quantify and compare the states and their switching rates. In order to overcome this obstacle, I applied the HMM model to concatenated data from these datasets to obtain a common underlying model. I then considered the average switching rate between all the states as a measure of temporal complexity that allowed me to compare the temporal dynamics between all subjects. I observed that the switching rate was increased when going from awake to moderate sedation and disorders of consciousness signifying a more unstable/noisy brain. From an information perspective, higher switching rate implies that information can reach a restricted number of states after a certain time while ignoring others (Ross, 1966). This finding is also in line with previous studies suggesting that the unconscious brain is associated with the emergence of self-reinforcing connectivity between a limited number of networks, a phenomenon that might disrupt whole-brain communication (Demetrio et al., 2013; Di Perri et al., 2013). In addition, analogous to the (spatial) network complexity, a brain with high switching rate would be visiting a restricted number of states while ignoring other states i.e. it is less temporally complex.

In light of these findings, an important question is whether there is a direct relationship of temporal complexity to spatial complexity discussed in chapter 5. A state obtained from the HMM model includes temporally coherent regions thus it is related to some form of spatial complexity across its regions. Therefore, it is possible that alterations in the temporal complexity between these states show relationship with alterations in their spatial complexity. Towards this direction, preliminary evidence exists from the EEG literature where studies have shown the brain dynamics switch swiftly between different EEG “microstates” which, at the same time, change their spatial distribution across the scalp (Van de Ville et al., 2010). Thus the reduced (spatial) complexity observed in chapter 4 might be also complemented by reduced temporal complexity although further evidence is required. Because the states obtained here were derived on group/concatenated data at each experiment, further analysis is required to obtain subject-specific states and associate their spatial complexity with the changes in temporal complexity.

---

## 5.5 Conclusion

Collectively, brain dynamics during sedation and disorders of consciousness were characterized by variety of states and inter-state transitions not previously seen by the static approach. Impaired consciousness showed increased switching rate and decreased temporal complexity suggesting that the unconscious brain becomes increasingly unstable. The additional temporal information yielded by this approach can provide new insights on the connectivity patterns sustaining sedation and disorders of consciousness. Future research can use the spatiotemporal complexity markers provided by this and the previous chapter, for example, in a classifier for diagnostic and prognostic purposes.

## **Chapter 6: Complexity of functional connectivity in the developing brain**

As I have discussed in the previous chapters, alterations in the spatial and temporal complexity of functional connectivity in the adult brain could explain anaesthetic-induced loss of consciousness and cases of disorders of consciousness. The last question of this thesis is whether these results can be transposed to the other end of the developmental trajectory i.e. the developing brain. Infants display features characteristic of what may be referred to as minimal consciousness; for example, avoiding painful stimuli or eye contact with their mother (Lagercrantz and Changeux, 2009). The state of consciousness in the developing brain is considered different than the adult one, as the infant brain is at a transitional state and its structural and functional connectivity have not been fully established (Nowakowski, 2006).

Together, these lead to the question of whether reduction in complexity is a marker for loss of consciousness only in the adult brain or it can be observed in the developing brain as well.

Functional connectivity networks in infants and children cannot be simply considered as precursors of adult brain networks as the developmental trajectory modulates their properties by a series of biological and experiential events (Power et al., 2011). Functional connectivity networks are less integrated in the first stages of development showing their lack of functional cross-communication (Fair et al., 2007b; Vértes and Bullmore, 2015). Differences between the functional connectivity of adults and infants render the translation of results obtained from anaesthetized adult brains to the developing brain an open question.

These observations necessitate a separate study of functional connectivity networks in the developing brain. fMRI studies looking at the functional connectivity in anaesthetised infants have been sparse mainly due to difficulties in image acquisition as well ethical implications of potential exposure of infants to any additional risks except those associated with medical conditions and associated treatment such as surgery (Mongerson et al., 2017). Bedside EEG is a useful, non-invasive tool for

tracking brain activity during infant anaesthesia. EEG captures oscillatory activity of pyramidal neuronal populations as signal of high temporal resolution from different sensors placed around the scalp (Cohen, 2017). EEG connectivity can be defined on the basis of coherent oscillatory activity. Towards this direction, I analysed a unique EEG dataset coming from infants spanning an age range of 0-3 years that were anaesthetized for clinical purposes (Cornelissen et al., 2015). This dataset allowed me to explore two research questions: First, the effects of loss of consciousness in the functional connectivity networks of the infant brain (anaesthetic trajectory). Second, the changes in the functional connectivity networks with increasing age within each anaesthetic state (developmental trajectory).

With respect to the anaesthetic trajectory, studies investigating the effects of anaesthesia in the developing brain have been limited by usually focusing on frontal connectivity. My plan is to expand this analysis to whole-brain EEG networks and, as in chapter 4 to show that EEG networks have reduced complexity during anaesthesia.

Regarding the developmental trajectory, fMRI studies have shown that the development of functional connectivity is characterized by changes in the integration of functionally specialized networks with the development of long-range connections (Fair et al., 2007b). In that regard, my plan is to use the previously proposed complexity measure to show that, within each anaesthetic state, brains become increasingly complex with increasing age.

Taken together these two research directions would not only suggest that complexity of brain connectivity in infants during anaesthesia is reduced, but also that these reductions are age-specific and related to each infant's brain functional development.

## **6.1 Introduction**

Part of our knowledge for the developing brain comes from post mortem histological or gene expression analysis suggesting that the basic structural and functional framework is rapidly reorganizing until the third year of life (Petanjek et al., 2011; Pletikos et al., 2014). Advances in imaging methods have resulted in an improvement

of *in vivo* quantification of brain functional connectivity networks during development. Such networks have been identified even before birth (van den Heuvel and Thomason 2016). Since their emergence, these networks are constantly maturing in terms of their functional connectivity in order to attain their functional specialization (Smyser et al., 2010). Studies looking at their trajectory in utero and after birth have shown that functional connectivity networks become more segregated as a result of their increased functional specialization and, at the same time, more integrated usually by recruiting long-range connections (Fair et al., 2007b; Thomason et al., 2015). Thalamocortical connectivity has also been shown to develop since infancy. At earlier stages of life, thalamic connectivity is mostly restricted to primary sensory systems whereas later at the first year of life, it includes higher-order and default mode systems (Alcauter et al., 2014).

How can functional connectivity in infants be studied? Initial attempts have used an fMRI experimental setting similar to adult fMRI (Fransson et al., 2007). However, recently it has been suggested that the neurovascular coupling is not mature in infants, thus BOLD based approaches might be misleading (Kozberg and Hillman, 2016). Alternative methods based on near-infrared spectroscopy (NIRS) have also been used to investigate differences in the organization of cortical networks in infants (Homae et al., 2010). Yet, NIRS requires light emission that penetrates through the superficial tissue (scalp and skull). In adults, this results in a restricted depth specificity typically reaching 5 to 10 millimetres beneath the inner surface of the skull. Thus NIRS is usually limited to very young infants where the superficial tissue development is still immature (Ferrari et al., 2014). Magnetoencephalography (MEG) is another approach that provides high temporal resolution as well as large coverage of the brain (Baillet, 2017). However, MEG devices are not optimized for infant head shapes, thus making their use in that age range impractical most of the times.

An alternative method of measuring brain activity is EEG, which quantifies the phase and amplitude of oscillatory activity of local neuronal assemblies in the human cortex by measuring membrane potential fluctuations of cortical pyramidal cells perpendicular to the skull (Cohen, 2017). Following historical conventions, oscillatory activity has been categorized in the following frequency bands of interest: delta (1-

4Hz), theta (4-8Hz), alpha (8-12Hz), beta (12-30Hz), low gamma (30-60Hz), high gamma (60-250Hz), and fast ripples (200-400Hz). Each rhythm is assumed to serve different brain functions (Buzsáki and Draguhn, 2004). For example, alpha oscillations are the most prominent rhythms to have been associated with higher-order communication between occipital, parietal, and central areas (Hipp et al., 2012) while theta frequency has been associated with memory retrieval by communicating information between the medial temporal lobe and the rest of the brain (Hipp et al., 2012). Importantly, coherence or connectivity between neuronal oscillations at different frequencies plays an important role in coordinating between-region communication (Fries, 2005).

What is the importance of these oscillations and their coherence in loss of consciousness? Evidence from power analysis has shown that loss of consciousness in adult anaesthesia is associated with increasing *alpha* oscillatory power (Purdon et al., 2013). What is more, these alpha oscillations are primarily localized in the frontal cortex and appear to be coherent, suggesting that alpha-based connectivity might prove more useful for understanding loss of consciousness (Brown et al., 2010). Models have now proposed that the emergence of alpha coherence in the frontal cortex is associated with the disruption of the thalamocortical loop and its impact in whole-brain communication (Llinás et al., 1998; Ching et al., 2010).

Evidence for this comes from studies using whole-brain EEG connectivity. These studies have shown changes in EEG alpha networks when transitioning to or recovering from loss of consciousness. Specifically, alpha EEG networks during anaesthetic-induced loss of consciousness become more fragmented and less efficient (Boly et al., 2012; Lee et al., 2013; Chennu et al., 2016). This has been linked to aforementioned theories regarding the shift in the balance of segregation and integration across the brain. Changes in alpha connectivity might produce a more segregated network as a result of impaired information integration across the whole-brain (Lee et al., 2017). Therefore, by using the complexity as an aggregate measure of segregation and integration, one can conclude that complexity of alpha networks might provide a marker for loss of consciousness.

Can similar conclusions be drawn for the infant brain? It is believed that infants reach a state of consciousness that is only basic compared to adults mainly because thalamocortical, higher-order, and DMN network connectivity has not fully developed (Lee et al., 2005; Nowakowski, 2006). However, if infants still exhibit some form of consciousness then a similar, to some extent, loss of balance of segregation and integration might take place in the infant brain during loss of consciousness (Merker, 2007; Lagercrantz and Changeux, 2009). Under this framework, decrease in the complexity of functional connectivity networks might discriminate between an unconscious and conscious infant brain even when functional connectivity still undergoes development.

To address this hypothesis, I conducted a connectivity analysis on novel EEG data in infants aged from 0 to 3 years old obtained from sevoflurane-induced anaesthesia during standard clinical procedure. Multi-electrode EEG caps were used to monitor brain activity during maintenance of surgical anaesthesia (MOSSA) and emergence from anaesthesia. To investigate whether complexity changed during loss of consciousness, I analysed EEG data from selected segments of two steady anaesthetic states with different levels of anaesthetic depth (MOSSA corresponding to moderate anaesthesia and MOSSA2 corresponding to deep anaesthesia) as well as emergence from anaesthesia. In turn, I constructed functional connectivity networks representing sensor activities focusing on alpha frequency for the reasons detailed previously. Based on the results of chapter 4, I hypothesized that the complexity of networks would decrease in anaesthesia compared to emergence.

The uniqueness of this dataset also allowed me to understand whether, within each anaesthetic state, the complexity of functional connectivity changed with age. Previous evidence coming from fMRI studies has shown that functional connectivity becomes more complex in network organization from 7-30 years in order to support higher-order functions (Fair et al., 2007b; Vertes and Bullmore, 2015). Expanding on this finding, my goal was to show that complexity increases since infancy reflecting a shift towards a more integrated configuration.

Taken together, I hypothesized that anaesthesia would not only cause an overall reduction in complexity, but also that this reduction would be age-specific reflecting different stages of brain development and function.

One limitation of EEG-derived connectivity is that it uses sensor data with limited coverage around the scalp. This can lead to ambiguities regarding the interpretation for how sensor-based connectivity is related to cortical connectivity. Source reconstruction techniques can deal with this issue by utilizing mathematical methods for deriving signals at the source level using the measured signals at the sensor level (Van Veen et al., 1997; Kybic et al., 2005). Thus functional connectivity in EEG data can be translated to cortical functional connectivity resembling, to some extent, functional connectivity derived from fMRI.

In that regard, I complemented the analysis at sensor level with analysis at the source level in order to demonstrate that alterations in complexity hold at both the sensor and source levels.

## **6.2 Materials and Methods**

### **6.2.1 Participants**

Part of the EEG data has been used previously reported in a study by Cornelissen and colleagues (Cornelissen et al., 2015). Multi-channel EEG data was collected from infants 0-3 years old who were scheduled for an elective surgical procedure at Boston Children's Hospital from December 2011 to August 2014. Eligibility criteria consisted of infants between 0 and 3 years postnatal age who required surgery below the neck and were recruited from the pre-operative clinic. Ethical approval was obtained from Boston Children's Hospital Institutional Review Board (Protocol Number IRB-P000003544) and classified as a 'no more than Minimal Risk' study. Informed written consent was obtained from parents/legal guardians before each study. No infants were prescribed midazolam or other premedication on the day of surgery. Data was obtained for two levels of surgical anaesthesia at different anaesthetic depths (MOSSA and MOSSA2) and during emergence (see section 6.2.2). After excluding missing data

and following strict artefact removal (see section 6.2.4 for details), the sample size was restricted to 39 infants for emergence (mean age 11.4, stdev +-8.8), 51 infants for MOSSA (mean age 11.4, stdev +-10.4), and 34 infants for MOSSA2 (mean age 10.5, stdev +-7.3). There were no statistically significant differences between the ages of these groups. To explore the developmental trajectory, participants were further divided into 5 age classes. The motivation from this stratification was as follows. First, age grouping was based on results of previous work studying infants of similar age ranges during surgical anaesthesia (Cornelissen et al., 2018). Second, age grouping was based on key stages of postnatal neurodevelopment (Kanold and Luhmann, 2010). Specifically, data was divided as following: i)  $\leq 3$  months ii) 4-6 months iii) 7-9 months iv) 10-14 months v)  $> 14$  months. The numbers of each group (n) for each condition were as follows: for emergence: i age group: (5) ii second age group: (9) iii age group: (7) iv age group: (9) v age group: (9) for MOSSA i age group: (8) ii age group: (9) iii age group: (12) iv age group: (9) v age group: (13) for MOSSA2 i age group: (3) ii age group: (5) iii age group: (10) iv age group: (7) v age group: (9).

## 6.2.2 Experimental design and procedure

Infants were anaesthetized with sevoflurane (SEVO)-induced anaesthesia to maintain a state of surgical anaesthesia (MOSSA). Nitrous oxide was added only if necessary (based on the anaesthesiologist's judgement) and was discontinued after intubation. EEG was recorded continuously before the surgery and after the infant was emerging from anaesthesia. From the recording, I chose segments corresponding to three anaesthetic states identified by steady states of sevoflurane concentration 1) Maintenance of state of surgical anaesthesia (MOSSA) when infants were anaesthetised with mean sevoflurane concentration 1.8-2.5% and plus minus nitrous oxide 2) Maintenance of state of surgical deep anaesthesia (MOSSA2) where infants were anaesthetised with mean sevoflurane concentration 2.6-3.3% plus minus nitrous oxide 3) emergence from anaesthesia with mean sevoflurane concentration 0-0.3%. Segments for MOSSA and MOSSA2 were chosen at least two minutes after incision time whereas segments for emergence were chosen at least two minutes after body movement for emergence. Mean and standard deviation SEVO concentration for the

MOSSA2 segments was 2.8 +/- 0.2%, for the MOSSA segments it was 2.2 +/- 0.2%, and for the recovery segments it was 0.1 +/- 0.3%.

An EEG cap was used to record EEG activity (WaveGuard EEG cap, Advanced NeuroTechnology, Enschede, Netherlands). There were two types of EEG caps used for this study: 33 recording electrodes were positioned according to the modified international 10/20 electrode placement system at Fz, FPz, FP1, FP2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Cz, CPz, C3, C4, CP1, CP2, CP5, CP6, Pz, P3, P4, P7, P8, T7, T8, M1, M2, POz, Oz, O1, and O2. Reference and ground electrodes were located at Fz and AFz, respectively. This cap was used for up to 6 months infants. For >6 months a 40 channel cap with electrodes Fz, FPz, FP1, FP2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Cz, CPz, C3, C4, CP1, CP2, CP5, CP6, Pz, P3, P4, P7, P8, T7, T8, M1, M2, POz, Oz, O1, and O2 and AF7, AF8, FT7, FT8, TP7, TP8, PO7 and PO8 was used. The impedance of the electrode-skin interface was kept to a minimum by massaging the skin with an EEG prepping gel (Nu-Prep gel, DO Weaver & Co., CO, USA), and conductive EEG gel was used to optimize contact with the electrodes (Onestep-Clear gel, H+H Medical Devices, Dulmen, Germany). EEG activity from 0 to 500Hz was recorded with an Xltek EEG recording system (EMU40EX, Natus Medical Inc., Ontario, Canada). Signals were digitized at a sampling rate of 1024Hz (or 256Hz in one case), and a resolution of 16 bits.

### **6.2.3 Clinical data collection**

Demographics and clinical information, including age, gender, surgical procedure, anaesthetic management, incision and body movement times were collected from the electronic medical records and from the in-house Anaesthesia Information Management System (AIMS). Sevoflurane, oxygen, and nitrous oxide concentrations were downloaded from the anaesthetic monitoring device (Drager Apollo, Draeger Medical Inc., Telford, PA) to a recording computer in real-time using ixTrend software (ixellence, Wildau, Germany). Signals were recorded at a sampling rate of 1 data point per second. Gross body movement was recorded with a camcorder that was time-locked to the EEG recording (Xltek DSP270x, Natus Medical Inc.).

#### 6.2.4 EEG data acquisition and preprocessing

EEG data was obtained for 100 seconds at two stages of MOSSA corresponding to different anaesthetic depths and one stage of emergence from MOSSA. The collected data underwent various preprocessing stages. First, I visually inspected the raw segments for any presence of strong artefacts. If that was the case, infants were not considered for further analysis. Then EEG signals were downsampled to 250Hz and bandpass filtered within the 1-50Hz range. The low edge of this range was defined for removing ultra-slow frequencies that can be contaminated by sweating etc. The upper edge was used in order to remove high frequency artefacts. EEG data was cleaned using the *clean\_rawdata* function implemented in the *EEGLAB* (Delorme and Makeig, 2004). Channels that were removed due to excessive artefact were interpolated using spherical interpolation implemented in *EEGLAB*. Data was re-montaged to a nearest-neighbour Laplacian reference using distances along the scalp surface in order to weight neighbouring electrode contributions. ICA was not required due to the small size of the data. After segmenting the continuous segment into 1-second epochs, I used the function *pop\_rejkurt* as implemented in *EEGLAB* (Delorme and Makeig, 2004) that rejects epochs based on the kurtosis of the data. I allowed no more than 10% of epochs to be rejected.

#### 6.2.5 EEG network construction and complexity

Decomposition of the time series in its oscillatory activity in different frequency bands (spectrum) was calculated using multi-taper analysis (Mitra and Bokil, 2008). Multi-taper analysis reduces the bias in obtaining the true underlying oscillatory activity caused by standard Fourier techniques (Mitra and Bokil, 2008). Multi-taper analysis uses tapers of EEG data and calculates the spectrum within each taper separately by using specific spectral functions (Slepian functions). For a taper of specific time length  $T$  and for a frequency band of interest  $W$  (tile of frequency and time), the time-bandwidth product  $T \times W$  corresponds to how many such functions will be used in this particular tile of frequency and time. I calculated the spectrum with  $T \times W = 3, K = 5$  tapers. These parameters were consistent with the previous studies published

using part of this data (Cornelissen et al., 2015) and have been shown to work better in terms of reducing bias in the calculation of the frequency decomposition (Purdon et al., 2013). I used the *chronux* toolbox to perform the multi-taper analysis (Mitra and Bokil, 2008).

Cross-spectral coherence is a complex number reflecting the relation between signals at each frequency and was calculated by correlating the multi-taper spectrums for all pairs of sensors  $i$  and  $j$ . For the sake of this study, I focused only on the alpha frequency band 8-12Hz (alpha) due to previous work arguing that patterns of alpha connectivity could constitute neural markers of impaired consciousness under anaesthesia (Purdon et al., 2013; Chennu et al., 2016).

One issue with using cross-spectral coherence relates to volume conduction. This refers to the case when two different electrodes might give spuriously high coherence just because they are measuring the same source (Stam et al., 2007). One way to overcome this problem is by keeping only the imaginary part of the coherence (Nolte et al., 2004). Two sensors measuring the same source cannot give non-zero imaginary coherence (Nolte et al., 2004). I thus obtained a  $33 \times 33$  functional connectivity matrix for each infant where each entry  $(i, j)$  represented the imaginary coherence between the two channels  $i$  and  $j$  at alpha frequency. M1 and M2 channels were excluded from these due to poor contact with the infants' heads leaving me with networks of  $31 \times 31$  electrodes. In the case where 40-channel EEG caps were utilised, I only used the same 31 electrodes as in the 33-channel EEG cap, thus obtaining a  $31 \times 31$  electrode matrix.

In turn, alpha-specific matrices for each infant were thresholded at a range 10-50% thresholds  $\tau$  at steps of 2%. The limits were chosen so as to prevent the network from being severely fragmented and from being random by introducing connections and they are consistent with recent literature in adult EEG networks (Chennu et al., 2016). Consequent network properties were calculated for an average over this range of thresholds.

I used the BCT toolbox to obtain the degree sample for each network and I calculated the complexity as described in chapter 3. Due to the limited number of sensors/nodes and the ambiguity involved in the interpretation of rare sensor degrees, I did not

investigate rare and frequent entropy as in chapter 4 but I focused on the total entropy of the degree distribution as in chapter 3. I also applied modularity analysis to help with visual representation of EEG networks. Modularity was calculated by the heuristic Louvain algorithm (as implemented in the BCT toolbox) and all measures derived therefrom were averaged over 50 repetitions.

### **6.2.6 Source reconstruction**

As I mentioned in the methods section, one way of linking EEG activity to source activity in the cortex is to use source reconstruction methods. The two steps for source reconstruction include forward modelling and its inverse solution. Forward modelling involves using Maxwell's equations to predict the electromagnetic field produced by the sources at a given electrode. This is done by deriving the leadfield matrix, a mathematical quantity that relates the measured activity at the electrode level with the source activity. In that regard, to calculate the forward model one needs to combine information regarding i) how electric activity spreads through different tissues (the head model) ii) the position and orientation of different dipoles (the source model), and iii) the electrodes' locations. As far as the head model is concerned, one approach is to use numerical solutions such as the Boundary Element Model (BEM) (Mosher et al., 1999) where the brain is compartmentalized into 3 tissues (brain, skull, scalp) each one being covered by a tessellation. To obtain such a geometrical description one needs anatomical information from the T1 images such as to segment out the brain, scalp, and skull tissues. When individual T1 data are not available, an alternative way is to use predefined templates. Under this framework, it is important to use age-appropriate brain templates and parameters to accurately quantify the localization and time course in each infant (Ortiz-Mantilla et al., 2012). Towards this direction, I used age-specific templates provided by the Richards laboratory (Sanchez et al., 2012) available here (upon request)

*(<http://jerlab.psych.sc.edu/NeurodevelopmentalMRIDatabase>)*

Specifically, templates were obtained for 3, 4.5, 6, 7.5, 9, 12, 18, 15, 24, 30, 36 and 48 months of age. Besides T1 and T2 structural images, scalp, brain, and skull

segmentations were also obtained. The specific details for these are described in (Sanchez et al., 2012).

For each one of these templates, I used standard conductivity settings and the '*bemcp*' option in *Fieldtrip* to obtain a BEM head model using the tessellation of the three compartments (brain, skull, scalp). I used a standard number of vertices for the construction of the head model as suggested by *Fieldtrip* (3000, 2000, and 1000 respectively). Each infant was matched to each head model as follows: If the age of each infant was  $\pm 2$  months from the age of the templates then it was used as a forward solution for that participant. Due to lack of data for  $\leq 3$  months infants, these were mapped to the 3-month template.

The next step required realignment of the electrodes' positions with the head model. This was performed manually using *Fieldtrip*'s graphical interface. Finally, for the source model, a 3-dimensional grid of dipoles with 1cm resolution was constructed. Electrodes, head model and source model were aligned and mapped to the same space. Following this, I obtained the leadfield matrix for each template.

Inverse solution refers to obtaining the source-level activity using the leadfield matrix and the acquired EEG data. One popular method for obtaining the inverse solution encompasses beamforming techniques (Van Veen et al., 1997). The basic principle of beamforming lies in obtaining a single source's activity by looking at how it contributes to the measured EEG activity compared to other sources (Van Veen et al., 1997). Using the forward solution obtained alongside the EEG data, I utilised the beamforming technique to obtain the activity of each dipole. I used the *Fieldtrip* toolbox with the option '*pcc*' to obtain the time course and spectrum of each dipole. I then calculated the *dipole*  $\times$  *dipole* connectivity matrix using the imaginary coherence as with electrode-based networks I discussed earlier in this chapter.

After obtaining the dipole matrices, I grouped the dipoles based on known cortical regions in order to assist with interpretation of the results. I used a specific parcellation obtained in the Richards templates to group dipoles to specific regions of interest. These include age-specific atlases for infants 3, 4.5, 6, 7.5, 9, 12, 18, 15, 24, 30, 36 and 48 months and are publicly available here

(<http://jerlab.psych.sc.edu/NeurodevelopmentalMRIDatabase>).

The parcellation was based on the macro-anatomical Hammers atlas previously used in adult literature (Hammers et al., 2003). Due to lack of data for infants  $\leq 3$  months, I mapped their source data to the closest age-matched atlas, i.e. the 3-month atlas.

I assigned each dipole to an ROI by overlapping the source model (the dipole positions) and the atlas images. I then averaged the connectivity values of the dipoles belonging to each ROI to obtain ROI-specific time course. This resulted in an ROI $\times$ ROI connectivity matrix for each infant.

Due to known difficulties in obtaining source signals from subcortical regions using beamforming techniques (Krishnaswamy et al., 2017), I excluded regions corresponding to the ventricles, cerebellum, brainstem, striatum, corpus callosum, hippocampus, thalamus, amygdala and insula.

Matrix thresholding was conducted as previously by looking at a range of thresholds 10-50% in steps of 2%. The complexity measure was applied in ROI networks as previously and averaged over the aforementioned range of thresholds. Due to the limited number of cortical regions, I did not investigate rare and frequent entropy (chapter 4) but I focused on the total entropy of the degree distribution as in chapter 3.

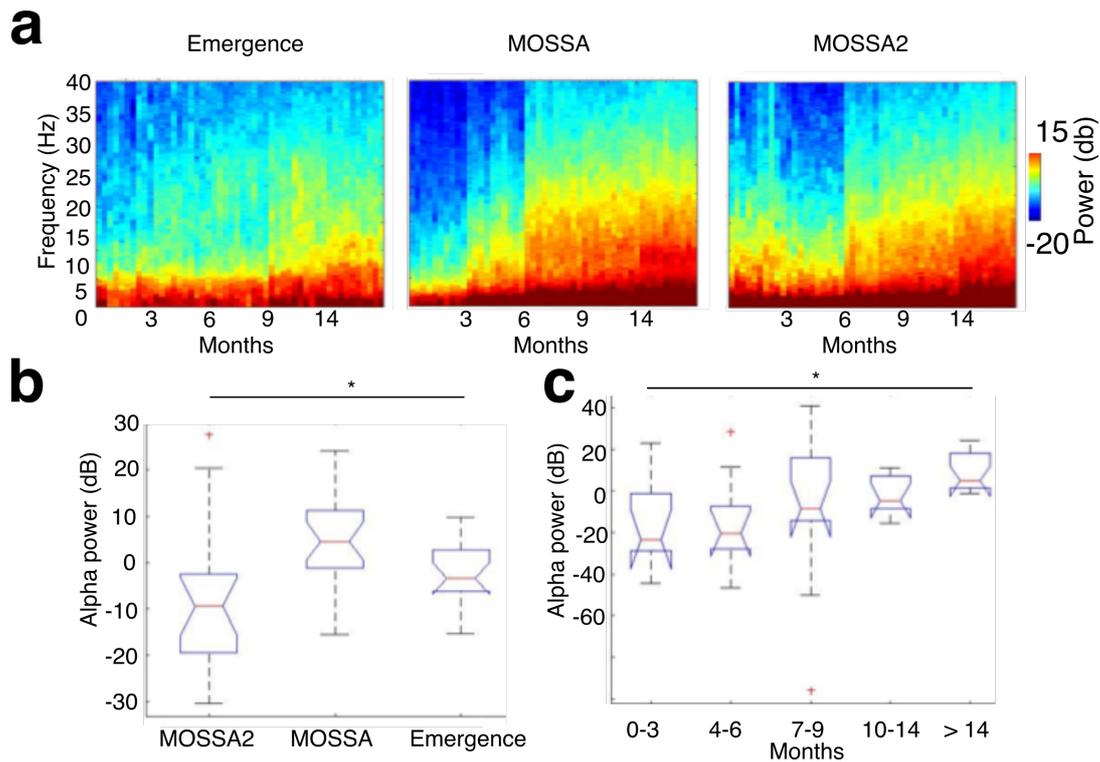
### **6.2.7 Statistical analysis**

Due to the wide age range, power and complexity data was not normally distributed. Thus, I used non-parametric ANOVA (Kruskal-Wallis) in order to test the effect of anaesthetic depth and age on power and complexity. P values were *post hoc* corrected using the Bonferroni criterion. I also reported effect sizes for ANOVA using the bayes factor and the eta squared quantities. Furthermore, I used correlation analysis to report a relationship between complexity and age. For these analyses I reported Spearman's correlation coefficient and statistical significance.

## 6.3 Results

### 6.3.1 Power during different ages and conditions

First, I analyzed data coming from power analysis. The goal was to quantify trends in alpha power as proposed by previous studies suggesting the important role of alpha in anaesthesia and loss of consciousness (Ching et al., 2010; Purdon et al., 2013). I calculated the median (over all infants) spectrograms for each condition and each age group which represent the power of each oscillation across time. First I show the spectrograms for only one frontal electrode (F8)(Fig. 6.1a) although similar trends were observed for other electrodes (Appendix Fig. 6.1). Qualitatively I observed some distinct patterns: 1) There was an increase in alpha power with increasing age for all conditions (Fig. 6.1a). One explanation for enhanced alpha power with increasing age is the maturation of thalamo-frontal connectivity (Cornelissen et al., 2015). Specifically, regarding changes in alpha power between different age groups within the same condition, I observed an effect of age on mean alpha power at MOSSA (chi square=12.28,  $p=0.0154$ , bayes factor=14.8914, eta squared=0.2349)(Fig. 6.1c). This was also the case for emergence (chi square=13.09,  $p=0.0108$ , bayes factor=16.1325, eta squared=0.2466) but not for MOSSA2. 2) I observed an increase in alpha power when comparing the anaesthetic state to emergence. Specifically, for all infant data pooled together, I observed an effect of sedation in mean (over all electrodes) alpha power (chi square=13.48,  $p=0.0012$ , bayes factor=8.3519, eta squared=0.2880) (Fig. 6.1b). *Post hoc* tests revealed an increase in mean power over all electrodes during MOSSA compared to emergence ( $p=0.0007$ ). These results replicated the observation that power increases with anaesthesia, previously observed in adults (Purdon et al., 2013). It is worth noting that during MOSSA2 power tended to decrease compared to MOSSA and emergence, potentially suggesting that alpha power might decrease during deep anaesthetic states (Hight et al., 2017).



**Figure 6.1 - Spectral power analysis for infant brain dynamics under anaesthesia and emergence from anaesthesia.** (a) Median (over all infants) spectrograms show the power in each frequency as a function of age in each of the three conditions. (b) Average alpha power over all sensors for all infants showed a significant increase with MOSSA. Panel (c) shows the dependence of alpha power to anaesthetic depth. Older infants had higher alpha power compared to younger infants. Boxplots' red lines show median values and whiskers represent 1.5 times the inter-fourth range. Single asterisk (\*) indicates  $p < 0.05$  significance of ANOVA.

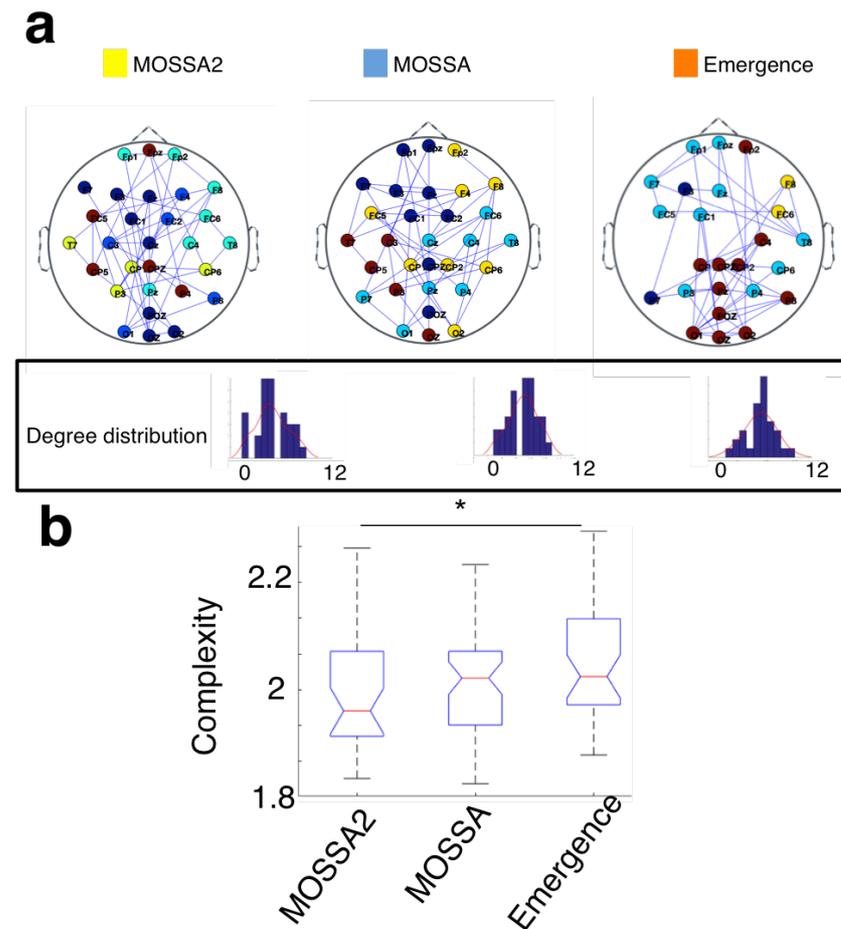
### 6.3.2 EEG network alterations in infant anaesthesia

#### *Sensor level*

The previous power analysis showed that there was an increase in alpha power at the MOSSA level. However, was this the case with complexity? According to our hypothesis, despite the observed effect of anaesthesia on power, complexity should decrease, reflecting a shift in the balance between segregation and integration towards a more segregated configuration. To show this, I investigated functional connectivity and its complexity at the electrode/sensor and source levels.

First, I examined the differences in complexity of functional connectivity networks across three anaesthetic conditions for all age groups pooled together. Average networks for all infants across the three conditions are shown in Figure 6.2a. Non-

parametric ANOVA showed a significant effect of anaesthesia on complexity (chi square=7.14,  $p=0.0281$ , bayes factor=3.2585, eta squared=0.1764)(Fig. 6.2b). Specifically, I observed a decrease in complexity during MOSSA2 compared to emergence ( $p=0.0207$ ) showing that networks were becoming homogenized and more segregated.

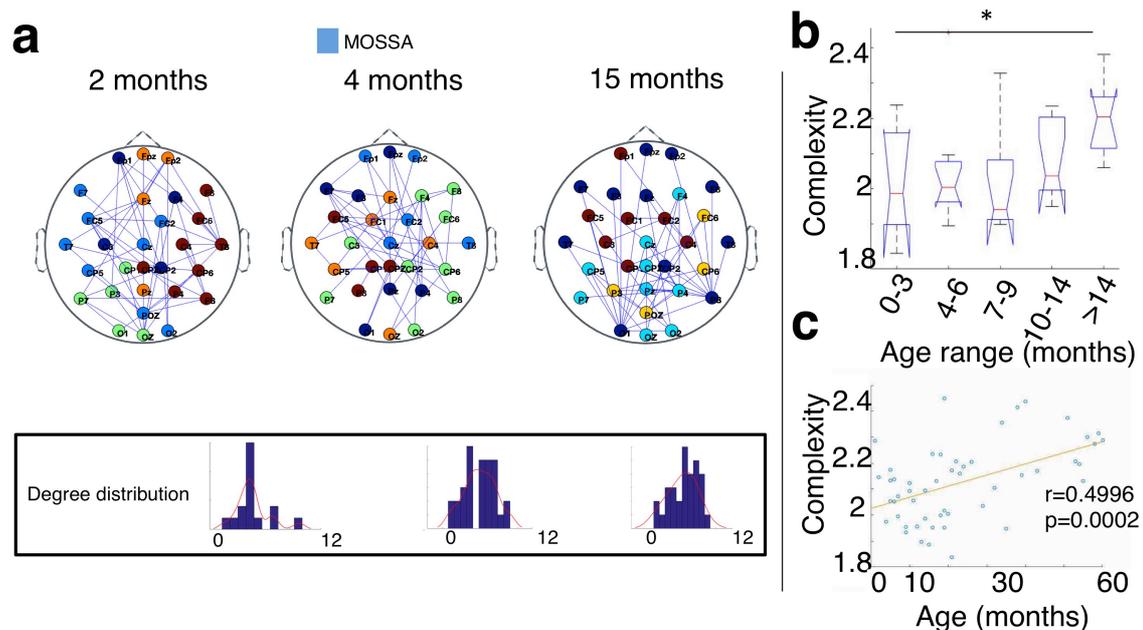


**Figure 6.2 - Complexity of alpha functional connectivity at the sensor level in the infant brain during anaesthesia and emergence from anaesthesia.** (a) Average connectivity networks at the sensor level and their respective degree distributions. Different colours for each sensor correspond to different modules in the network. (b) There was a significant effect of anaesthetic depth on complexity with a significant decrease when comparing MOSSA2 to emergence. Boxplots' red lines show median values and whiskers represent 1.5 times the inter-fourth range. Single asterisk (\*) indicates  $p < 0.05$  significance of ANOVA.

Next I quantified the effect of age on complexity within each condition. Representative EEG networks for single infants from 3 age groups at MOSSA are

shown in Fig. 6.3a. I observed a significant effect of age on complexity at MOSSA (chi square=9.46,  $p=0.04$ , bayes factor=10.5063, eta squared=0.1913)(Fig. 6.3b). Similar results were observed at emergence (chi square=19.73,  $p=0.0006$ , bayes factor=25.9662, eta squared=0.3303) and at MOSSA2 (chi square=10.39,  $p=0.0344$ , bayes factor=11.9629, eta squared=0.2062).

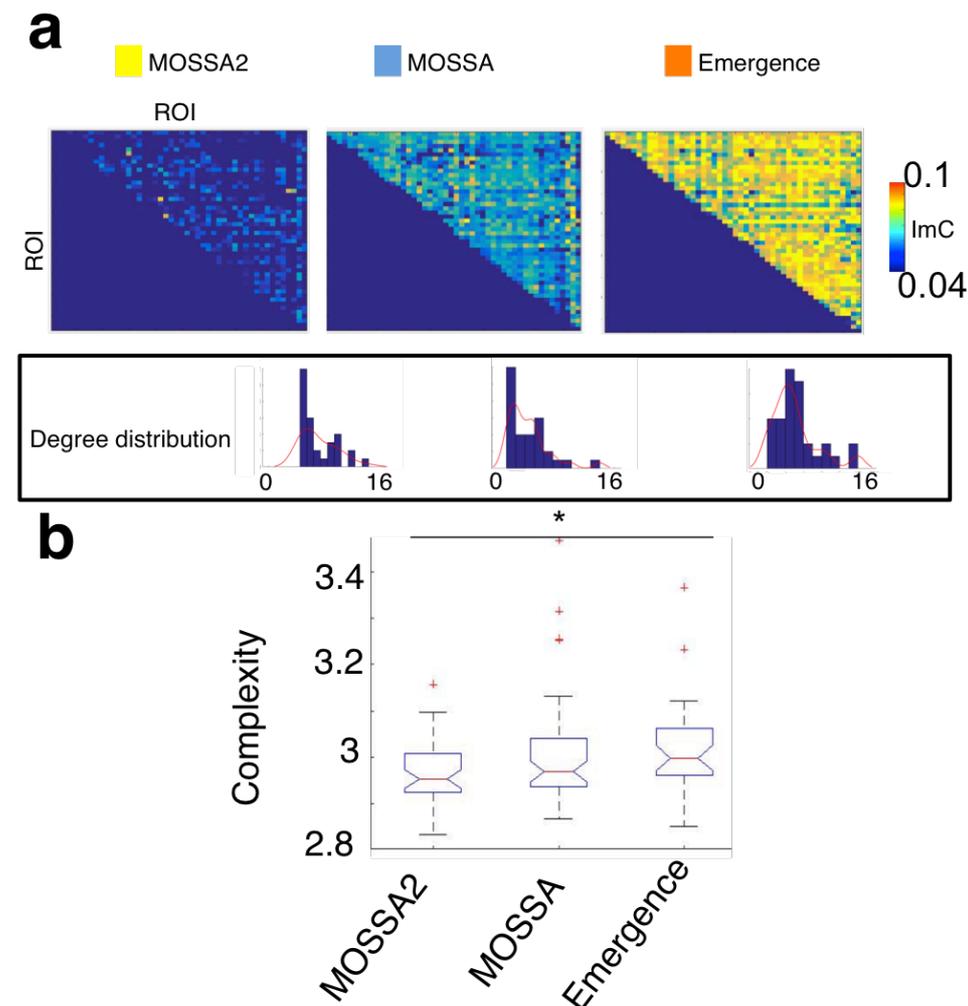
In addition, I found that complexity was positively correlated with age when infants emerged from anaesthesia ( $r=0.3527$ ,  $p=0.0277$ ). The positive correlation was also maintained during MOSSA ( $r=0.4996$ ,  $p=0.0002$ ) (Fig. 6.3c) and MOSSA2 ( $r=0.4038$ ,  $p=0.0179$ ) conditions. These results showed that, even during anaesthesia, there was an increase in complexity with increasing age potentially reflecting the tendency of the infant brain to become more integrated as age increases.



**Figure 6.3 - Complexity of alpha functional connectivity at the sensor level in the infant brain at different ages.** (a) Connectivity networks at the sensor level and their respective degree distributions for individual infants at 2 months, 4 months, and 15 months during MOSSA. Different colours for each sensor correspond to different clusters in the network. (b) There was a significant effect of age on complexity with older infants showing higher complexity than younger infants. (c) There was also a positive correlation of age with complexity of alpha connectivity during MOSSA. Boxplots' red lines show median values and whiskers represent 1.5 times the inter-fourth range. Single asterisk (\*) indicates  $p<0.05$  significance of ANOVA.  $r$  shows Spearman's correlation coefficient.

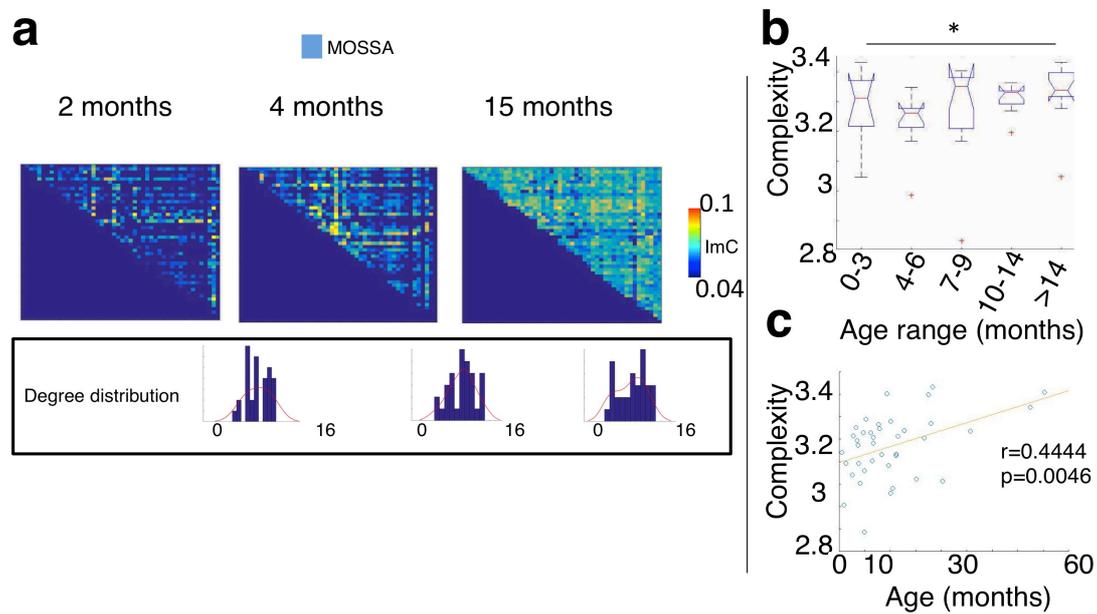
**Source level**

Next, I used source reconstruction to obtain cortical functional connectivity networks for each infant. Average matrices for all the infants across the three conditions are shown in Fig. 6.4a. First, I investigated the effect of anaesthetic depth on complexity of alpha functional connectivity in order to see whether results on alpha connectivity at the sensor level would be similar for the source level. I observed a significant effect of anaesthetic depth on complexity (chi square=6.03,  $p=0.0423$ , bayes factor=2.3415, eta squared=0.1532, Fig. 6.4b) with *post hoc* tests revealing a significant difference in complexity between emergence and MOSSA2 ( $p=0.03$ ). Similar to the results observed at the sensor level, cortical networks became less complex and more segregated with increasing anaesthetic depth.



**Figure 6.4 - Complexity of alpha functional connectivity at the source level in the infant brain during anaesthesia and emergence from anaesthesia.** (a) Average connectivity networks between pairs of cortical ROIs and their respective degree distributions. Brighter colours at the yellow orange end of the colour scale correspond to high imaginary coherence between pairs of ROIs (b) There was a significant effect of anaesthetic depth on complexity with a significant decrease when comparing MOSSA2 to emergence. Boxplots' red lines show median values and whiskers represent 1.5 times the inter-fourth range. Single asterisk (\*) indicates  $p < 0.05$  significance of ANOVA.

I then quantified the effect of age on complexity within each condition. Representative EEG networks for single infants from 3 age groups during MOSSA are shown in Fig. 6.5a. I observed a significant effect of age on complexity during MOSSA (chi square=8.02,  $p=0.043$ , bayes factor=8.2326, eta squared=0.1670)(Fig. 6.5b) and no effect of age on complexity during emergence and MOSSA2. Similar to the sensor level, I observed a positive correlation of complexity with age during emergence ( $r=0.27$ ,  $p=0.04$ ) that was maintained during MOSSA ( $r=0.4444$ ,  $p=0.0046$ )(Fig. 6.5c) and followed a trend during MOSSA2 ( $r=0.2619$ ,  $p=0.0633$ ).



**Figure 6.5 - Complexity of alpha functional connectivity at the source level in the infant brain at different ages.** (a) Alpha functional connectivity networks between cortical ROIs and their respective degree distributions for individual infants at the age of 2 months, 4 months, and 15 months during MOSSA. (b) There was a significant effect of age on complexity with older infants showing higher complexity than younger infants. (c) There was also a positive correlation of age with complexity of alpha connectivity during MOSSA. Boxplots' red lines show median values and whiskers represent 1.5

times the inter-fourth range. Single asterisk (\*) indicates  $p < 0.05$  significance of ANOVA.  $r$  shows Spearman's correlation coefficient.

#### 6.4 Discussion

In this experimental chapter I analysed a unique dataset of infants 0-3 years old undergoing surgical anaesthesia and I examined changes in complexity of functional connectivity in two anaesthetic states and emergence from anaesthesia. I observed that the complexity of functional connectivity was reduced with increasing anaesthetic depth. This finding is in line with the results observed in chapter 4 where I showed reduction in whole-brain complexity in propofol-induced sedation and anaesthesia. In addition, I observed that complexity positively correlated with age showing that functional connectivity is becoming increasingly complex as the human brain develops.

Converging results from the adult literature show that GABAergic anaesthetics (ether-based agents, barbiturates, propofol) modulate alpha oscillations (Ní Mhuirheartaigh et al., 2013; Purdon et al., 2013). The established importance of alpha oscillations has led to the idea of using alpha-based connectivity networks as marker of loss of consciousness under anaesthesia. Previous EEG studies in adult propofol-induced anaesthesia have shown alterations in a variety of network metrics. At the local/regional level, studies have shown a decrease in the degree of network hubs (Lee et al., 2010; Lee et al., 2013). At the meso level studies in propofol-induced anaesthesia show that alpha EEG networks become more fragmented during anaesthesia (Chennu et al., 2016). Finally, at the global/whole-brain level EEG networks are becoming less globally efficient (Blain-Moraes et al., 2017). Together, these results suggest that, despite increase in alpha power, alpha networks during loss of consciousness are biased towards a less global/more local configuration.

An important question is whether such brain network configuration can be translated to infant anaesthesia. Few studies have studied brain activity and connectivity in infants under anaesthesia. A previous study using part of this data focused only on frontal electrode coherence (Cornelissen et al., 2015). Expanding on these findings, I applied whole-brain EEG network modelling and showed that loss of consciousness is

characterized by reduced complexity both at the sensor and source levels. Thus, despite the increase in alpha power during MOSSA, I showed that the balance between segregation and integration was shifted with increasing anaesthetic depth with networks becoming less complex and more segregated.

Can this inconsistency between the increase in alpha power and reduction in network complexity explained? Alpha oscillations likely represent activity between the forebrain and the thalamus (Ching et al., 2010). Specifically, GABA potentiation induces coherent alpha oscillations that can explain the appearance of alpha power primarily in the frontal cortex. However, although the thalamus might be locally coherent during anaesthesia, models have suggested that information efficiency in the whole-brain is inhibited because of the inability of the rest of the brain to become synchronized with the “fast” oscillatory thalamo-frontal connectivity (Ching et al., 2010). Consistent with this observation are theories discussing loss of consciousness as a shift in the balance of segregation and integration in the adult brain. Although anaesthetics might make the thalamus to oscillate at coherent alpha frequencies (Brown et al., 2010), loss of consciousness occurs due to its impact on whole-brain integration (Alkire et al., 2008). Taken together, it is plausible that the same mechanistic explanation i.e. a shift in the balance between segregation and integration towards a more segregated configuration can explain loss of consciousness in the infant brain. However, it is worth noting that, due to differences in thalamocortical maturation between infants and adults, more investigation is required regarding the effect of thalamocortical connectivity in whole-brain complexity.

In addition to these results, I showed that the reduction in complexity was age-specific as complexity was related to age within each anaesthetic state. Specifically, I observed that alpha complexity increased monotonically with age. Previous results on the development of resting-state fMRI in infants lend credence to this idea. During development, networks become more specialized while inter-network (and often long-range) connections are developed in order to integrate different specialized modules (Gao et al., 2009). Thus the developing brain becomes increasingly complex and more integrated. To what extent this increase in functional connectivity complexity is driven by developmental phenomena? Changes in structural

---

connectivity occur rapidly since birth while myelination starts around the 3-4 months of life (Brody et al., 1987). However structural connectivity is adult-like around 9 months after birth while myelination and synaptogenesis peak around 12 months (Conel, 1963). Thus increasing complexity cannot be solely attributed to developmental changes but also reflects the brain's tendency to integrate functionally specialized modules, potentially due to increasing adaptation to environmental and cognitive demands (Lagercrantz and Changeux, 2009).

## **6.5 Conclusion**

Taken together, I observed a monotonic decrease in complexity in the infant brain with increasing anaesthetic depth. I concluded that the decrease in complexity reflects the idea of a shift in the balance between integration and segregation towards a more segregated configuration. It is therefore possible that the mechanistic explanation for losing consciousness found for adults can be transposed (to some extent) to the infant brain. Furthermore, I found that complexity increased with age, suggesting age-specific brain responses during anaesthesia, potentially due to different levels of brain network segregation and integration. Additional studies at this age range will pave the way for understanding how infants lose consciousness during anaesthesia as well as for deciphering brain function during development.

## Chapter 7: Discussion

### 7.1 Summary

This work attempted to answer the following question: How do we lose consciousness? Theoretical models propose that consciousness is a phenomenon that depends both on the segregated and integrated aspects of experience (Tononi, 2004). By the same token, neural activity should be segregated and integrated to produce consciousness (Crick and Koch, 2003) while loss of consciousness should allude to disruption of how segregation and integration are interrelated (Alkire et al., 2008).

While consciousness cannot be quantified directly, indirect measures have associated the coexistence of segregation and integration with the complexity inherent in brain function (Tononi et al., 1994). Complexity is low when the components of a system are either completely independent (segregated) or when the components are completely dependent (integrated) (Tononi et al., 1994; Tononi et al., 1998). But what is the case for the human brain? Studies using graph-theoretic models have assumed that complexity of the human brain is at a “sweet spot”, encompassing a critical balance between segregation and integration (Zamora-López et al., 2016). In that regard, chapter 3 provided evidence that the complexity of functional connectivity was critical, as it was derived on a Nash equilibrium of optimal connections supporting efficient communication at both the local and whole-brain levels.

Following this rationale, I assumed that loss of consciousness is associated with shifts from this balance. Towards this direction, in chapter 4 I showed evidence that complexity of functional connectivity was decreased during loss of consciousness. Using a dynamic connectivity approach, in chapter 5 I investigated the temporal dimension of complexity where I found that temporal complexity was also decreased in the unconscious brain. Finally, in chapter 6 I extended this framework in the developing brain where I showed that complexity was decreased in anaesthetized infants.

Overall, by collecting a variety of evidence, from the spatial and temporal complexity of the adult brain to the complexity in the infant brain, I argue that alterations in

complexity of functional connectivity might offer a potential mechanistic explanation for how we are losing consciousness.

## **7.2 Complexity of functional connectivity**

The healthy brain depends on acquiring information from different neuronal systems and integrating these into a coherent module (Fuster, 2003). Thus on the one hand, advances in neuroscience have shown the functional specialization/segregation of different regions and groups of neurons (for example the regions in the primary visual cortex, Fujita et al., 1992). On the other hand it is now clear that information from these specialized systems needs to be functionally integrated in order to foster adaptive behaviour (for example in the visual cortex information from different specialized regions are integrated to yield the feature of the visual stimuli, Kanizsa, 1979). In that regard, complexity of brain function is believed to support the coexistence of segregation and integration (Sporns and Tononi, 2001), a property that has been deemed essential for survival (Damasio, 1989).

Ideas about brain complexity, a property of brain function that encapsulates the balanced capacity of the brain to segregate and integrate information, have transformed the way neuroscientists think about consciousness and fostered theoretical and practical endeavours. These broadly tried to answer two complementary questions: First what are the properties of the brain as a system in order to support consciousness? Second how does conscious access take place in the brain? Regarding the first question, theories proposed by Tononi and colleagues suggest that consciousness emerges because of the integration of segregated subsystems in the brain (Tononi and Edelman, 1998). In that sense, in order for the brain as a system to support consciousness it should present a high level of differentiation, as well as of information integration (Tononi, 2004). Regarding the second question, theoretical models have proposed that conscious access relies upon a common data structure upon which information is shared from different specialized modules. The term proposed for this framework is the “global workspace” where information from different specialized systems is available and integrated in a

common/global space (Baars, 2002; Baars and Franklin, 2003). Expanding on this framework, Dehaene and colleagues have argued about the idea of conscious processing as a neuronal process related to how information from different specialized systems is integrated in the global workspace via long-range connections (Dehaene et al., 2006; Dehaene, 2014). Converging neuroimaging data, acquired when contrasting conscious and non-conscious processing, point towards this direction. For example, experiments studying exposure to conscious and unconscious visual stimuli suggest that unconscious words activate only the visual cortex while conscious word evoke widespread “global-workspace-like” activations including visual, parietal and frontal cortices (Dehaene and Changeux, 2011).

The aforementioned theories propose the idea that, for neuronal activity to sustain consciousness it must be functionally integrated and at the same time highly differentiated (Tononi and Edelman, 1998). In order to experimentally test these concepts, the use of formal measures has been introduced aiming at capturing the balance of segregation and integration using statistical and graph-theoretic approaches. Previously, measures of complexity investigating information exchange between neuronal populations were proposed but were usually limited to few cortical areas (Tononi et al., 1994). The emergence of networks and large-scale connectivity allowed assessing interactions between multiple regions as whole-brain networks/graphs. Based on these models, I used the entropy of the degree distribution to measure complexity in functional connectivity. The degree distribution quantifies the co-existence of differentially connected regions across the human brain (Bullmore and Sporns, 2012, Vértes et al., 2012). Shifts in the entropy of the degree distribution result in brain networks with differential organization. For example, decreased entropy is associated with increased segregation and decreased integration, as information in the network is mostly communicated locally rather than globally (Power et al., 2010). Opposite trends are observed when networks are biased towards a more integrated configuration thus supporting global rather than local communication (Power et al., 2010). Above and beyond these extremes, the entropy of the degree distribution in functional connectivity networks show a specific “heavy-tailed” pattern, believed to reflect a critical balance between segregation and

integration to support both local and global efficient communication (Vértes et al., 2012; Zamora-López et al., 2016).

How does one find empirical evidence for this balance? Criticality is a useful framework for studying how properties in a system arise at the point of balance between opposite tendencies (Christensen and Moloney, 2005). To study these, one needs to look at the collective behaviour of the systems' elements; i.e. how each element's specific tendencies are employed in order to produce complex patterns. When applied to brains, models suggest that brain complexity can be predicted using the dynamics of brain regions and how these are entangled upon the premises of their structural connectivity (Deco and Jirsa, 2012). Expanding on these ideas, theories now propose that critical dynamics pertain to the way each brain region communicates information, and how these communication paths are efficiently interweaved to support functional complexity (Avena-Koenigsberger et al., 2017).

In light of this, in chapter 3 I assessed criticality of functional connectivity by showing that it emerges from the Nash equilibrium of connections supporting communication transfer in the most optimal way (optimal connections). These connections were differentially distributed across the cortex serving different functional roles. LSNs such as the DMN, that have been deemed important for integrating information, had a high number of inter-network optimal connections while more segregated networks, such as the SM network, had a high number of intra-network optimal connections. Thus I claimed that functional connectivity in the healthy brain lies in a “sweet spot” of complexity as it can be derived upon the collective properties of optimal connections, reflecting the critical balance between segregation and integration.

### **7.3 Alterations of complexity in functional connectivity**

The notion of complexity as a critical balance between segregation and integration suggests that even few perturbations can cause alterations in complexity. For example, perturbation of long-range connections under anaesthetic induction is sufficient to produce a disconnected brain and reduce its complexity (Alkire et al., 2008). Towards

this direction, initial evidence using computer simulations have shown that transitions to a state of high anaesthesia is consistent with a breakdown of connections responsible of network integration (Steyn-Ross et al., 2001). In light of this, in chapter 4 I used complexity in order to characterize states of anaesthetic-induced unconsciousness and disorders of consciousness, under the assumption that the unconscious brain would shift towards a less complex and more segregated configuration. I refined the quantification of complexity by considering regions with different degrees in order to capture where the reduction in complexity would take place. I found that loss of consciousness during sedation and anaesthesia was characterized by decreased complexity with alterations more prominent in the complexity of sparsely connected regions. I showed that these regions were important for network integration by looking at their impact on the largest connected component, and proposed that their alterations in complexity resulted in a segregated brain network. In disorders of consciousness this phenomenon was even more prevalent suggesting that decreasing states of consciousness might be associated with increasing segregation.

Interestingly, these connections were primarily found in the DMN suggesting its enhanced role in network integration. This result aligns with the findings of chapter 3, where I showed that the DMN was optimally connected with other networks, potentially supporting inter-network communication and whole-brain integration. Additional stratification of the characteristics of these connections might shed more light on their role in loss of consciousness. One hypothesis is that sparsely connected regions also incorporate thalamocortical connections as evident by the increased number of sparse connections in subcortical regions during sedation. Indeed, the thalamus been shown as a primary target for anaesthetics (Brown et al., 2010) and is affected in patients with disorders of consciousness (Laureys et al., 2000a). At the same time the thalamus is important for network integration (Hwang et al., 2017) showing that alterations of its connectivity might play an increased role in disrupting whole-brain communication (Alkire et al., 2008). An additional hypothesis is that sparsely connected regions include regions with long-range connections-for example those between the anterior and posterior parts of the DMN-that are important for communicating information between distant parts of the brain (Bassett and Bullmore,

2016). Confirmatory evidence comes from previous studies showing that long-range connectivity changes in both anaesthesia and disorders of consciousness (Schröter et al., 2012; MacDonald et al., 2015). Thus changes in the complexity of long-range connectivity might impair network integration during loss of consciousness.

#### **7.4 Dynamic functional connectivity**

Evidence from anaesthetised primates and humans show that even during anaesthesia functional connectivity is not static but shows fluctuations. But what do these brain dynamics tell about loss of consciousness? Are brain dynamics altered during loss of consciousness and is this more prevalent in specific networks?

Prominent theories suggest that consciousness is related to the diversity of brain connectivity across time (Tononi, 2004) and that anaesthesia might suppress consciousness by shrinking this functional repertoire (Alkire et al., 2008). Towards this direction, previous research using dynamic connectivity methods, attempted to quantify the dynamic changes in connectivity patterns during loss of consciousness (Barttfeld et al., 2015a; Ma et al., 2017). This research discretizes functional connectivity by calculating connectivity over different time windows. The disadvantage of using sliding window methods relates to whether it can actually capture the entire repertoire of functional connectivity.

To overcome this obstacle, in chapter 5 I used a dynamic connectivity model in order to identify states that the unconscious brain would occupy during at each time point. For the sedation data, I found that brain dynamics during sedation were characterized by distinct temporal patterns as the brain would predominantly visit certain states and avoid visiting others. Interestingly, a state that was visited scarcely during sedation was associated with default mode/task-negative regions, while a state that was visited more was associated with frontal-parietal/task-positive regions. Previous studies have highlighted how switching between these networks might regulate information sharing in the brain (Fox et al., 2005). In light of this, I showed that their inter-dynamics related to the state of consciousness during sedation as higher toggling

between the two states was related to higher levels of responsiveness. Studies of anaesthesia in humans and monkeys have shown that default mode network connectivity with task positive networks was lost after propofol administration and was regained after recovery of consciousness (Boveroux et al., 2010; Barttfeld et al., 2015a), suggesting that these brain dynamics are important for transition to and emergence from loss of consciousness. Similarly, in the case of disorders of consciousness I found two states with different spatial characteristics dominating the brain dynamics of the patients. These two states also showed overlap with regions of positive (task-negative regions) and negative DMN connectivity (task-positive regions) observed in patients' functional data (Di Perri et al., 2016). Moreover, the switching rate between these states positively correlated with the CRS-R score showing that their inter-state dynamics related to their state of consciousness.

Taken together, it was evident that the switching rate between these states could discriminate between different state of consciousness during sedation or clinical assessment of. By complementing the results of chapter 4 on the effect of sedation on DMN complexity and the loss of integration, these findings show that the temporal aspect of DMN connectivity also changes during loss of consciousness. One interesting hypothesis is that loss of consciousness is characterized not only by decreased DMN complexity and its effect on whole-brain integration, but also by its pathological dynamic interactions with the task-positive networks. Together, these two distinct phenomena might diminish the spatiotemporal repertoire of connectivity, thus reducing the amount of information a brain has access to (Alkire et al., 2008). However, a study unifying these two methodologies is required to further address this.

A natural extension of this approach relates to whether the switching rate can discriminate between different states of consciousness. Towards this direction, I obtained a collection of common states for both of the sedation and disorders of consciousness populations. Because no state was population-specific, I calculated the switching rate between all these states as an aggregate measure of temporal complexity. Higher switching rate would imply that the brain would visit only a specific number of states while ignoring others i.e. a less temporally complex/more

unstable brain. In that regard, I found higher switching rate when comparing individuals during awake and sedation levels to patients with disorders of consciousness.

Thus in conjunction with the spatial complexity results presented in chapter 4, it is possible that there exist two aspects of complexity reduction during loss of consciousness: spatial as evident for more segregated networks and temporal as evident by the greater local temporal dynamics. Such a two-faceted shift in complexity has been postulated by theoretical models claiming the importance of the spatiotemporal brain repertoire in shaping conscious experience (Tononi, 2004). The involvement of default mode regions in both of these processes suggests that the spatiotemporal shift towards a more segregated configuration might have a common basis similar to what has been termed as the “dynamic core” of segregation and integration in consciousness (Tononi and Edelman, 1998). The DMN could be the neural instantiation of this common basis; it encapsulates unified neural processes of high complexity organized at a very refined temporal scale (Raichle, 2015) that emphasize integration both at the static (Vatansever et al., 2015b) and the dynamic level (Tang et al., 2017). Further studies implementing a common approach for calculating the spatiotemporal complexity can confirm this hypothesis.

### **7.5 Complexity of functional connectivity in the developing brain**

The spatiotemporal grain of consciousness and complexity is not only a feature of the adult brain. Consciousness is present since infancy; human infants experience pain and react to stimuli showing features of what is termed as basic consciousness (Lagercrantz and Changeux, 2009). Interestingly, anaesthesia can abolish consciousness in infants using the same process as in adults. In light of the work presented in previous chapters, this poses the question of whether reduction of complexity is also a characteristic of the anaesthetized infant brain.

In that regard, in chapter 6 I used EEG based connectivity derived from data recorded in infants during different levels of maintenance of surgical anaesthesia and

emergence from anaesthesia. I focused on alpha frequency connectivity for various reasons. Power analyses have suggested the emergence of alpha power during unconsciousness (Purdon et al., 2013). These alpha oscillations are believed to decrease network integration by acting on thalamocortical loops (Alkire et al., 2008). Results from adult anaesthesia and disorders of consciousness using whole-brain alpha EEG recordings show that networks during anaesthesia are more fragmented and less globally efficient (Chennu et al., 2016; Chennu et al., 2017). Thus, the way alpha-based whole-brain connectivity changes might provide a suitable framework for quantifying alterations in complexity.

First, I observed that alpha oscillations changed during anaesthesia. When comparing anaesthesia to emergence, alpha power increased during anaesthesia potentially reflecting the emergence of alpha-specific thalamocortical oscillations previously reported in adult literature (Purdon et al., 2013). However, confirming the initial hypothesis, the complexity in alpha-derived networks was reduced in anaesthesia compared to emergence. This result was verified using source localization methods showing that decrease in complexity also takes place at the source level. These results validate and extend the findings presented in chapter 4 where I showed reduced complexity in the anaesthetized adult brain. Further analysis can elucidate whether reduction in complexity in the infant brain is also more evident in default mode regions in a similar way to what has been previously proposed in this thesis. Towards this direction, the DMN is one of the first higher-order networks to show a well-established structure in the developmental context by comprising robust functional connectivity across distant medial frontal, medial/lateral parietal, and medial/lateral temporal regions (Gao et al., 2015). This finding shows that the DMN likely serves as a foundation for other higher-order functions to integrate with each other; thus it is possible that its complexity has an increased role in consciousness since infancy. In addition, I found that alpha complexity was positively correlated with age during anaesthesia. Increased complexity indicated that brain networks become more integrated, potentially showing that the brain is developing in order to attain a higher level of conscious processing (Fair et al., 2007b). The developmental trajectory of functional connectivity starts with the emergence of primary sensory systems followed by the formation of higher-order and default mode networks (Gao et al., 2009). Thus

the ongoing integration of the developing brain could be attributed to the rapid development of networks that are responsible for combining multi-modal information coming from primary sensory systems (Gao et al., 2015). In conjunction with the previous findings in this thesis, the DMN might have an enhanced role in how this information is integrated and, thus, a profound impact on the increasing brain complexity during development.

Overall, alpha complexity was able to discriminate between stages of unconsciousness in the infant brain showing the possibility of translating the results derived from the adult brain to the infant brain. At the same time, the complexity increased with increasing age showing different aspects of functional maturation across development.

## **7.6 Clinical insights**

As evident by previous remarks, complexity was used for discriminating between different states of consciousness. Refined approaches showed that in order to perturbate whole-brain complexity, alterations in the complexity of few regions important for network integration might suffice.

The complexity approach proposed here could therefore find application as a tool for identifying different states of consciousness, complementary to the clinical assessment. All of the data analysed were obtained during resting-state condition suggesting that the state of consciousness could potentially be assessed irrespectively of the individual's behaviour (e.g. response to stimuli or commands). However, a question remains as to how the current resting-state functional connectivity approach is enough to identify the neural correlates of consciousness. A drawback of the resting-state paradigms is that they do not shed light into how complex information processing breaks down during loss of consciousness. For example studies have shown that fronto-parietal network is disrupted in unresponsive patients showing a lack of executive function in these patients (Naci et al., 2014). Until we have a clear picture of the relationship between task-based and resting-state functional connectivity (such efforts start to emerge; see Tavor et al., 2016), complementary task-based studies will

need to take place. Eventually, I claim that the goal of obtaining consciousness markers relying only on resting-state data can provide a valuable avenue for understanding consciousness not only from a practical perspective (ease of acquisition, no need for task design) but also from a clinical perspective (use as input to classifiers for categorizing disorders of consciousness cohorts, e.g. see Demertzi et al., 2015).

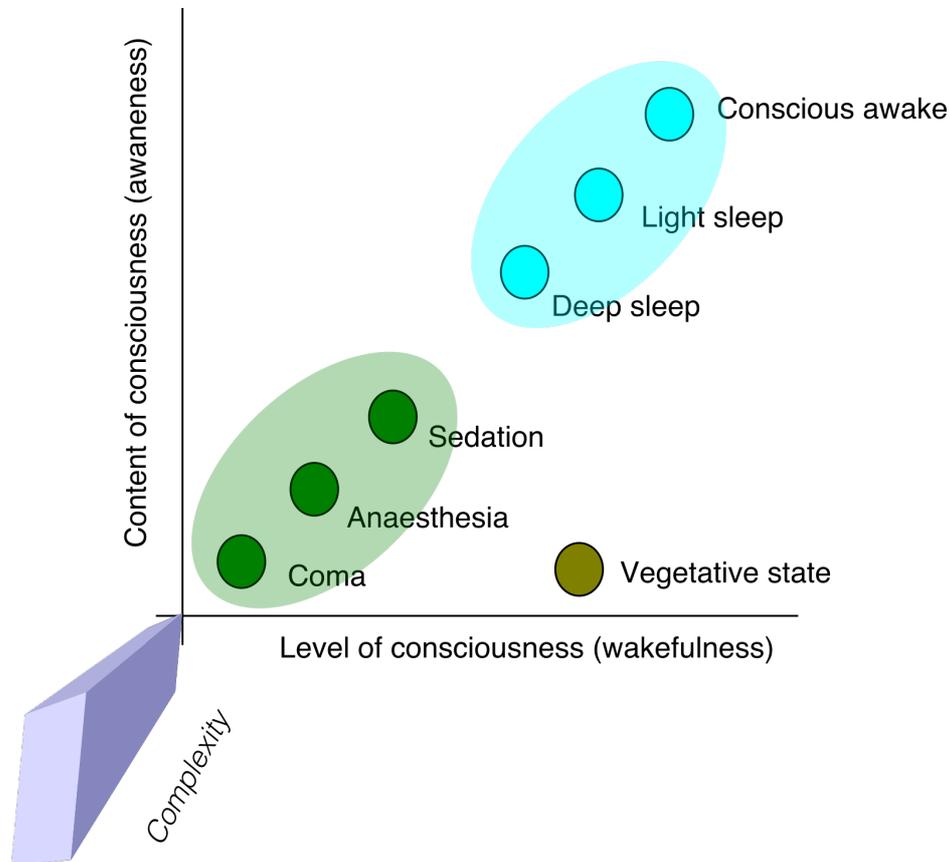
In that note, further research and validation of complexity in multiple datasets might assist with machine-learning classification between different states of consciousness in novel datasets.

The importance of complexity was also evident in infant anaesthesia. Despite the vast clinical need, there are still no robust markers for tracking brain dynamics in infant anaesthesia. I hope that the results presented in this thesis will pave the way for using complexity as a means of understanding and identifying anaesthetic states in the developing brain.

Finally it is worth noting a particular limitation of the experimental method followed in this thesis. The goal of this work was to introduce spatio-temporal complexity as a way of capturing loss of consciousness. In order to verify the utility of the complexity measure, correlations took place between complexity and the level of responsiveness (in sedation) or CRS-R scores (in the DOC cohort). However, such measures might not always tell the full story about whether someone is unconscious or not; for example it could be the case that some individuals are in a state of consciousness in which many of the cognitive and behavioural systems are still 'on-line' (Bayne et al., 2016) Thus additional research is required to investigate the multifaceted cognitive and behavioural aspects that are lost during loss of consciousness and how these relate to complexity.

Hopefully, by introducing and verifying complexity measures as objective markers of loss consciousness we can assist with the multi-dimensional problem of classifying states of consciousness. The end goal is to reduce the known taxonomy of consciousness to a single dimension above and beyond awareness and wakefulness (Fig. 7.1). This is something that fits nicely with the basic motivation that this thesis was built upon in the introduction: the clinic. Specifically, I hope that the use of

complexity will pave a new, objective determination of consciousness that can assist with individuals' prognosis and recovery at the bedside without the need for multiple and complex behavioural and cognitive assessments.



**Figure 7.1. - Complexity as an objective marker of consciousness.** The goal is to use complexity as a different dimension of mapping consciousness above and beyond the two-dimensional map of wakefulness and awareness.

### 7.7 Future work

There are additional future directions that can extend the work of each chapter.

The proposed NNG model in chapter 3 for explaining the criticality of complexity of functional connectivity networks took into account the topographic location of each region and employed a wiring minimization rule to produce a network of optimal connections. One can argue that besides the topographic location, other morphological features of brain organization (for example axon tensions, wiring

constraints, etc.) might need to be taken into consideration and compared against the current NNG model.

In chapter 4, I argued about the complexity of sparsely connected regions as a measure of classifying different states of consciousness. I made a specific hypothesis that the complexity of thalamic connectivity would change as its role in network integration has been previously demonstrated (Hwang et al., 2017). In light of this, a more comprehensive analysis including a parcellation of the thalamus might provide more information regarding the complexity of thalamocortical connectivity and how it is being affected during loss of consciousness.

Moreover, I would like to use the complexity of sparsely connected regions to further investigate the classification between patients *within* the disorders of consciousness cohort (for example classifying between MCS and VS patients). This could be an additional extension to previous EEG studies regarding classification between different states of consciousness in patients (Sitt et al., 2014; Chennu et al., 2017).

In chapter 5, I used the HMM model to derive a set of states that characterized brain dynamics. A major issue of these models is that they assume that the BOLD data is stationary (Liégeois et al., 2017) meaning that its mean and variance do not change over time. However, studies have suggested that this is not always the case (Laumann, et al. 2017). Thus I would like to utilize other dynamic models that can address this problem in a more systematic way.

Finally, by combining methods from chapters 4 and 5, I would like to derive a common model that can incorporate both spatial and temporal complexity. This will allow me to address hypotheses regarding the spatiotemporal reduction in complexity during loss of consciousness as well as to search for a common core of regions that can cause both loss of spatial and temporal complexity (for example DMN regions).

---

## **7.8 Conclusion**

How do we lose consciousness? Evidence presented in this thesis attempted to answer this question by showing that loss of consciousness occurs because of a drastic decrease in complexity. This measure was able to discriminate between different states of consciousness in the adult and infant brain, by exploring both spatial and temporal domains. I believe that this provides a new avenue for understanding the relevance of resting-state functional connectivity for the assessment of different states of consciousness irrespectively of the behavioural capacity or age of the individual.

---

## References

Absalom AR, Menon DK (2009) BIS and spectral entropy monitoring during sedation with midazolam/remifentanil and dexmedetomidine/remifentanil. *Crit Care* 13:137.

Achard S, Delon-Martin C, Vértes PE, Renard F, Schenck M, Schneider F, Heinrich C, Kremer S, Bullmore ET (2012) Hubs of brain functional networks are radically reorganized in comatose patients. *Proc Natl Acad Sci USA* 109:20608-20613.

Adapa RM, Davis MH, Stamatakis EA, Absalom AR, Menon DK (2014) Neural correlates of successful semantic processing during propofol sedation. *Hum Brain Mapp* 35:2935-2949.

Akalin-Acar Z, Gençer NG (2004) An advanced boundary element method (BEM) implementation for the forward problem of electromagnetic source imaging. *Phys Med Biol* 49:5011-5028.

Alexander A (2011) Deterministic white matter tractography. In: *Diffusion MRI: Theory, Methods, and Applications* (Jones DK, ed). Oxford University Press, Oxford UK.

Alcauter S, Lin W, Smith JK, Short SJ, Goldman BD, Reznick JS, Gilmore JH, Gao W (2014) Development of thalamocortical connectivity during infancy and its cognitive correlations. *J Neurosci* 34:9067-9075.

Alkire MT, Haier RJ, Barker SJ, Shah NK, Wu JC, Kao YJ (1995) Cerebral metabolism during propofol anesthesia in humans studied with positron emission tomography. *Anesthesiology* 82:393-403.

Alkire MT, Hudetz AG, Tononi G (2008) Consciousness and anaesthesia. *Science* 322:876-880.

Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T, Calhoun VD (2012) Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex* 24:663-676.

- Alter T, Walter S (2007) Phenomenal concepts and phenomenal knowledge. University Press Scholarship Online.
- Amaro E, Jr, Barker GJ (2006) Study design in fMRI: basic principles. *Brain Cogn* 60:220-232.
- Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26:839-851.
- Ashburner J (2009) Preparing fMRI data for statistical analysis. In: *fMRI Techniques and Protocols* (Filippi M, ed). 151-178: Humana Press.
- Avena-Koenigsberger A, Misic B, Sporns O (2017) Communication dynamics in complex brain networks. *Nat Rev Neurosci* 19:17-33.
- Baars BJ (1988) *A cognitive theory of consciousness*. Cambridge University Press.
- Baars BJ (2002) The conscious access hypothesis: origins and recent evidence. *Trends Cogn Sci* 6:47-52.
- Bai L, Ren P, Bai X, Hancock ER (2014) A graph kernel from the depth-based representation. In: *Structural, Syntactic, and Statistical Pattern Recognition. S+SSPR 2014. Lecture Notes in Computer Science* (Fränti P, Brown G, Loog M, Escolano F, Pelillo M, eds). Vol 8621, Springer, Berlin, Heidelberg.
- Baillet S (2017) Magnetoencephalography for brain electrophysiology and imaging. *Nat Neurosci* 20:327-339.
- Bak P, Paczuski M (1995) Complexity, contingency, and criticality. *Proc Natl Acad Sci USA* 92: 6689-6696.
- Bak P (1996) *How nature works: the science of self-organised criticality*. New York, NY: Copernicus Press.
- Baker AP, Brookes MJ, Rezek IA, Smith SM, Behrens T, Probert Smith PJ, Woolrich M (2014) Fast transient networks in spontaneous human brain activity. *Elife* 3:e01867.

- Balduzzi D, Tononi G (2008) Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLOS Comput Biol* 4:e1000091.
- Barrett AB, Seth AK (2011) Practical measures of integrated information for time-series data. *PLOS Comput Biol* 7:e1001052.
- Barttfeld P, Uhrig L, Sitt JD, Sigman M, Jarraya B, Dehaene S (2015a) Signature of consciousness in the dynamics of resting-state brain activity. *Proc Natl Acad Sci USA* 112:887-892.
- Barttfeld P, Bekinschtein TA, Salles A, Stamatakis EA, Adapa R, Menon DK, Sigman M (2015b) Factoring the brain signatures of anesthesia concentration and level of arousal across individuals. *Neuroimage Clin* 9:385-391.
- Bassett DS, Bullmore ET (2006) Small-world brain networks. *Neuroscientist* 12:512-523.
- Bassett DS, Bullmore ET (2016) Small-world brain networks revisited. *Neuroscientist* 3:499-516.
- Batchelor PG, Moakher M, Atkinson D, Calamante F, Connelly A (2005) A rigorous framework for diffusion tensor calculus. *Magnetic Resonance in Medicine* 53:221-225.
- Bayne T, Hohwy J, Owen AM (2016) Are there levels of consciousness? *Trends Cogn Sci* 20:405-413.
- Becker CO, Pequito S, Pappas GJ, Miller MB, Grafton ST, Bassett DS, Preciado VM (2018) Spectral mapping of brain functional connectivity from diffusion imaging. *Sci Rep* 8:1411.
- Beckmann CF, DeLuca M, Devlin JT, Smith SM (2005) Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci* 360:1001-1013.
- Beckmann CF, Mackay CE, Filippini N, Smith SM (2009) Group comparison of resting-state fMRI data using multi-subject ICA and dual regression. *Neuroimage* 47:S148.

- Beggs JM, Plenz D (2003) Neuronal avalanches. *J Neurosci* 23:11167-11177.
- Beggs JM (2008) The criticality hypothesis: how local cortical networks might optimize information processing. *Philos Transact A Math Phys Eng Sci* 366:329-343.
- Behrens TE, Berg HJ, Jbabdi S, Rushworth MF, Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34:144-155.
- Behzadi Y, Restom K, Liao J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37:90-101.
- Berger H (1929) Über das Elektroencephalogramm des Menschen. *Arch Psychiatr Nervenkr* 87:527-570.
- Betzl RF, Bassett DS (2016) Multi-scale brain networks. *Neuroimage* 160:73-83.
- Betzl RF, Medaglia J, Papadopoulos L, Baum G, Gur R, Gur R, Roalf D, Satterthwaite T, Bassett DS (2017) The modular organization of human anatomical brain networks: Accounting for the cost of wiring. *Network Neurosci* 1:42-68.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Rao SM, Cox RW (1999) Conceptual processing during the conscious resting state. A functional MRI study. *J Cogn Neurosci* 11:80-95.
- Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34:537-541.
- Blain-Moraes S, Tarnal V, Vanini G, Bel-Behar T, Janke E, Picton P, Golmirzaie G, Palanca BJA, Avidan MS, Kelz MB, Mashour GA (2017) Network efficiency and posterior alpha patterns are markers of recovery from general anesthesia: a high-density electroencephalography study in healthy volunteers. *Front Hum Neurosci* 11:328.
- Blink EJ (2004) Basic MRI Physics. In. <http://www.mri-physics.net/>

Bloch F, Hansen WW, Packard M (1946) Nuclear induction. *Physical Review* 70:474-489.

Bokil H, Andrews P, Kulkarni JE, Mehta S, Mitra PP (2010) Chronux: a platform for analyzing neural signals. *J Neurosci Methods* 192:146-151.

Boly M, Tshibanda L, Vanhaudenhuyse A, Noirhomme Q, Schnakers C, Ledoux D, Boveroux P, Garweg C, Lambermont B, Phillips C, Luxen A, Moonen G, Basseti C, Maquet P, Laureys S (2009) Functional connectivity in the default network during resting state is preserved in a vegetative but not in a brain dead patient *Hum Brain Mapp* 30:2393-2400.

Boly M, Seth AK (2012) Modes and models in disorders of consciousness science. *Arch Ital Biol* 150:172-184.

Boly M, Moran R, Murphy M, Boveroux P, Bruno MA, Noirhomme Q, Ledoux D, Bonhomme V, Brichant JF, Tononi G, Laureys S, Friston K (2012) Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. *J Neurosci* 32:7082-7090.

Boveroux P, Vanhaudenhuyse A, Bruno MA, Noirhomme Q, Lauwick S, Luxen A, Degueldre C, Plenevaux A, Schnakers C, Phillips C, Brichant JF, Bonhomme V, Maquet P, Greicius MD, Laureys S, Boly M (2010) Breakdown of within- and between-network resting state functional magnetic resonance imaging connectivity during propofol-induced loss of consciousness. *Anesthesiology* 113:1038-1053.

Boyd S, Vandenberghe L (2007) *Convex optimization*. Cambridge University Press, New York.

Bressler SL, Tognoli E (2006) Operational principles of neurocognitive networks. *Int J Psychophysiol* 60:139-148.

Bressler SL, Menon V (2010) Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn Sci* 14:277-290.

- Brodmann K (1909) Vergleichende lokalisationslehre der großhirnrinde. Leipzig: J. A. Barth.
- Brody BA, Kinney HC, Kloman AS, Gilles FH (1987) Sequence of central nervous system myelination in human infancy. I. An autopsy study of myelination. *Journal of Neuropathology and Experimental Neurology*.
- Brown EN, Lydic R, Schiff ND (2010) General anesthesia, sleep, and coma. *N Engl J Med* 363:2638-2650.
- Brown EN, Purdon PL, Van Dort CJ (2011) General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annu Rev Neurosci* 34:601-628.
- Brunel N (2016) Is cortical connectivity optimized for storing information? *Nat Neurosci* 19:749-755.
- Bruno MA, Vanhaudenhuyse A, Thibaut A, Moonen G, Laureys S (2011) From unresponsive wakefulness to minimally conscious PLUS and functional locked-in syndromes: recent advances in our understanding of disorders of consciousness. *J Neurol* 258:1373-1384.
- Bullmore ET, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10:186-198.
- Bullmore ET, Sporns O (2012) The economy of brain network organization. *Nat Rev Neurosci* 13:336-349.
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304:1926-1929.
- Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. *Nat Rev Neurosci* 13:407-420.
- Cajal SR (1995) *Histology of the nervous system of man and vertebrates*. Oxford University Press, New York.

Cammoun L, Gigandet X, Meskaldji D, Thiran JP, Sporns O, Do KQ, Maeder P, Meuli R, Hagmann P (2012) Mapping the human connectome at multiple scales with diffusion spectrum MRI. *J Neurosci Methods* 203:386-397.

Carhart-Harris RL, Leech R, Hellyer PJ, Shanahan M, Fielding A, Tagliazucchi E, Chialvo DR, Nutt D (2014) The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front Hum Neurosci* 8:20.

Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Cassali KR, Casarotto S, Bruno MA, Laureys S, Tononi G, Massimini M (2013) A theoretically based index of consciousness independent of sensory processing and behavior. *Sci Transl Med* 5:198ra105.

Casarotto S, Comanducci A, Rosanova M, Sarasso S, Fecchio M, Napolitani M, Pigorini, Casali AG, Trimarchi PD, Boly M, Gosseries O, Bodart O, Curto F, Landi C, Mariotti M, Devalle G, Laureys S, Tononi G, Massimini M (2016) Stratification of unresponsive patients by an independently validated index of brain complexity. *Ann Neurol* 80:718-729.

Catani M, De Schotten MT, Slater D, Dell'Acqua F (2013) Connectomic approaches before the connectome. *Neuroimage* 80:2-13.

Chai XJ, Castañón A, Ongür D, Whitfield-Gabrieli S (2012) Anticorrelations in resting state networks without global signal regression. *Neuroimage* 59:1420-1428.

Chalmers DJ (1995) Facing up to the problems of consciousness. *J Conscious Stud* 2:200-219.

Chalmers DJ (1996) *The conscious mind: In search of a fundamental theory*. Oxford University Press.

Chang C, Liu Z, Chen MC, Liu X, Duyn JH (2013) EEG correlates of time-varying BOLD functional connectivity. *Neuroimage* 72:227-236.

Chen YJ, Lo YC, Hsu YC, Fan CC, Hwang TJ, Liu CM, Chien YL, Hsieh MH, Liu CC, Hwu HG, Tseng WY (2015) Automatic whole brain tract - based analysis using

- predefined tracts in a diffusion spectrum imaging template and an accurate registration strategy. *Hum Brain Mapp* 36:3441-3458.
- Chennu S, O'Connor S, Adapa RM, Menon DK, Bekinschtein TA (2016) Brain connectivity dissociates responsiveness from drug exposure during propofol-induced transitions of consciousness. *PLOS Comput Biol* 12:e1004669.
- Chennu S, Annen J, Wannez S, Thibaut A, Chatelle C, Cassol H, Martens G, Schnakers C, Gosseries O, Menon DK, Laureys S (2017) Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness. *Brain* 140: 2120-2132.
- Cherniak C (1994) Component placement optimization in the brain. *J Neurosci* 14:2418-2427.
- Cherniak C, Mokhtarzada Z, Rodriguez-Esteban R, Changizi K (2004) Global optimization of cerebral cortex layout. *Proc Natl Acad Sci USA* 101:1081-1086.
- Ching S, Cimenser A, Purdon PL, Brown EN, Kopell NJ (2010) Thalamocortical model for a propofol-induced alpha-rhythm associated with loss of consciousness. *Proc Natl Acad Sci USA* 107:22665-22670.
- Chklovskii DB, Schikorski T, Stevens CF (2002) Wiring optimization in cortical circuits. *Neuron* 34:341-347.
- Christensen K, Moloney MR (2005) Complexity and criticality. World Scientific Publishing Company.
- Chung FRK (1997) Spectral graph theory. American Mathematical Society.
- Cohen-Adad J, Descoteaux M, Rossignol S, Hoge RD, Deriche R (2008) Detection of multiple pathways in the spinal cord using q-ball imaging. *Neuroimage* 42:739-749.
- Cohen MX (2017) Where does EEG come from and what does it mean? *Trends Neurosci* 40:208-218.

- Cole DM, Smith SM, Beckmann CF (2010) Advances and pitfalls in the analysis and interpretation of resting-state fMRI data. *Front Syst Neurosci* 4:8.
- Cole MW, Repovš G, Anticevic A (2014a) The frontoparietal control system: a central role in mental health. *Neuroscientist* 20:652-664.
- Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE (2014b) Intrinsic and task-evoked network architectures of the human brain. *Neuron* 83:238-251.
- Conel JL (1963) The postnatal development of the human cerebral cortex. Harvard University Press. Cambridge, MA.
- Conklin J, Eliasmith C (2005) An attractor network model of path integration in the rat. *Journal of Computational Neuroscience* 18:183-203.
- Contreras D, Destexhe A, Sejnowski TJ, Steriade M (1996) Control of spatiotemporal coherence of a thalamic oscillation by corticothalamic feedback. *Science* 274:771-774.
- Cornelissen L, Kim SE, Purdon PL, Brown EN, Berde CB (2015) Age-dependent electroencephalogram (EEG) patterns during sevoflurane general anesthesia in infants. *Elife*. 4:e06513.
- Cornelissen L, Kim SE, Lee JM, Brown EN, Purdon PL, Berde CB (2018) Electroencephalographic markers of brain development during sevoflurane anaesthesia in children up to 3 years old. *Br J Anaesth* 120:1274-1286.
- Crick F, Koch C (1990) Towards a neurobiological theory of consciousness. *Semin Neurosci* 2:263-275.
- Crick F, Koch C (2003) A framework for consciousness. *Nat Neurosci* 6:119-126.
- Crone JS, Soddu A, Höller Y, Vanhaudenhuyse A, Schurz M, Bergmann J, Schmid E, Trinka E, Laureys S, Kronbichler M (2014) Altered network properties of the frontoparietal network and the thalamus in impaired consciousness. *Neuroimage Clin* 4:240-248.

- Curley EM (1986) Spinoza 's geometric method - notes. *Studia Spinozana: an International and Interdisciplinary Series* 2:164.
- Daducci A, Gerhard S, Griffa A, Lemkaddem A, Cammoun L, Gigandet X, Meuli R, Hagmann P, Thiran JP (2012) The Connectome Mapper: An open-source processing pipeline to map connectomes with MRI. *PLOS ONE* 7:e48121.
- Damadian R (1971) Tumor detection by nuclear magnetic resonance. *Science* 171:1151-1153.
- Damasio AR (1989) The brain binds entities and events by multiregional activation from convergence zones. *Neural Comp* 1:123-132.
- Darvas F, Pantazis D, Kucukaltun-Yildirim E, Leahy RM (2004) Mapping human brain function with MEG and EEG: methods and validation. *Neuroimage* 23 Suppl 1:S289-99.
- Deco G, Jirsa V (2012) Ongoing cortical activity at rest: criticality, multistability and ghost attractors. *J Neurosci* 32:3366-3375.
- Dehaene S, Naccache L (2001) Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79:1-37.
- Dehaene S, Changeux JP, Naccache L, Sackur J, Sergent C (2006) Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn Sci* 10:204-211.
- Dehaene S, Changeux JP (2011) Experimental and theoretical approaches to conscious processing. *Neuron* 70:200-227.
- Dehaene S (2014) *Consciousness and the brain: deciphering how the brain codes our thoughts*. New York, New York: Viking Adult.
- Deichmann R (2009) Principles of MRI and functional MRI. In: *fMRI techniques and protocols* (Filippi M, ed). 3-29: Humana Press.

- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics *Journal of Neurosci Methods* 134:9-21.
- Del Ferraro G, Moreno A., Min B, Morone F, Pérez-Ramírez Ú, Pérez-Cervera L, Parra LC, Holodny A, Canals S, Makse HA (2018) Finding influential nodes for integration in brain networks using optimal percolation theory. *Nat Commun* 9:2274.
- Del Giudice M, Booth T, Irwing P (2012) The distance between mars and venus: measuring global sex differences in personality. *PLOS ONE* 7:e29265.
- Demertzi A, Vanhaudenhuyse A, Brédart S, Heine L, di Perri C, Laureys S (2013) Looking for the self in pathological unconsciousness. *Front Hum Neurosci* 7:538.
- Demertzi A, Antonopoulos G, Heine L, Voss HU, Crone JS, de Los Angeles C, Bahri MA, Di Perri C, Vanhaudenhuyse A, Charland-Verville V, Kronbichler M, Trinka E, Phillips C, Gomez F, Tshibanda L, Soddu A, Schiff ND, Whitfield-Gabrieli S, Laureys S (2015) Intrinsic functional connectivity differentiates minimally conscious from unresponsive patients. *Brain* 138:2619-2631.
- Dennett DC (1988) Quining qualia. *Conscious Mod Sci* 1-26.
- Dennett DC (2005) *Sweet dreams: Philosophical obstacles to a science of consciousness*. MIT Press.
- de Pasquale F, Della Penna S, Snyder AZ, Lewis C, Mantini D, Marzetti L, Belardinelli P, Ciancetta L, Pizzella V, Romani GL, Corbetta M (2010) Temporal dynamics of spontaneous MEG activity in brain networks. *Proc Natl Acad Sci USA* 107:6040-6045.
- Di Perri C, Bastianello S, Bartsch AJ, Pistarini C, Maggioni G, Magrassi L, Imberti R, Pichiecchio A, Vitali P, Laureys S, Di Salle F (2013) Limbic hyperconnectivity in the vegetative state. *Neurology* 81:1417-1424.
- Di Perri C, Bahri MA, Amico E, Thibaut A, Heine L, Antonopoulos G, Charland-Verville V, Wannez S, Gomez F, Hustinx R, Tshibanda L, Demertzi A, Soddu A, Laureys S (2016) Neural correlates of consciousness in patients who have emerged

from a minimally conscious state: a cross-sectional multimodal imaging study. *Lancet Neurol* 15:830-842.

Di Perri C, Amico E, Heine L, Annen J, Martial C, Larroque SK, Soddu A, Marinazzo D, Laureys S (2018) Multifaceted brain networks reconfiguration in disorders of consciousness uncovered by co-activation patterns. *Hum Brain Mapp* 39:89-103.

Dixon ML, Andrews-Hanna JR, Spreng RN, Irving ZC, Mills C, Girn M, Christoff K (2017) Interactions between the default network and dorsal attention network vary across default subsystems, time, and cognitive states. *Neuroimage* 147:632-649.

Dosenbach DU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE (2007) Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci USA* 104:11073-11078.

Einstein A (1956) *Investigations on the theory of the brownian movement*. Dover publications.

Ercsey-Ravasz M, Markov NT, Lamy C, Van Essen DC, Knoblauch K, Toroczkai Z, Kennedy H (2013) A predictive network model of cerebral cortical connectivity based on a distance rule. *Neuron* 89:184-197.

Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM (1993) 3d statistical neuroanatomical models from 305 MRI volumes. *Nuclear Science Symposium & Medical Imaging Conference Vols 1-3*:1813-1817.

Fabrikant A, Luthra A, Maneva E, Papadimitriou CH, Shenker S (2003) On a network creation game. In: *Podc'03 (Proceedings of the Twenty-Second Annual Symposium on Principles of Distributed Computing)* 347-351, Boston, Massachusetts.

Fair DA, Schlaggar BL, Cohen AL, Miezin FM, Dosenbach NU, Wenger KK, Fox MD, Snyder AZ, Raichle ME, Petersen SE (2007a) A method for using blocked and event-related fMRI data to study "resting state" functional connectivity. *Neuroimage* 35:396-405.

Fair DA, Dosenbach NU, Church JA, Cohen AL, Brahmbhatt S, Miezin FM, Barch DM, Raichle ME, Petersen SE, Schlaggar BL (2007b) Development of distinct control networks through segregation and integration. *Proc Natl Acad Sci USA* 104:13507-13512.

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1-47.

Fernández-Espejo D, Soddu A, Cruse D, Palacios EM, Junque C, Vanhaudenhuyse A, Rivas E, Newcombe V, Menon DK, Pickard JD, Laureys S, Owen AM (2012) A role for the default mode network in the bases of disorders of consciousness. *Ann Neurol* 72:335-343.

Ferrari M, Mottola L, Quaresima V (2014) Principles, techniques, and limitations of near infrared spectroscopy. *Can J Appl Physiol* 29:463-487.

Fischl B, Sereno MI, Tootell RB, Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* 8:272-284.

Fiset P, Paus T, Daloze T, Plourde G, Meuret P, Bonhomme V, Hajj-Ali N, Backman SB, Evans AC (1999) Brain mechanisms of propofol-induced loss of consciousness in humans: a positron emission tomographic study. *J Neurosci* 19:5506-5513.

Fornito A, Zalesky A, Bullmore ET (2016) *Fundamentals of brain network analysis*. Elsevier academic press.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102:9673-9678.

Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8:700-711.

Fox MD, Greicius M (2010) Clinical applications of resting state functional connectivity. *Front Syst Neurosci* 4:19.

- Fox PT, Raichle ME (1986) Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proc Natl Acad Sci USA* 83:1140-1144.
- Fox PT (1995) Spatial normalization origins: objectives, applications, and alternatives. *Hum Brain Mapp* 3:161-164.
- Frahm J, Merboldt KD, Hancicke W (1993) Functional MRI of human brain activation at high spatial-resolution. *Magn Reson Med* 29:139-144.
- Fransson P, Skiöld B, Horsch S, Nordell A, Blennow M, Lagercrantz H, Aden U. Resting-state networks in the infant brain (2007) *Proc Natl Acad Sci USA* 104:15531-15536.
- Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci* 9:474-480.
- Fujita I, Tanaka K, Ito M, Cheng K (1992) Columns for visual features of objects in monkey inferotemporal cortex. *Nature* 360:343-346.
- Fuster JM (2003) *Cortex and mind: unifying cognition*. New York: Oxford University Press.
- Gallos LK, Makse HA, Sigman M (2012) A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proc Natl Acad Sci USA* 109:2825-2830.
- Gao W, Zhu H, Giovanello KS, Smith JK, Shen D, Gilmore JH, Lin W (2009) Evidence on the emergence of the brain's default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proc Natl Acad Sci USA* 106:6790-6795.
- Gao W, Alcauter S, Smith JK, Gilmore JH, Lin W (2015) Development of human brain cortical network architecture during infancy. *Brain Struct Funct* 220:1173-1186.
- Gerrard P, Zafonte R, Giacino JT (2014) Coma recovery scale-revised: evidentiary support for hierarchical grading of level of consciousness. *Arch Phys Med Rehabil* 95: 2335-2341.

- Giacino JT, Ashwal S, Childs N, Cranford R, Jennett B, Katz DI, Kelly JP, Rosenberg JH, Whyte J, Zafonte RD, Zasler ND (2002) The minimally conscious state: definition and diagnostic criteria. *Neurology* 58:349-353.
- Giacino JT, Kalmar K, Whyte J (2004) The JFK coma recovery scale-revised: measurement characteristics and diagnostic utility. *Arch Phys Med Rehabil* 85:2020-2029.
- Goense JB, Logothetis NK (2008) Neurophysiology of the BOLD fMRI signal in awake monkeys. *Curr Biol* 18:631-640.
- Gosseries O, Bruno MA, Chatelle C, Vanhaudenhuyse A, Schnakers C, Soddu A, Laureys S (2011) Disorders of consciousness: what's in a name? *NeuroRehabilitation* 28:3-14.
- Goulas A, Schaefer A, Margulies DS (2015) The strength of weak connections in the macaque cortico-cortical network. *Brain Struct Funct* 220:2939-2951.
- Granovetter MS (1973) The strength of weak ties. *American Journal of Sociology* 78:1360-1380.
- Grave de Peralta Menendez R, Gonzalez Andino SL (1999) Backus and Gilbert method for vector fields. *Hum Brain Mapp* 7:161-165.
- Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M, WU-Minn HCP Consortium (2013) The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80:105-124.
- Grassberger P (2012) Randomness, information, and complexity. *arXiv:1208.3459v1*.
- Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci USA* 100:253-258.

Greicius MD, Kiviniemi V, Tervonen O, Vainionpää V, Alahuhta S, Reiss AL, Menon V (2008) Persistent default-mode network connectivity during light sedation. *Hum Brain Mapp* 29:839-847.

Greicius MD, Supekar K, Menon V, Dougherty RF (2009) Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb Cortex* 19:72-78.

Gugino LD, Chabot RJ, Prichep LS, John ER, Formanek V, Agho LS (2001) Quantitative EEG changes associated with loss and return of consciousness in healthy adult volunteers anaesthetized with propofol or sevoflurane. *Br J Anaesth* 87:421-428.

Guldenmund P, Demertzi A, Boveroux P, Boly M, Vanhaudenhuyse A, Bruno MA, Gosseries O, Noirhomme O, Brichant JF, Bonhomme V, Laureys S, Soddu A (2013) Thalamus, brainstem and salience network connectivity changes during propofol-induced sedation and unconsciousness. *Brain Connect* 3:273-285.

Guldenmund P, Gantner IS, Baquero K, Das T, Demertzi A, Boveroux P, Bonhomme V, Vanhaudenhuyse A, Bruno MA, Gosseries O, Noirhomme Q, Kirsch M, Boly M, Owen AM, Laureys S, Gómez F, Soddu A (2016) Propofol-induced frontal cortex disconnection: a study of resting-state networks, total brain connectivity, and mean BOLD signal oscillation frequencies. *Brain Connect* 6:225-237.

Gulyás A, Bíró JJ, Kőrösi A, Rétvári G, Krioukov D (2015) Navigable networks as Nash equilibria of navigation games. *Nat Commun* 6:7651.

Gusnard DA, Raichle ME (2001) Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685-694.

Hagmann P, Jonasson L, Maeder P, Thiran JP, Wedeen VJ, Meuli R (2006) Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics* 26 Suppl 1:S205-223.

Hagmann P, Kurant M, Gigandet X, Thiran P, Wedeen VJ, Meuli R, Thiran JP (2007) Mapping human whole-brain structural networks with diffusion MRI. *PLOS ONE* 2:e597.

- Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O (2008) Mapping the structural core of the human cerebral cortex. *PLOS Biology* 6:e159.
- Haimovici A, Tagliazucchi E, Balenzuela P, Chialvo DR (2013) Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Phys Rev Lett* 110:178101.
- Hammers A, Allom R, Koeppe MJ, Free SL, Myers R, Lemieux L, Mitchell TN, Brooks DJ, Duncan JS (2003) Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp* 19:224-247.
- Harris B (1975) The statistical estimation of entropy in the non-parametric case. *NTIS*.
- Hatfield G (2000) Descartes' Naturalism about the Mental. In: *Descartes' Natural Philosophy* (Gaukroger S, Schuster J, Sutton J, eds). 630-658. London: Routledge.
- Heine L, Soddu A, Gómez F, Vanhaudenhuyse A, Tshibanda L, Thonnard M, Charland-Verville V, Kirsch M, Laureys S, Demertzi A (2012) Resting state networks and consciousness: alterations of multiple resting state network connectivity in physiological, pharmacological, and pathological consciousness states. *Front Psychol* 3:295.
- Henderson CJ, Butler SR, Glass (1975) The localization of equivalent dipoles of EEG sources by the application of electrical field theory. *Electroencephalogr Clin Neurophysiol* 39:117-130.
- Hight D, Voss LJ, Garcia PS, Sleight J (2017) Changes in alpha frequency and power of the electroencephalogram during volatile-based general anesthesia. *Front Syst Neurosci* 11:36.
- Hipp JF, Hawellek DJ, Corbetta M, Siegel M, Engel AK (2012) Large-scale cortical correlation structure of spontaneous oscillatory activity *Nat Neurosci* 15:10.1038/nn.3101.

- Homae F, Watanabe H, Otobe T, Nakano T, Go T, Konishi Y, Taga G (2010) Development of global cortical networks in early infancy. *J Neurosci* 30:4877-4882.
- Hormuzdi SG, Filippov MA, Mitropoulou G, Monyer H, Bruzzone R (2004) Electrical synapses: a dynamic signaling system that shapes the activity of neuronal networks. *Biochim Biophys Acta* 1662:113-137.
- Hudetz AG (2012) General anaesthesia and human brain connectivity. *Brain Connect* 2:291-302.
- Hudetz AG, Liu X, Pillay S (2015) Dynamic repertoire of intrinsic brain states is reduced in propofol-induced unconsciousness *Brain Connect* 5:10-22.
- Hudson AE, Calderon DP, Pfaff DW, Proekt A (2014) Recovery of consciousness is mediated by a network of discrete metastable activity states. *Proc Natl Acad Sci USA* 111:9283-9288.
- Huettel SA, Obembe J, Song A, Woldorff M (2004) The BOLD fMRI refractory effect is specific to stimulus attributes: Evidence from a visual motion paradigm. *Neuroimage* 21:402-408.
- Hughes SW, Crunelli V (2005) Thalamic mechanisms of EEG alpha rhythms and their pathological implications. *Neuroscientist* 11:357-372.
- Hutcheon B, Yarom Y (2000) Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends Neurosci* 23:216-222.
- Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Della Penna S, Duyn JH, Glover GH, Gonzalez-Castillo J, Handwerker DA, Keilholz S, Kiviniemi V, Leopold DA, de Pasquale F, Sporns O, Walter M, Chang C (2013) Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80:360-378.
- Hwang K, Bertolero MA, Liu WB, D'Esposito M (2017) The human thalamus is an integrative hub for functional brain networks. *J Neurosci* 37:5594-5607.
- Jackson F (1982) Epiphenomenal qualia. *Philos Q* 32:127.

- James W (1890) *The principles of psychology*. Dover Publications.
- Jatoi MA, Kamel N, Malik AS, Faye I, Begum T (2014) A survey of methods used for source localization using EEG signals. *Biomedical Signal Processing and Control* 11:42-52.
- Jevtovic-Todorovic V (2016) General anesthetics and neurotoxicity: how much do we know? *Anesthesiol Clin* 34:439-451.
- Jones DK, Catani M, Pierpaoli C, Reeves SJ, Shergill SS, O'Sullivan M, Maguire P, Horsfield MA, Simmons A, Williams SC, Howard RJ (2005) A diffusion tensor magnetic resonance imaging study of frontal cortex connections in very-late-onset schizophrenia-like psychosis. *Am J Geriatr Psychiatry* 13:1092-1099.
- Kaiser M, Martin R, Andras P, Young MP (2007) Simulation of robustness against lesions of cortical networks. *Eur J Neurosci* 25:3185-3192.
- Kanizsa G (1979) *Organization in Vision*. Praeger.
- Kanold PO, Luhmann HJ (2010) The subplate and early cortical circuits. *Annu Rev Neurosci* 33:23-48.
- Khanna A, Pascual-Leone A, Michel CM, Farzan F (2015) Microstates in resting-state EEG: current status and future directions. *Neurosci Biobehav Rev* 49:105-113.
- Kirschstein T, Köhling R (2009) What is the source of the EEG? *Clin EEG Neurosci* 40:146-149.
- Klyachko VA, Stephens CF (2003) Connectivity optimization and the positioning of cortical areas. *Proc Natl Acad Sci USA* 100:7937-7941.
- Koch MA, Norris DG, Hund-Georgiadis M (2002) An investigation of functional and anatomical connectivity using magnetic resonance imaging. *Neuroimage* 16:241-250.
- Kotchoubey B, Lang S, Mezger G, Schmalohr D, Schneck M, Semmler A, Bostanov V, Birbaumer N (2005) Information processing in severe disorders of consciousness: vegetative state and minimally conscious state. *Clin Neurophysiol* 116:2441-2453.

- Kötter R, Sommer FT (2000) Global relationship between structural connectivity propagation in the cerebral cortex. *Philos Trans R Soc Lond B Biol Sci* 355:127-134.
- Koulakov AA, Chklovskii DB (2001) Orientation preference patterns in mammalian visual cortex: a wire length minimization approach. *Neuron* 29:519-527.
- Kozberg M, Hillman E (2016) Neurovascular coupling and energy metabolism in the developing brain. *Prog Brain Res* 225:213-242.
- Krishnaswamy P, Obregon-Henao G, Ahveninen J, Khan S, Babadi B, Iglesias JE, Hämäläinen MS, Purdon PL (2017) Sparsity enables estimation of both subcortical and cortical activity from MEG and EEG. *Proc Natl Acad Sci USA* 114:E10465-E10474.
- Kybic J, Clerc M, Abboud T, Faugeras O, Keriven R, Papadopoulos T (2005) A common formalism for the integral formulations of the forward EEG problem. *IEEE Trans Med Imaging* 24:12-28.
- Lagercrantz H, Changeux JP (2009) The emergence of human consciousness: from fetal to neonatal life. *Pediatr Res* 65:255-260.
- Lancichinetti A, Fortunato S (2012) Consensus clustering in complex networks. *Sci Rep* 2:336.
- Långsjö JW, Alkire MT, Kaskinoro K, Hayama H, Maksimow A, Kaisti KK, Aalto S, Aantaa R, Jääskeläinen SK, Revonsuo A, Scheinin H (2012) Returning from oblivion: imaging the neural core of consciousness. *J Neurosci* 32:4935-4943.
- Laughlin SB, Sejnowski TJ (2003) Communication in neuronal networks. *Science* 301:1870-1874.
- Laumann TO, Snyder AZ, Mitra A, Gordon EM, Gratton C, Adeyemo B, Gilmore AW, Nelson SM, Berg JJ, Greene DJ, McCarthy J, Tagliazucchi E, Laufs H, Schlaggar BL, Dosenbach NUF, Petersen SE (2017) On the stability of BOLD fMRI correlations. *Cereb Cortex* 27:4719-4732.

- Laureys S, Faymonville ME, Luxen A, Lamy M, Franck G, Maquet P (2000a) Restoration of thalamocortical connectivity after recovery from persistent vegetative state. *Lancet* 355:1790-1791.
- Laureys S, Faymonville ME, Lamy M (2000b) Cerebral function in vegetative state studied by positron emission tomography. In: *Yearbook of Intensive Care and Emergency Medicine*.
- Laureys S, Owen AM, Schiff ND (2004) Brain function in coma, vegetative state, and related disorders. *Lancet Neurol* 3:537-546.
- Laureys S (2005) The neural correlate of (un)awareness: lessons from the vegetative state. *Trends Cogn Sci* 9:556-559.
- Laureys S, Celesia GG, Cohadon F, Lavrijsen J, León-Carrión J, Sannita WG, Szabon L, Schmutzhard E, von Wild KR, Zeman A, Dolce G, European Task Force on Disorders of Consciousness (2010) Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. *BMC Med* 8:68.
- Lazar M (2010) Mapping brain anatomical connectivity using white matter tractography. *NMR Biomed* 23:821-835.
- Lee M, Sanders RD, Yeom S, Won D, Seo K, Kim HJ, Tononi G, Lee S (2017) Network properties in transitions of consciousness during propofol-induced sedation. *Sci Rep* 7:16791.
- Lee SJ, Ralston HJ, Drey EA, Partridge JC, Rosen MA (2005) Fetal pain: a systematic multidisciplinary review of the evidence. *JAMA* 294:947-954.
- Lee U, Oh G, Kim S, Noh G, Choi B, Mashour GA (2010) Brain networks maintain a scale-free organization across consciousness, anesthesia, and recovery: evidence for adaptive reconfiguration. *Anesthesiology* 113:1081-1091.
- Lee U, Ku S, Noh G, Baek S, Choi B, Mashour GA (2013) Disruption of frontal-parietal communication by ketamine, propofol, and sevoflurane. *Anesthesiology* 118:1264-75.

- Leszl W (2006) Democritus' works: From their titles to their contents. Brill.
- Li M, Han D, Lu S, Liu Z, Zhong N (2012) Associative information processing in parahippocampal place area (PPA): an fMRI study. *Brain Informatics* 7670:1-9.
- Liang H, Wang H (2017) Structure-Function network mapping and its assessment via persistent homology. *PLOS Comput Biol* 13:e1005325.
- Liégeois R, Laumann TO, Snyder AZ, Zhou J, Yeo BTT (2017) Interpreting temporal fluctuations in resting-state functional connectivity MRI. *Neuroimage* 163:437-455.
- Lin P, Yang Y, Gao J, De Pisapia N, Ge S, Wang X, Zuo CS, Jonathan Levitt J, Niu C (2017) Dynamic default mode network across different brain states *Sci Rep* 7:46088.
- Lindquist KA, Barrett LF (2012) A functional architecture of the human brain: emerging insights from the science of emotion. *Trends Cogn Sci* 16:533-540.
- Lindquist MA, Xu Y, Nebel MB, Caffo BS (2014) Evaluating dynamic bivariate correlations in resting-state fMRI: a comparison study and a new approach. *Neuroimage* 101:531-46.
- Llinás R, Ribary U, Contreras D, Pedroarena C (1998) The neuronal basis for consciousness. *Philos Trans R Soc Lond B Biol Sci* 353: 1841-1849.
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412:150-157.
- Lynall ME, Bassett DS, Kerwin R, McKenna PJ, Kitzbichler M, Muller U, Bullmore ET (2010) Functional connectivity and brain networks in schizophrenia. *J Neurosci* 30:9477-9487.
- Ma Y, Hamilton C, Zhang N (2017) Dynamic connectivity patterns in conscious and unconscious brain. *Brain Connect* 7:1-12.
- MacDonald AA, Naci L, MacDonald PA, Owen AM (2015) Anesthesia and neuroimaging: investigating the neural correlates of unconsciousness. *Trends Cogn Sci* 19:100-107.

- Makris N, Schlerf JE, Hodge SM, Haselgrove C, Albaugh MD, Seidman LJ, Rauch SL, Harris G, Biederman J, Caviness VS Jr, Kennedy DN, Schmahmann JD (2005) MRI-based surface-assisted parcellation of human cerebellar cortex: an anatomically specified method with estimate of reliability. *Neuroimage* 25:1146-1160.
- Mansfield P (1977) Multi-planar image formation using NMR spin echoes. *J Phys C: Solid State Phys.*
- Margulies DS, Ghosh SS, Goulas A, Falkiewicz M, Huntenburg JM, Langs G, Bezgin G, Eickhoff SB, Castellanos FX, Petrides M, Jefferies E, Smallwood J (2016) Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci USA* 113:12574-12579.
- Markov NT, Misery P, Falchier A, Lamy C, Vezoli J, Quilodran R, Gariel MA, Giroud P, Ercsey-Ravasz M, Pilaz LJ, Huissoud C, Barone P, Dehay C, Toroczkai Z, Van Essen DC, Kennedy H, Knoblauch K (2011) Weight consistency specifies regularities of macaque cortical networks. *Cereb Cortex* 21:1254-1272.
- Marsh B, White M, Morton N, Kenny GNC (1991) Pharmacokinetic model driven infusion of propofol in children. *Br J Anaesth* 67:41-48.
- Martuzzi R, Ramani R, Qiu M, Rajeevan N, Constable RT (2010) Functional connectivity and alterations in baseline brain state in humans. *Neuroimage* 49:823-834.
- Mashour GA (2014) Top-down mechanisms of anesthetic-induced unconsciousness. *Front Syst Neurosci* 8:115.
- Mashour GA (2004) Consciousness unbound: Toward a paradigm of general anesthesia. *Anesthesiology* 100:428-433.
- Mazoyer B, Zago L, Mellet E, Bricogne S, Etard O, Houdé O, Crivello F, Joliot M, Petit L, Tzourio-Mazoyer N (2001) Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res Bull* 54:287-298.

- Merker B (2007) Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behav Brain Sci* 30:63-81.
- Mhuirheartaigh RN, Rosenorn-Lanng D, Wise R, Jbabdi S, Rogers R, Tracey I (2010) Cortical and subcortical connectivity changes during decreasing levels of consciousness in humans: a functional magnetic resonance imaging study using propofol. *J Neurosci* 30:9095-9102.
- Micheloyannis S, Vourkas M, Tsirka V, Karakonstantaki E, Kanatsouli K, Stam CJ (2009) The influence of ageing on complex brain networks: a graph theoretical analysis. *Hum Brain Mapp* 30:200-208.
- Mitra P, Bokil H (2008) *Observed brain dynamics*. Oxford University Press, New York.
- Mongerson CRL, Jennings RW, Borsook D, Becerra L, Bajic D (2017) Resting-state functional connectivity in the infant brain: methods, pitfalls, and potentiality. *Front Pediatr* 5:159.
- Monti MM (2012) Cognition in the vegetative state. *Annu Rev Clin Psychol* 8:431-454.
- Monti MM, Rosenberg M, Finoia P, Kamau E, Pickard JD, Owen AM (2015) Thalamo-frontal connectivity mediates top-down cognitive functions in disorders of consciousness. *Neurology* 84:167-173.
- Mosher JC, Leahy RM, Lewis PS (1999) EEG and MEG: forward solutions for inverse methods. *IEEE Trans Biomed Eng* 46:245-259.
- Mowshowitz A, Dehmer M (2012) Entropy and complexity of graphs revisited. *Entropy* 14:559-570.
- Naccache L (2006) Psychology. Is she conscious? *Science* 313:1395-1396.
- Naccache L (2017) Minimally conscious or cortically mediated state? *Brain* 141:949-960.

- Naci L, Cusack R, Anello M, Owen AM (2014) A common neural code for similar conscious experiences in different individuals. *Proc Natl Acad Sci USA* 111: 14277-14282.
- Nagel T (1974) What is it like to be a bat? *Philosophical Review* LXXXIII 4: 435-450.
- Naro A, Bramanti A, Leo A, Cacciola A, Manuli A, Bramanti P, Calabrò RS (2018) Shedding new light on disorders of consciousness diagnosis: the dynamic functional connectivity. *Cortex* 103:316-328.
- Newman M, Strogatz S, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *arXiv:cond-mat/0007235*.
- Niedermeyer E, Lopes da Silva FH (2004) *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams and Wilkins, New York.
- Ní Mhuirheartaigh R, Warnaby C, Rogers R, Jbabdi S, Tracey I (2013) Slow-wave activity saturation and thalamocortical isolation during propofol anesthesia in humans. *Sci Transl Med* 5:208ra148.
- Nisan N (2007) *Algorithmic game theory*. Cambridge University Press, New York.
- Nir Y, Fisch L, Mukamel R, Gelbard-Sagiv H, Arieli A, Fried I, Malach R (2007) Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Curr Biol* 17:1275-1285.
- Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S, Hallett M (2004) Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clin Neurophysiol* 115:2292-2307.
- Norman DA, Shallice T (1980) Attention to action. Willed and automatic control of behavior. University of California San Diego CHIP Report 99.
- Nowakowski R (2006) Stable neuron numbers from cradle to grave. *Proc Natl Acad Sci USA* 103:12219-12220.

- Ogawa S, Lee TM, Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA* 87:9868-9872.
- Oizumi M, Albantakis L, Tononi G (2014) From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. *PLOS Comput Biol* 10:e1003588.
- Ongur D, Price JL (2010) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex* 10:206-219.
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological Data. *Computational Intelligence and Neuroscience* 156869.
- Orsini C, Dankulov MM, Colomer-de-Simón P, Jamakovic A, Mahadevan P, Vahdat A, Bassler KE, Toroczkai Z, Boguñá M, Caldarelli G, Fortunato S, Krioukov D (2015) Quantifying randomness in real networks. *Nat Commun* 6:8627.
- Ortiz-Mantilla S, Hämäläinen JA, Benasich AA (2012) Time course of ERP generators to syllables in infants: a source localization study using age-appropriate brain templates. *Neuroimage* 59:3275-3287.
- Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD (2006) Detecting awareness in the vegetative state. *Science* 313:1402.
- Paasonen J, Stenroos P, Salo RA, Kiviniemi V, Gröhn O (2018) Functional connectivity under six anesthesia protocols and the awake condition in rat brain. *Neuroimage* 172:9-20.
- Park HJ, Friston K (2013) Structural and functional brain networks: from connections to cognition. *Science* 342:123841.
- Parkes L, Fulcher B, Yücel M, Fornito A (2018) An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *Neuroimage* 171:415-436.

Parvizi J, Van Hoesen GW, Buckwalter J, Damasio A (2006) Neural connections of the posteromedial cortex in the macaque. *Proc Natl Acad Sci USA* 103:1563-1568.

Pascual-Marqui RD (1999) Review of methods for solving the EEG inverse problem. *IJBEM* 1.

Pauling L, Coryell CD (1936) The magnetic properties and structure of hemoglobin, oxyhemoglobin and carbonmonoxyhemoglobin. *Proc Natl Acad Sci USA* 22:210-216.

Petanjek Z, Judaš M, Šimic G, Rasin MR, Uylings HB, Rakic P, Kostovic I (2011) Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proc Natl Acad Sci USA* 108:13281-13286.

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME (1988) Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331:585-589.

Pletikos M, Sousa AM, Sedmak G, Meyer KA, Zhu Y, Cheng F, Li M, Kawasawa YI, Sestan N (2014) Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron* 81:321-332.

Plourde G, Belin P, Chartrand D, Fiset P, Backman SB, Xie G, Zatorre RJ (2006) Cortical processing of complex auditory stimuli during alterations of consciousness with the general anesthetic propofol. *Anesthesiology* 104:448-457.

Poldrack RA, Mumford JA, Nichols TE (2011) *Handbook of functional MRI data analysis*. Cambridge University Press.

Posner MI, Snyder CRR (1975) Facilitation and inhibition in the processing of signals. In: *Attention and performance volume* (Rabbitt PMA, Dornic S, eds). *Attention and performance V*, Academic Press, New York.

Posner MI, Raichle ME (1994) *Images of mind*. New York: Scientific American Library.

Posner MI, Rothbart MK (1998) Attention, self-regulation and consciousness. *Philos Trans R Soc Lond B Biol Sci* 353:1915-1927.

- Power JD, Fair DA, Schlaggar BL, Petersen SE (2010) The development of human functional networks. *Neuron* 67:735-748.
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann, TO, Miezin FM, Schlaggar BL, Petersen SE (2011) Functional network organization of the human brain. *Neuron* 72:665-678.
- Preti MG, Bolton TA, Van De Ville D (2017) The dynamic functional connectome: state-of-the-art and perspectives. *Neuroimage* 160:41-54.
- Price CJ (2012) A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62:816-847.
- Puce A, Hämäläinen MS (2017) A review of issues related to data acquisition and analysis in EEG/MEG Studies. *Brain Sci* 7.
- Purdon PL, Pierce ET, Mukamel EA, Prerau MJ, Walsh JL, Wong KF, Salazar-Gomez AF, Harrell PG, Sampson AL, Cimenser A, Ching S, Kopell NJ, Tavares-Stoeckel C, Habeeb K, Merhar R, Brown EN (2013) Electroencephalogram signatures of loss and recovery of consciousness from propofol. *Proc Natl Acad Sci USA* 110:E1142-51.
- Rabinovich MI, Afraimovich VS, Bick C, Varona P (2012) Information flow dynamics in the brain. *Phys Life Rev* 9:51-73.
- Raichle ME (1998) Behind the scenes of functional brain imaging: a historical and physiological perspective. *Proc Natl Acad Sci USA* 95:765-772.
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. *Proc Natl Acad Sci USA* 98:676-682.
- Raichle ME (2010) Two views of brain function. *Trends Cogn Sci* 14:180-190.
- Raichle ME (2015) The brain's default mode network. *Annu Rev Neurosci* 38:433-447.
- Ramsay MA, Savege TM, Simpson BR, Goodwin R (1974) Controlled sedation with alphaxalone-alphadolone. *Br Med J* 2:656-659.

- Remnant P, Bennett J (1996) *Leibniz: new essays on human understanding*. Cambridge University Press.
- Reus and van den Heuvel (2013) The parcellation-based connectome: limitations and extensions. *NeuroImage* 80 (2013) 397-404
- Rochat P (2003) Five levels of self-awareness as they unfold early in life. *Conscious Cogn* 12:717-731.
- Ross S (1996) *Stochastic processes*. Wiley.
- Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52:1059-1069.
- Rubinov M, Sporns O, Thivierge JP, Breakspear M (2011) Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons. *PLOS Comput Biol* 7:e1002038.
- Rubinov M, Bullmore E (2013) Fledgling pathoconnectomics of psychiatric disorders. *Trends Cogn Sci* 17:641-647.
- Sanchez C, Richards JE, Almlí RC (2012) Neurodevelopmental MRI brain templates for children from 2 weeks to 4 years of age. *Dev Psychobiol* 54:77-91.
- Sarasso S, Boly M, Napolitani M, Gosseries O, Charland-Verville V, Casarotto S, Rosanova M, Casali AG, Brichant JF, Boveroux P, Rex S, Tononi G, Laureys S, Massimini M (2015) Consciousness and complexity during unresponsiveness induced by propofol, xenon, and ketamine. *Curr Biol* 25:3099-3105.
- Scannell JW, Blakemore C, Young MP (1995) Analysis of connectivity in the cat cerebral cortex. *J Neurosci* 15:1463-1483.
- Schiff ND, Ribary U, Moreno DR, Beattie B, Kronberg E, Blasberg R, Giacino J, McCagg C, Fins JJ, Llinás R, Plum F (2002) Residual cerebral activity and behavioural fragments can remain in the persistently vegetative brain. *Brain* 125:1210-1234.

- Schiff ND (2008) Central thalamic contributions to arousal regulation and neurological disorders of consciousness. *Ann NY Acad Sci* 1129:105-118.
- Schnakers C, Vanhaudenhuyse A, Giacino J, Ventura M, Boly M, Majerus S, Moonen G, Laureys S (2009) Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC Neurol* 9:35.
- Schönemann PH (1968) On two-sided orthogonal Procrustes problems. *Psychometrika* 33.
- Schrijver A (1987) *Theory of linear and integer programming*. John Wiley & Sons, Inc., New York.
- Schröter MS, Spormaker VI, Schorer A, Wohlschläger A, Czisch M, Kochs EF, Zimmer C, Hemmer B, Schneider G, Jordan D, Ilg R (2012) Spatiotemporal reconfiguration of large-scale brain functional networks during propofol-induced loss of consciousness. *J Neurosci* 32:12832-12840.
- Schrouff J, Perlberg V, Boly M, Marrelec G, Boveroux P, Vanhaudenhuyse A, Bruno M. A, Laureys S, Phillips C, Péligrini-Issac M, Maquet P, Benali H (2011) Brain functional integration decreases during propofol-induced loss of consciousness. *Neuroimage* 57:198-205.
- Sen PB, Basser PJ (2005) A model for diffusion in white matter in the brain. *Biophys J* 89: 2927-2938.
- Seth AK, Barrett AB, Barnett L (2011) Causal density and integrated information as measures of conscious level. *Philos Trans A Math Phys Eng Sci* 369:3748-3767.
- Shallice T (1972) Dual functions of consciousness. *Psychological Review* 79:383-393.
- Shallice T (1988) *From neuropsychology to mental structure*. Cambridge University Press, New York.
- Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* 27:379-423.

- Shannon CE, Weaver W (1949) The mathematical theory of communication. Urbana, IL: The University of Illinois Press. 1-117.
- Shew WL, Yang H, Yu S, Roy Rajarshi, Plenz D (2011) Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *J Neurosci* 31:55-63.
- Shew WL, Plenz D (2013) The functional benefits of criticality in the cortex. *Neuroscientist* 19:88-100.
- Shmuel A, Leopold DA (2008) Neuronal correlates of spontaneous fluctuations in fMRI signals in monkey visual cortex: Implications for functional connectivity at rest. *Hum Brain Mapp* 29:751-761.
- Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE (1997) Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J Cogn Neurosci* 9:648-663.
- Sihvola J (2007) The problem of consciousness in Aristotle's psychology. In: *Consciousness. Studies in the history of philosophy of mind* (Heinämaa S, Lähteenmäki V, Remes P, eds). Vol 4, Springer, Dordrecht.
- Sitt JD, King JR, El Karoui I, Rohaut B, Faugeras F, Gramfort A, Cohen L, Sigman M, Dehaene S, Naccache L (2014) Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain* 137:2258-2270.
- Sirotin YB, Das A (2009) Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature* 457:475-479.
- Sladky R, Friston KJ, Trostl J, Cunnington R, Moser E, Windischberger C (2011) Slice-timing effects and their correction in functional MRI. *Neuroimage* 58:588-594.
- Smallwood J, Brown K, Baird B, Schooler JW (2012) Cooperation between the default mode network and the frontal-parietal network in the production of an internal train of thought. *Brain Res* 1428:60-70.

- Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci USA* 106:13040-13045.
- Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen S, Power J, Salimi-Khorshidi G, Snyder AZ, Vu AT, Woolrich MW, Xu J, Yacoub E, Ugurbil K, Van Essen DC, Glasser MF, WU-Minn HCP Consortium (2013) Resting-state fMRI in the Human Connectome Project. *Neuroimage* 80:144-168.
- Smyser CD, Inder TE, Shimony JS, Hill JE, Degnan AJ, Snyder AZ, Neil JJ (2010) Longitudinal analysis of neural network development in preterm infants. *Cerebral Cortex* 20:2852-286.
- Soddu A, Vanhauzenhuysse A, Demertzi A, Marie-Aurelie B, Tshibanda L, Di H, Melanie B, Papa M, Laureys S, Noirhomme Q (2011) Resting state activity in patients with disorders of consciousness. *Funct Neurol* 26:37-43.
- Sotiropoulos SN, Jbabdi S, Xu J, Andersson JL, Moeller S, Auerbach EJ, Glasser MF, Hernandez M, Sapiro G, Jenkinson M, Feinberg DA, Yacoub E, Lenglet C, Van Essen DC, Ugurbil K, Behrens TE, WU-Minn HCP Consortium (2013) Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *Neuroimage* 80:125-143.
- Sporns O, Tononi G, Edelman GM (2000) Connectivity and complexity: the relationship between neuroanatomy and brain dynamics. *Neural Netw* 13:909-922.
- Sporns O, Tononi G (2001) Classes of network connectivity and dynamics. *Complexity* 7:28-38.
- Sporns O (2013) Structure and function of complex brain networks. *Dialogues Clin Neurosci* 15:247-262.
- Spreng NR (2012) The fallacy of a "task-negative" network. *Front Psychol* 3:145.

- Spruston N (2008) Pyramidal neurons: dendritic structure and synaptic integration. *Nat Rev Neurosci* 9:206-221.
- Stam CJ, Nolte G, Daffertshofer A (2007) Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum Brain Mapp* 28:1178-1193.
- Stamatakis EA, Adapa RM, Absalom AR, Menon DK (2010) Changes in resting neural connectivity during propofol sedation. *PLOS ONE* 5:e1424.
- Stamatakis EA, Orfanidou E, Papanicolaou AC (2017) Functional Magnetic Resonance Imaging. In: *The Oxford handbook of functional brain imaging in neuropsychology and cognitive neurosciences* (Papanikolaou AC, ed). Oxford University Press.
- Steyn-Ross ML, Steyn-Ross DA, Sleight JW, Wilcocks LC (2001) Toward a theory of the general-anesthetic-induced phase transition of the cerebral cortex. I. A thermodynamics analogy. *Phys Rev E Stat Nonlin Soft Matter Phys* 64:011917.
- Tagliazucchi, Chialvo DR (2012) Brain complexity born out of criticality. <https://arxiv.org/pdf/1211.0309.pdf>.
- Tagliazucchi E, Carhart-Harris R, Leech R, Nutt D, Chialvo DR (2014) Enhanced repertoire of brain dynamical states during the psychedelic experience. *Hum Brain Mapp* 35:5442-5456.
- Tagliazucchi E, Chialvo DR, Siniatchkin M, Amico E, Brichant JF, Bonhomme V, Noirhomme Q, Laufs H, Laureys S (2016) Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics. *J R Soc Interface* 13:20151027.
- Talairach J, Tournoux P (1988) *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. Stuttgart, New York: Georg Thieme.
- Tang YY, Tang Y, Tang R, Lewis-Peacock JA (2017) Brief mental training reorganizes large-scale brain networks. *Front Syst Neurosci* 11:6.

- Tavor I, Parker Jones O, Mars RB, Smith SM, Behrens TE, Jbabdi S (2016) Task-free MRI predicts individual differences in brain activity during task performance. *Science* 352:216-220.
- Thomason ME, Grove LE, Lozon TA Jr, Vila AM, Ye Y, Nye MJ, Manning JH, Pappas A, Hernandez-Andrade E, Yeo L, Mody S, Berman S, Hassan SS, Romero R (2015) Age-related increases in long-range connectivity in fetal functional neural connectivity networks in utero. *Dev Cogn Neurosci* 11:96-104.
- Tononi G, Sporns O, Edelman GM (1994) A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci USA* 91: 5033-5037.
- Tononi G, Edelman GM, Sporns O (1998) Complexity and coherency: integrating information in the brain. *Trends Cogn Sci* 2:474-484.
- Tononi G, Edelman GM (1998) Consciousness and complexity. *Science* 282:1846-1851.
- Tononi G (2004) An information integration theory of consciousness. *BMC Neurosci* 5:42.
- Tononi G (2008) Consciousness as integrated information: a provisional manifesto. *Biol Bull* 215:216-242.
- Tononi G, Koch C (2015) Consciousness: here, there, and everywhere? *Philos Trans R Soc Lond B Biol Sci* 370:1668.
- Tuch DS, Reese TG, Wiegell MR, Makris N, Belliveau JW, Wedeen VJ (2002) High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn Reson Med* 48:577-582.
- Tuch DS, Reese TG, Wiegell MR, Wedeen VJ (2003) Diffusion MRI of complex neural architecture. *Neuron* 40:885-895.
- Tuch DS (2004) Q-ball imaging. *Magnetic Resonance in Medicine* 52:1358-1372.

- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273-289.
- Uğurbil K, Xu J, Auerbach EJ, Moeller S, Vu AT, Duarte-Carvajalino JM, Lenglet C, Wu X, Schmitter S, Van de Moortele PF, Strupp J, Sapiro G, De Martino F, Wang D, Harel N, Garwood M, Chen L, Feinberg DA, Smith SM, Miller KL, Sotiropoulos SN, Jbabdi S, Andersson JL, Behrens TE, Glasser MF, Van Essen DC, Yacoub E, WU-Minn HCP Consortium (2013) Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *Neuroimage* 80:80-104.
- Uhlhaas PJ, Haenschel C, Nikolić D, Singer W (2008) The role of oscillations and synchrony in cortical networks and their putative relevance for the pathophysiology of schizophrenia. *Schizophr Bull* 34:927-943.
- Vacas S, Kurien P, Maze M (2013) Sleep and anesthesia-common mechanisms of action. *Sleep Med Clin* 8:1-9.
- Valiant P, Valiant G (2013) Estimating the unseen: improved estimators for entropy and other properties. In: *Advances in Neural Information Processing Systems* 26, (NIPS).
- Van de Ville D, Britz J, Michel CM (2010) EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proc Natl Acad Sci USA* 107:18179-18184.
- van den Heuvel MP, Sporns O (2013) Network hubs in the human brain. *Trends Cogn Sci* 12:683-696.
- van den Heuvel MI, Thomason ME (2016) Functional connectivity of the human brain in utero. *Trends Cogn Sci* 20: 931-939.
- van den Heuvel MP, de Lange SC, Zalesky A, Seguin C, Yeo BTT, Schmidt R (2017) Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations. *Neuroimage* 152:437-449.

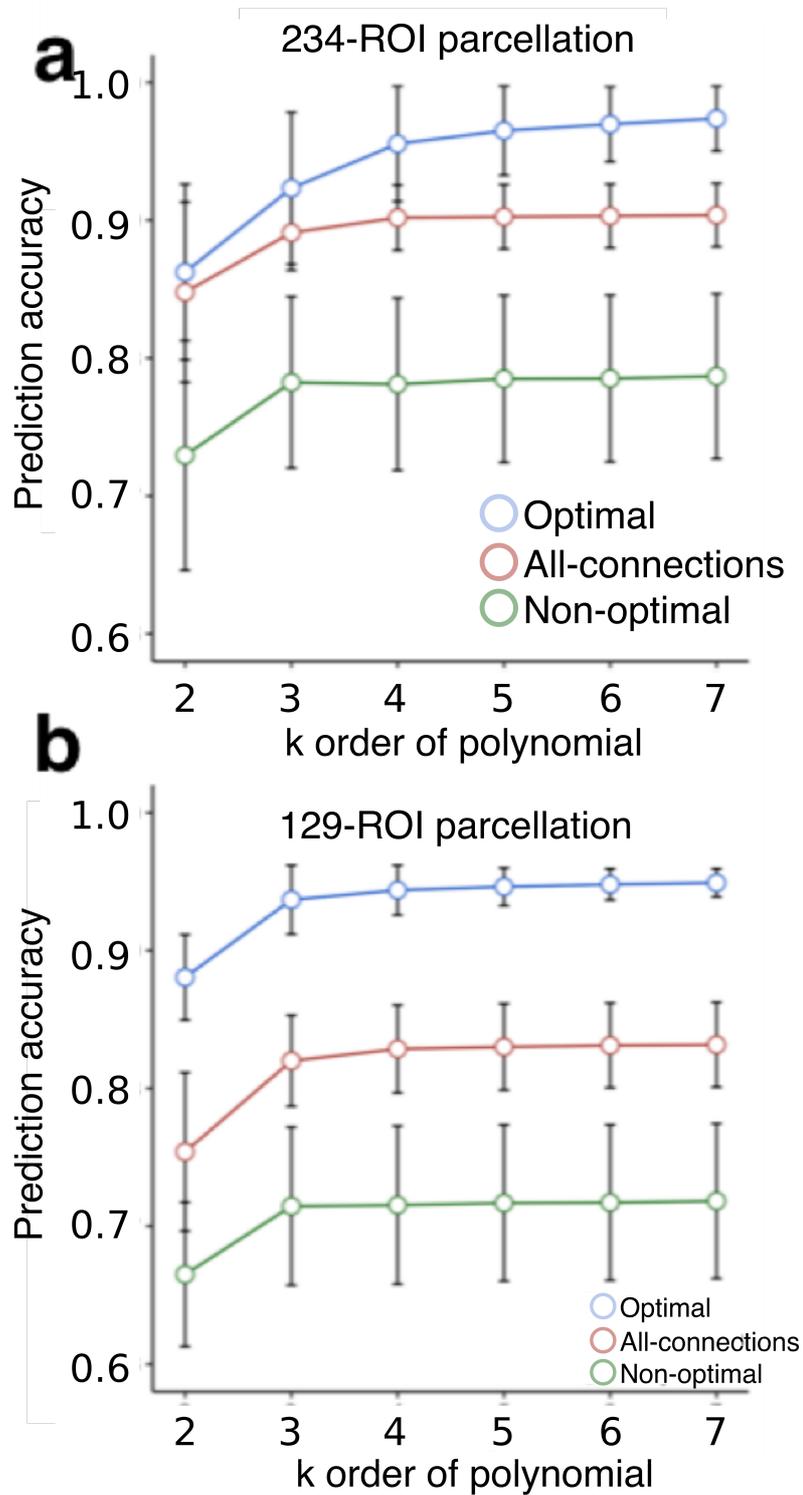
- van Emden MH (1975) An analysis of complexity. Mathematisch Centrum.
- Van Essen DC (1997) A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature* 385:313-318.
- Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, and the WU-Minn HCP Consortium (2013) The WU-Minn Human Connectome Project: an overview. *Neuroimage* 80:62-79.
- Vanhaudenhuyse A, Noirhomme Q, Tshibanda LJ, Bruno MA, Boveroux P, Schnakers C, Soddu A, Perlberg V, Ledoux D, Brichant JF, Moonen G, Maquet P, Greicius MD, Laureys S, Boly M (2010) Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. *Brain* 133:161-171.
- Váša F, Shanahan M, Hellyer PJ, Scott G, Cabral J, Leech R (2015) Effects of lesions on synchrony and metastability in cortical networks. *Neuroimage* 118:456-467.
- Van Veen BD, Van Drongelen W, Yuchtman M, Suzuki A (1997) Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng* 44:867-880.
- van Vugt MK, Sederberg PB, Kahanac MJ (2007) Comparison of spectral analysis methods for characterizing brain oscillations *J Neurosci Methods* 162:49-63.
- Vatansever D, Menon DK, Manktelow AE, Sahakian BJ, Stamatakis EA (2015a) Default mode network connectivity during task execution. *Neuroimage* 122:96-104.
- Vatansever D, Menon DK, Manktelow AE, Sahakian BJ, Stamatakis EA (2015b) Default mode dynamics for global functional integration. *J Neurosci* 35:15254-15262.
- Vatansever D, Menon DK, Stamatakis EA (2017) Default mode contributions to automated information processing. *Proc Natl Acad Sci USA* 114:12821-12826.
- Vértes PE, Alexander-Bloch AF, Gogtay N, Giedd JN, Rapoport JL, Bullmore ET (2012) Simple models of human brain functional networks. *Proc Natl Acad Sci USA* 109:5868-5873.

- Vértés PE, Bullmore ET (2015) Annual research review: growth connectomics--the organization and reorganization of brain networks during normal and abnormal development. *J Child Psychol Psychiatry* 56:299-320.
- Vidaurre D, Smith SM, Woolrich MW (2017a) Brain network dynamics are hierarchically organized in time. *Proc Natl Acad Sci USA* 114:12827-12832.
- Vidaurre D, Abeysuriya R, Becker R, Quinn AJ, Alfaro-Almagro F, Smith SM, Woolrich MW (2017b) Discovering dynamic brain networks from big data in rest and task. *Neuroimage* pii: S1053-8119:30548-30557.
- Vincent JL, Patel GH, Fox MD, Snyder AZ, Baker JT, Van Essen DC, Zempel JM, Snyder LH, Corbetta M, Raichle ME (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447:83-86.
- Vinck M, Oostenveld R, van Wingerden M, Battaglia F, Pennartz CM (2011) An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *Neuroimage* 55:1548-1565.
- Xie G, Deschamps A, Backman SB, Fiset P, Chartrand D, Dagher A, Plourde G (2011) Critical involvement of the thalamus and precuneus during restoration of consciousness with physostigmine in humans during propofol anaesthesia: a positron emission tomography study. *Br J Anaesth* 106:548-557.
- Yeh FC, Van Jay W, Wen-Yih IT (2010) Generalized q-sampling imaging. *Medical Imaging, IEEE Transactions* 29:1626-1635.
- Yeh FC, Verstynen TD, Wang Y, Fernández-Miranda JC, Tseng WYI (2013) Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLOS ONE* 8:e80713.
- Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125-1165.

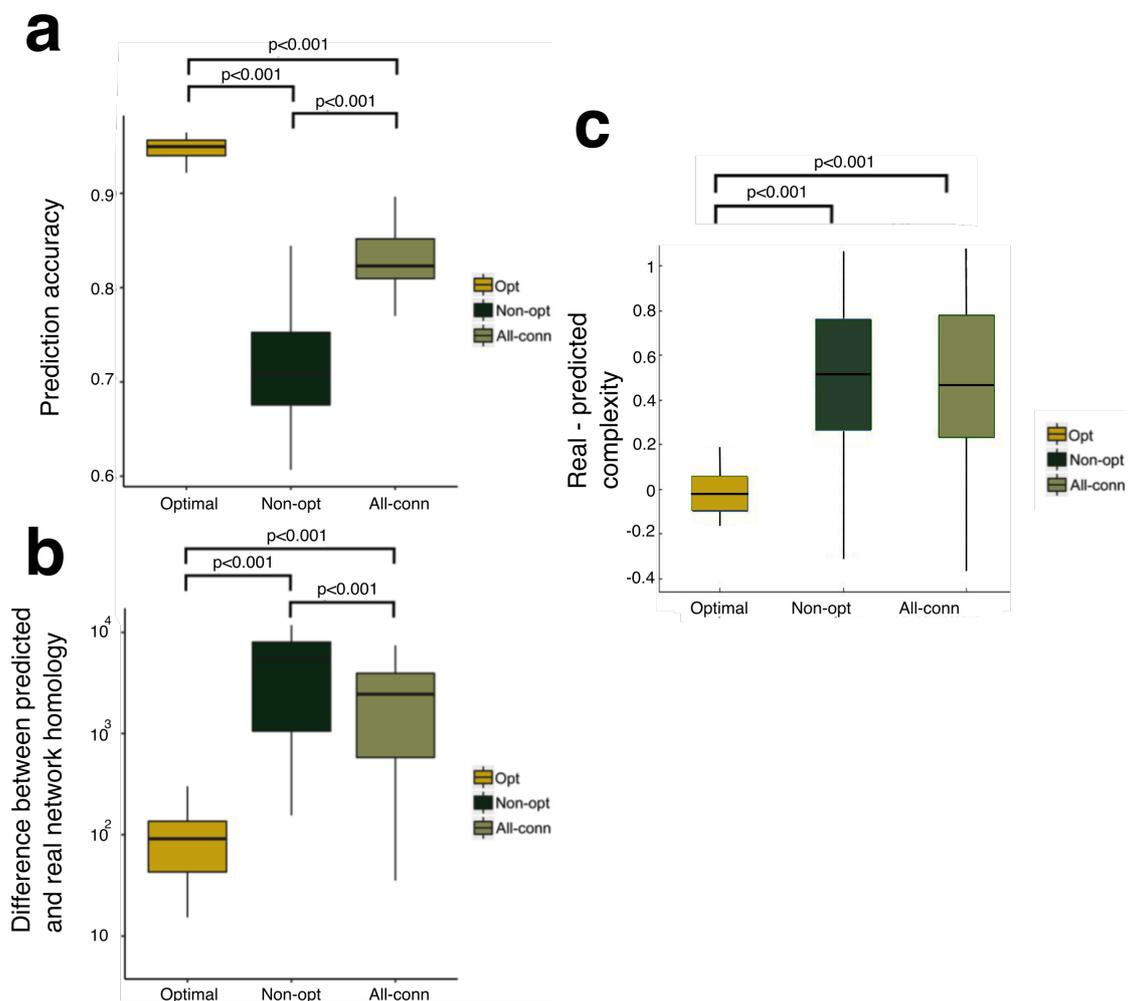
- 
- Wainwright MJ, Jordan MI (2008) Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning: Vol 1: No. 1-2:1-305*.
- Warnaby CE, Sleigh JW, Hight D, Jbabdi S, Tracey I (2017) Investigation of slow-wave activity saturation during surgical anesthesia reveals a signature of neural inertia in humans. *Anesthesiology* 127:645-657.
- Watts DJ, Strogatz, SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393:440-442.
- Wedeen VJ, Wang RP, Schmahmann JD, Benner T, Tseng WY, Dai G, Pandya DN, Hagmann P, D'Arceuil H, de Crespigny AJ (2008) Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage* 41:1267-1277.
- Wells WM, 3rd, Viola P, Atsumi H, Nakajima S, Kikinis R (1996) Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1:35-51.
- Whitfield-Gabrieli S, Nieto-Castanon A (2012) Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect* 2:125-141.
- Wu YC, Alexander AL (2007) Hybrid diffusion imaging. *Neuroimage* 36:617-629.
- Zamora-López G, Chen Y, Deco G, Kringelbach ML, Zhou C (2016) Functional complexity emerging from anatomical constraints in the brain: the significance of network modularity and rich-clubs. *Sci Rep* 6:38424.
- Zhao K, Halu A, Severini S, Bianconi G (2011) Entropy rate of nonequilibrium growing networks. *Physical Review E* 84:066113.

## Appendix 1

## Chapter 3



**Appendix Figure 3.1 - Predicting functional connectivity from structural connectivity matrices using different orders of polynomial transformation.** The prediction algorithm uses a polynomial transformation (of order  $k$ ) of the structural connectivity matrix to predict the functional connectivity matrix. (a) shows the prediction accuracy, as expressed by correlation of the predicted and real functional connectivity matrices, for the 234-ROI parcellation across different values of  $k$ . We observed that above  $k = 5$  the results plateaued showing that incorporating higher powers of polynomial transformation did not contribute further to the prediction. (b) shows similar results for the 129-ROI parcellation. Circles show average over  $n=50$  subjects and lines show  $\pm$  standard error of the mean.



**Appendix Figure 3.2 - Predicting functional connectivity from structural connectivity matrices at the 129-ROI resolution.** (a) Each box plot shows the variation of the prediction accuracy across  $n=50$  subjects as expressed in terms of correlation between the real and predicted functional connectivity matrices (one-way repeated-measures ANOVA  $F(2,98)=837$ ,  $p < 0.0001$  Bonferroni corrected multiple comparisons test, bayes factor= $8.0399e-30$ , eta squared= $0.9447$ ). The predicted functional connectivity matrix was obtained using a polynomial order of  $k = 5$  of the structural connectivity matrix. Higher

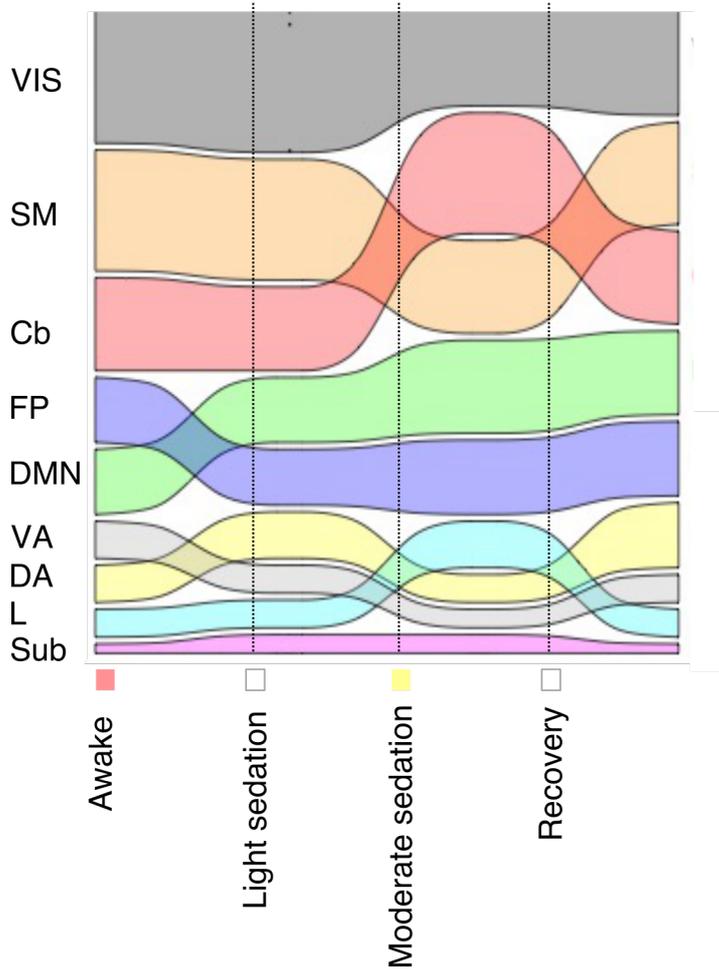
scores show higher correlations between the real and the predicted functional connectivity matrices. (b) Here I used methods from homology theory to assess the similarities between predicted and real functional connectivity matrices (one-way repeated-measures ANOVA  $F(2,98)=82.16$ ,  $p<0.0001$ , Bonferroni corrected multiple comparisons test). Lower scores reflect a smaller difference between the real and predicted functional connectivity matrices. Both prediction scores show that using only optimal connections was significantly more predictive of whole-brain functional connectivity than using all structural or non-optimal connections. Boxplots' thick lines show the median value while whiskers reflect the maximum and minimum values of the data. (c) Difference between predicted complexity and complexity obtained from real functional connectivity matrices using optimal, non-optimal, and whole-brain connections. Complexity was obtained as the entropy of degree distribution of functional connectivity networks. I observed that optimal connections could retrieve real complexity significantly better than non-optimal and whole-brain connections (ANOVA  $F(2,98)=44.84$ ,  $p<0.0001$ , Bonferroni corrected multiple comparisons test). The smaller the value to 0 the better the prediction of complexity is. Boxplots' thick lines show the median value while whiskers reflect the maximum and minimum values of the data.

**Appendix Table 3.1 - Optimality statistics for the structural networks at the 129-ROI resolution.**

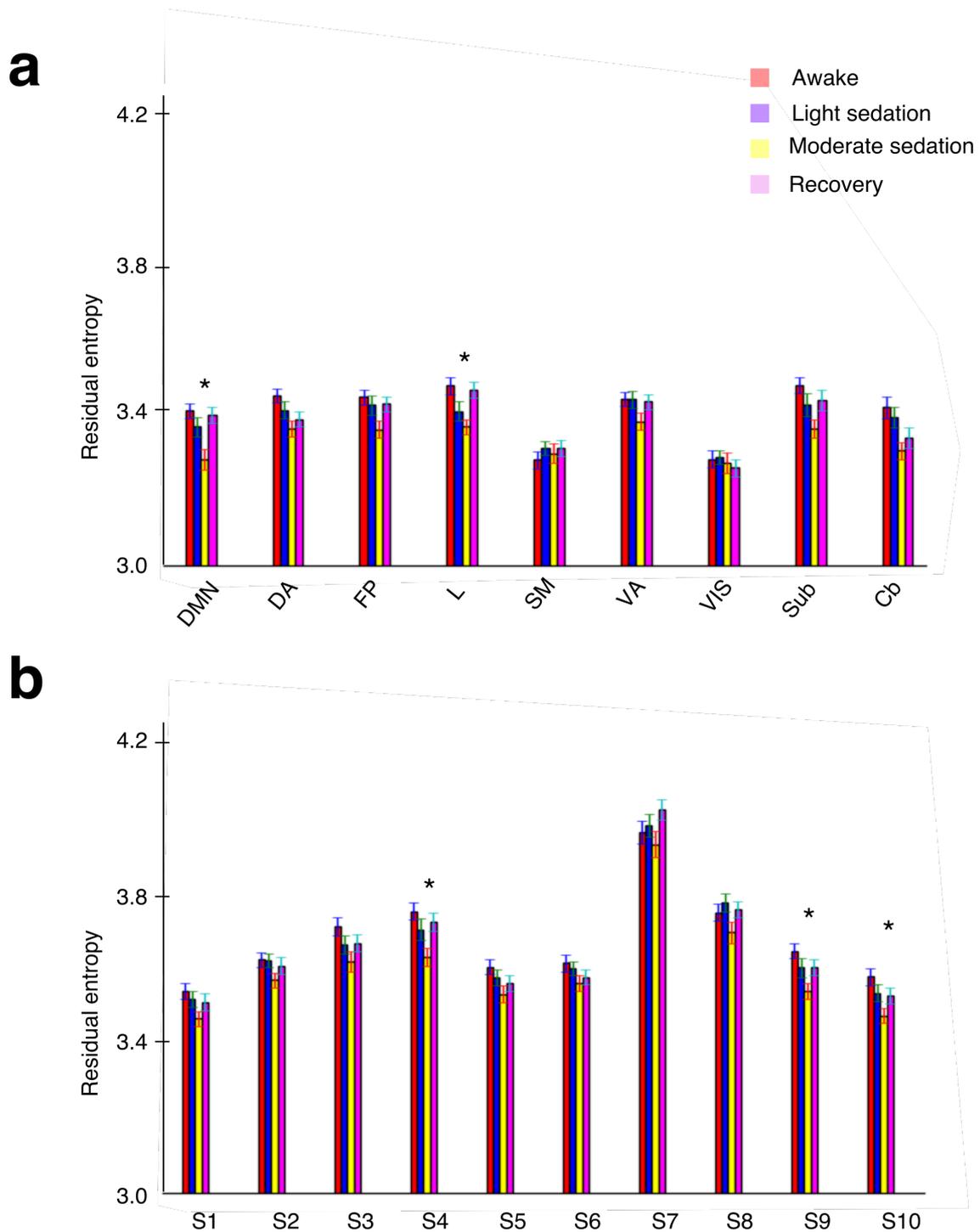
Optimality was calculated by counting the number of optimal edges (connections that existed in both the NNG and real networks  $|T|$ ) divided by the number of total edges the NNG model ( $|M|$ ) produced. False positives remained relatively low compared to the total number of edges  $|M|$  of the NNG model. Results are presented in the form of mean (stdev) over  $n=50$  subjects.

|                       | Edge Statistics |
|-----------------------|-----------------|
| Number of nodes       | 129             |
| Real edges $ R $      | 1725.6 (255.3)  |
| NNG edges $ M $       | 334.1 (5.4)     |
| Optimal edges $ T $   | 239.1 (14.7)    |
| False positives $ F $ | 93.8 (15.4)     |
| Optimality $ T / M $  | 0.72 (0.04)     |

## Chapter 4

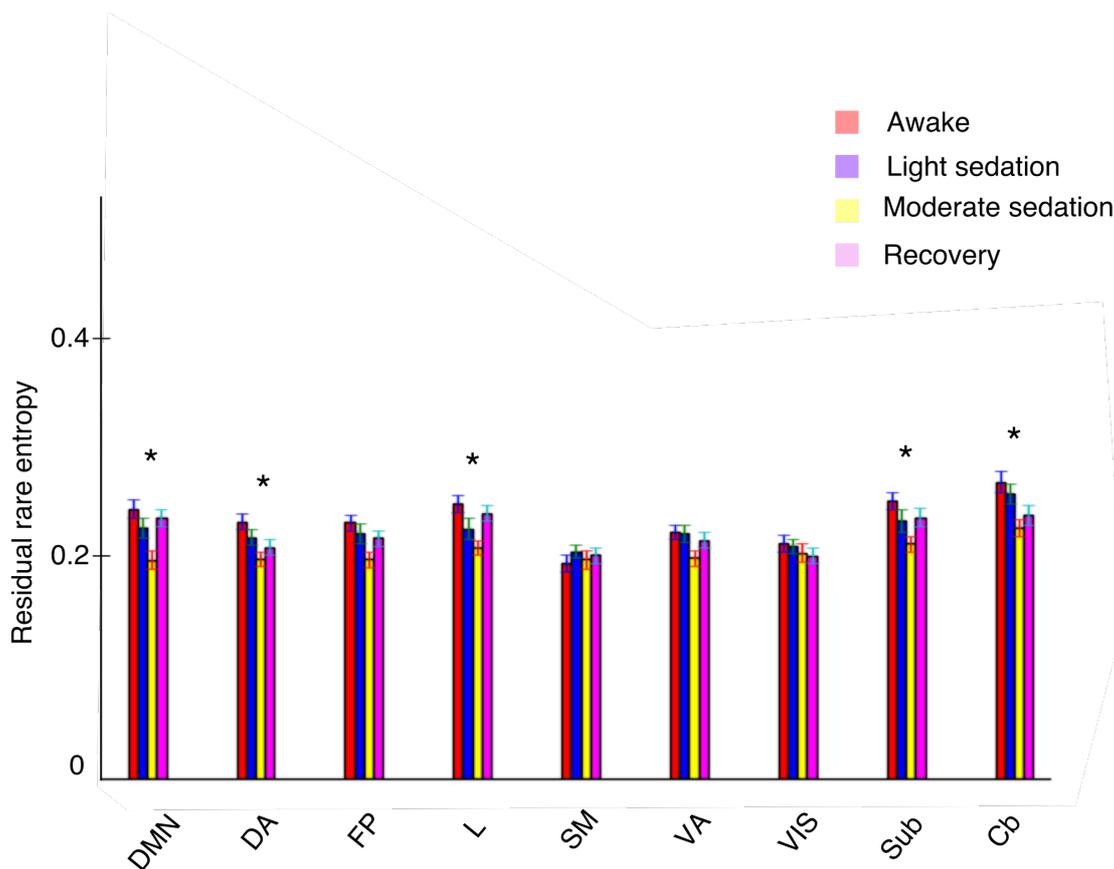


**Appendix Figure 4.1 - Alterations in sparsely connected regions' connectivity during sedation.** The alluvial diagram shows changes in the median over n=25 subjects amount (width of colour band) of participation of the endpoints of sparsely connected regions in different LSNs with increasing sedation.



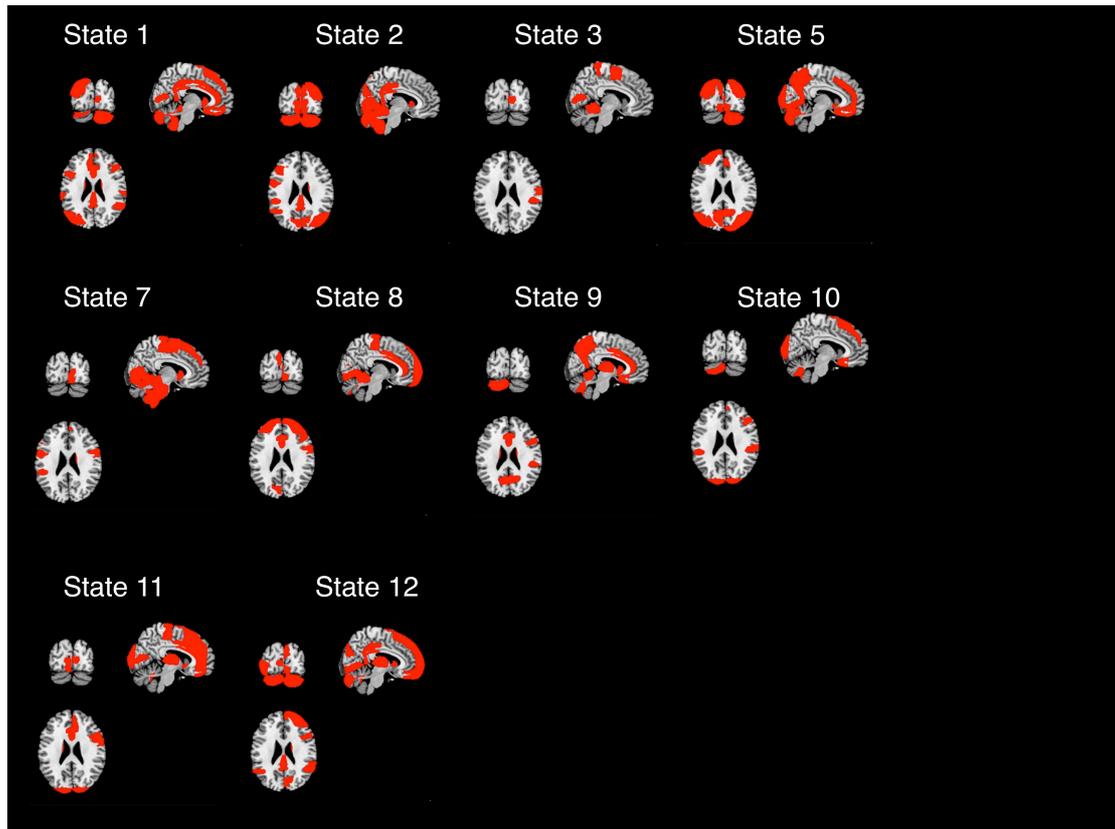
**Appendix Figure 4.2 - Residual entropy during moderate sedation.** Part (a) shows residual entropy that was calculated by virtually removing from the whole-brain network 7 pre-defined networks (and 2 manually defined subcortical and cerebellar networks) coming from Yeo et al., (2011) and calculating global entropies in the respective residual networks. Barplots show mean values and error bars show the standard error of the mean for  $n=25$  subjects. Asterisk (\*) implies significance  $p<0.05$  that survived Bonferroni correction within sedation levels and across 9 networks. *Abbreviations:* DMN: Default-

Mode Network, DA: Dorsal Attention Network, FP: Fronto-Parietal Network, L: Limbic Network, SM: Somato-Motor Network, VA: Ventral Attention Network, VIS: Visual Network, Sub: Subcortical, Cb: Cerebellum. Part (b) shows residual entropy that was calculated by removing 10 pre-defined networks from Smith et al. Barplots show mean values and error bars show the standard error of the mean. Asterisk (\*) implies significance  $p < 0.05$  that survived Bonferroni correction within sedation levels and across 10 networks. Abbreviations stand for the Smith et al. (2009) networks with the order presented in the Smith et al. (2009) work.

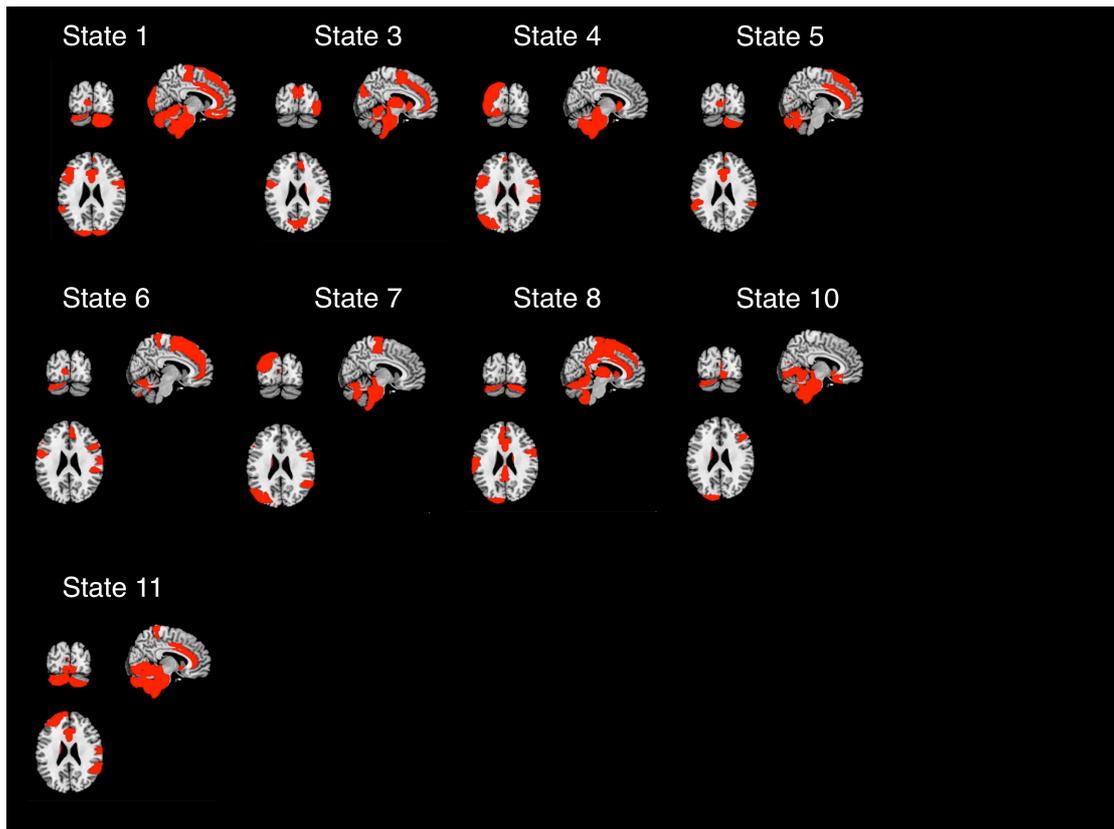


**Appendix Figure 4.3 - Residual rare entropy during moderate sedation.** Figure shows residual rare entropy that was calculated by removing from the whole-brain network 7 pre-defined networks coming from Yeo et al. (2011) (and 2 manually defined subcortical and cerebellar networks) and calculating rare entropy in the respective lesioned networks. Barplots show mean values and error bars show the standard error of the mean for  $n=25$  subjects. Asterisk (\*) implies significance  $p < 0.05$  that survived Bonferroni correction within sedation levels and across 9 networks. *Abbreviations:* DMN: Default-Mode Network, DA: Dorsal Attention Network, FP: Fronto-Parietal Network, L: Limbic Network, SM: Somato-Motor Network, VA: Ventral Attention Network, VIS: Visual, Sub: Subcortical, Cb: Cerebellum.

## Chapter 5

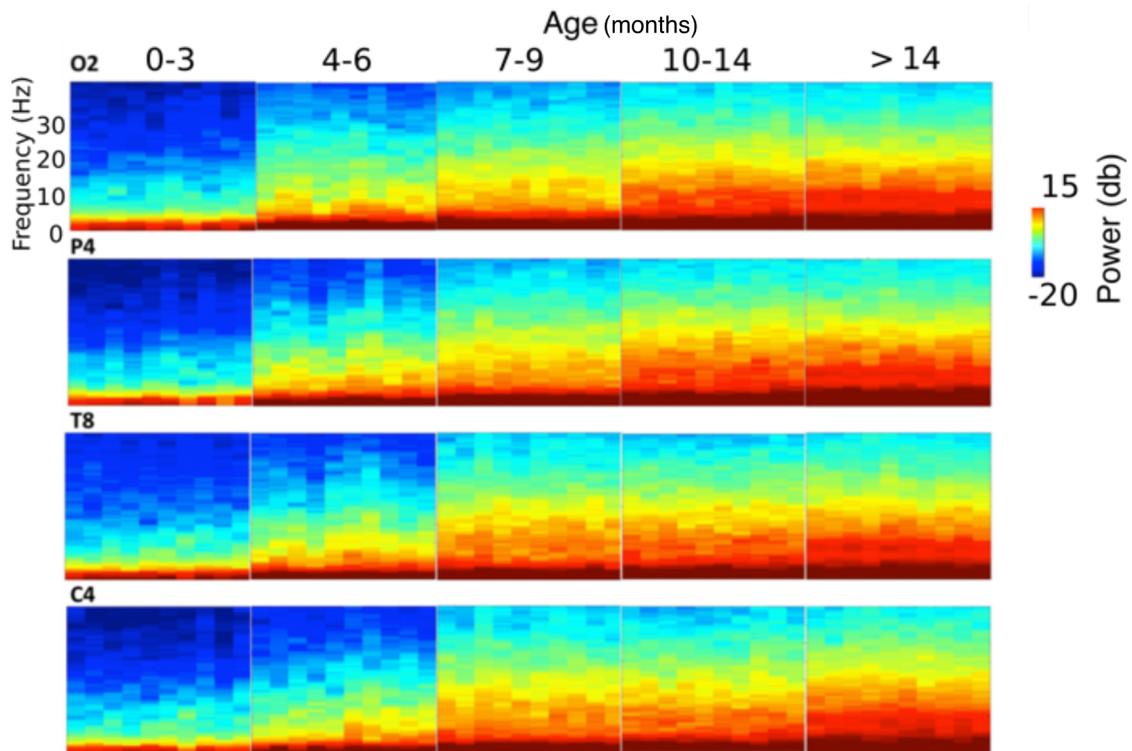


**Appendix Figure 5.1 - Brain states extracted from the Markov model for the sedation experiment.** Each panel shows slices at  $x=98$   $y=41$   $z=97$  of the states identified from the HMM output applied on the sedation data.



**Appendix Figure 5.2 - Brain states extracted from the Markov model for the DOC experiment.** Each panel shows slices at  $x=98$   $y=41$   $z=97$  of the states identified from the HMM output applied on the DOC data.

## Chapter 6



**Appendix Figure 6.1 - Spectral power analysis for infant brain dynamics under anaesthesia.** Median (over all infants) spectrograms show the power in each frequency as a function of age during MOSSA for four different channels O2, P4, T8, C4.