
Electronic Thesis and Dissertation Repository

4-22-2022 2:00 PM

An Implementation of Integrated Information Theory in Resting-State fMRI

Idan E. Nemirovsky, *The University of Western Ontario*

Supervisor: Dr. Andrea Soddu, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Physics

© Idan E. Nemirovsky 2022

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Anesthesia and Analgesia Commons](#), [Cognitive Neuroscience Commons](#), [Computational Neuroscience Commons](#), and the [Discrete Mathematics and Combinatorics Commons](#)

Recommended Citation

Nemirovsky, Idan E., "An Implementation of Integrated Information Theory in Resting-State fMRI" (2022). *Electronic Thesis and Dissertation Repository*. 8516.
<https://ir.lib.uwo.ca/etd/8516>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Integrated Information Theory (IIT) is a framework developed to explain consciousness, arguing that conscious systems consist of interacting elements that are integrated through their causal properties. In this study, we present the first application of IIT to functional magnetic resonance imaging (fMRI) data and investigate whether its principal metric, Φ , can meaningfully quantify resting-state cortical activity patterns. Data was acquired from 17 healthy subjects who underwent sedation with propofol, a short acting anesthetic. Using PyPhi, a software package developed for IIT, we thoroughly analyze how Φ varies across different networks and throughout sedation. Our findings indicate that variations in Φ closely reflect the conscious state of patients in the frontoparietal and dorsal attention networks, which are responsible for higher-order cognitive functions. Despite existing limitations in applying IIT to empirical data, we obtained promising results that merit further applications of this framework to fMRI.

Keywords

Integrated information theory (IIT), consciousness, neural correlates of consciousness, causality, functional magnetic resonance imaging (fMRI), resting-state networks, sedation, propofol

Summary for Lay Audience

Consciousness is a highly debated subject in neuroscience, and many different theories have been proposed to explain it. In this area of research, Integrated Information Theory (IIT) has emerged as a leading framework. IIT argues that a system, such as the brain, generates “consciousness” through the integration of its elements, which can be computed using the causes and effects associated with the system’s interactions. The principal metric of this framework is integrated information, or Φ , which measures the extent of a system’s integration. Although promising developments have been made so far, studies applying IIT to empirical neuroimaging data are limited. In this work, we provide a seminal application of this framework to resting-state functional magnetic resonance imaging data (fMRI). fMRI is a technique that allows for analysis of brain activity patterns over space and time. “Resting-state” (also known as “task-negative”) acquisition describes measurements of spontaneous brain activity patterns when subjects are not involved in a particular task. Studies employing this technique have identified a series of resting-state networks (RSNs), which are collections of correlated regions associated with the brain’s functions at rest. We apply our analysis to several RSNs acquired from subjects who underwent sedation with propofol, a short acting anesthetic. To test whether Φ is a valid marker of consciousness, we thoroughly analyzed how it varies throughout the different states of awareness induced by the sedative. In our discussions, we relate our results to existing literature on these networks and how they are affected by propofol. Most notably, we found that Φ corresponds to the conscious state of subjects in higher-order networks of the brain that are responsible for awareness and perception. Altogether, our implementation presents a promising procedure for computing integrated information from fMRI data. Ultimately, our goal is to provide a foundation for future neuroimaging studies that apply IIT to better understand neurological disorders and other states of altered consciousness.

Co-Authorship Statement

Although not published at the time of writing, the material in Chapters 2 and 3 is primarily based on the following manuscript:

I.E. Nemirovsky, N.J.M Popiel, J. Rudas, M. Caius, L.Naci, N.D. Schiff, A.M. Owen, A. Soddu, An Implementation of Integrated Information Theory in Resting-State fMRI. (2022).

Our group's initial attempts to quantify integrated information for resting-state fMRI were made by Nicholas J.M. Popiel and Jorge Rudas, who also made significant contributions to developing a neuroimaging pipeline that was used to process the data acquired in this study.

fMRI data was acquired from a previous study conducted by Dr. Lorina Naci and Dr. Adrian M. Owen. The methodology for fMRI acquisition is provided in the supplementary material (Chapter 6) and is primarily based on the following manuscripts:

Naci, L., Haugg, A., MacDonald, A., Anello, M., Houldin, E., Naqshbandi, S., Gonzalez-Lara, L. E., Arango, M., Harle, C., Cusack, R., & Owen, A. M. (2018). Functional diversity of brain networks supports consciousness and verbal intelligence. *Scientific Reports*, 8(1), 13259.

Kandeeban, S., Rudas, J., Gomez, F., Stojanoski, B., Valluri, S., Owen, A. M., Naci, L., Nichols, E. S., & Soddu, A. (2020). Modeling an auditory stimulated brain under altered states of consciousness using the generalized Ising model. *NeuroImage*, 223, 117367.

The material in Chapter 4 on the Generalized Ising Model is based on simulations ran by Matthew Caius, who has also worked with our group on integrated information theory.

Finally, the first chapter gives background information on disorders of consciousness. Although it is not the main focus of this thesis, part of the research conducted for this master's involved writing a review on how these disorders are diagnosed, which I published with shared first-authorship.

Porcaro, C., Nemirovsky, I. E., Riganello, F., Mansour, Z., Cerasa, A., Tonin, P., Stojanoski, B., & Soddu, A. (2022). Diagnostic Developments in Differentiating Unresponsive Wakefulness Syndrome and the Minimally Conscious State. *Frontiers in Neurology*, 12.

Acknowledgments

I would like to begin by thanking my supervisor, Professor Andrea Soddu, for providing me with an incredible research experience. All the progress I made would not have been possible without your encouragement, and I am extremely thankful to have learned from your years of experience as a physicist and neuroscientist.

I would also like to express my appreciation for Western University's Physics and Astronomy department. Everyone here, including students, staff, and faculty, strives to make this an inclusive and welcoming environment, and I will forever be thankful for the people I had the chance to meet and work with.

Having completed my master's during the COVID-19 pandemic, there is no denying that circumstances related to the virus have made these last two years challenging. However, I am grateful to have had a loving family and group of friends who helped make these months pass by easier.

To my parents, thank you for your unconditional love and support. You have always believed in me, even at times when I doubted myself, and I owe all my accomplishments to how you raised me.

Table of Contents

Abstract.....	ii
Summary for Lay Audience	iii
Co-Authorship Statement	iv
Acknowledgments	v
Table of Contents.....	vi
List of Tables	ix
Chapter 1 : Introduction and Background	1
1.1. Overview	1
1.2. Background on Neuroscience and Neuroimaging	3
1.2.1. The Nervous System.....	3
1.2.2. Functional Magnetic Resonance Imaging	4
1.2.3. Anesthesia with Propofol.....	8
1.3. Integrated Information Theory	10
1.3.1. Defining Consciousness	10
1.3.2. Quantifying Integrated Information for a Single Mechanism	14
1.3.3. Quantifying Integrated Information for a System of Mechanisms.....	19
1.4. References	22
Chapter 2 : Integrated Information in Resting-State Networks.....	26
2.1. Overview	26
2.2. Methods	27
2.2.1. Obtaining binary signals and transition probability matrices.....	27
2.2.2. Concatenating time-series.....	29
2.2.3. Statistical test for the Markov property	30
2.2.4. Test for conditional independence test	33

2.2.5. Calculating $\mu[\Phi^{\max}]$ for a time-series	33
2.2.6. Analysis of modulation by propofol.....	35
2.2.7. Interregional correlations.....	36
2.2.8. Spatial and temporal control procedures	36
2.3. Results	37
2.3.1. Markov property and conditional independence	37
2.3.2. Modulation of $\mu[\Phi^{\max}]$ by propofol.....	38
2.3.3. Spatial and temporal control procedures	41
2.4. Discussion.....	44
2.5. References	47
Chapter 3 : Resting State Networks and their Temporal Structure	49
3.1. Overview	49
3.2. Methods	50
3.2.1. Grouped permutations	50
3.2.2. Diagonal strength and diagonally selective permutations	52
3.3. Results	55
3.3.1. Grouped permutations	55
3.3.2. Diagonal strength and diagonally selective permutations	60
3.4. Discussion.....	63
3.5. References	66
Chapter 4 : Integrated Information and the Length of a Time-Series	67
4.1. Overview	67
4.2. Methods	68
4.2.1. Concatenations and their effects on $\mu[\Phi^{\max}]$	68
4.2.2. The Generalized Ising Model	68
4.2.3. $\mu[\Phi^{\max}]$ and the length of a simulated time-series.....	70

4.3. Results	71
4.3.1. Concatenations and their effects on $\mu[\Phi^{\max}]$	71
4.3.2. Simulated and empirical time-series.....	72
4.4. Discussion.....	73
4.5. References	75
Chapter 5 : Summary, Conclusions, and Future Work.....	76
5.1. Summary.....	76
5.2. Conclusions and Future Work	78
5.3. References	79
Chapter 6 : Supplementary Material.....	80
6.1. Propofol Administration and fMRI Acquisition.....	80
6.1.1. Participants and Ethics.....	80
6.1.2. Administration of Propofol.....	80
6.1.3. fMRI Acquisition and Neuroimaging Data Preprocessing.....	82
6.1.4. Extraction of Representative Regions from RSNs	82
6.2. Propofol Administration and fMRI Acquisition.....	85
6.2.1. Markov Property Test.....	85
6.2.2. Modulation of $\mu[\Phi^{\max}]$ by Propofol.....	86
6.3. References	85
Curriculum Vitae	86

List of Tables

Table 1.2.1. Descriptions of higher-order cortical networks	8
Table 1.2.2. Descriptions of motor and sensory cortical networks	8
Table 2.2.1. Contingency table for counting the sequences of states in a time-series	32
Table 2.3.1. Relative distance between original TPMs and their conditionally independent versions	37
Table 6.2.1. Results for the Markov property contingency table test.....	86
Table 6.2.2. Results for statistical comparison of $\mu[\Phi^{\max}]$ over the four states of awareness in each of the 11 RSNs	84

List of Figures

Figure 1.2.1. Fundamental structures of the nervous system	3
Figure 1.2.2. Overview of functional MRI.....	6
Figure 1.2.3. Propofol and the cerebral cortex	9
Figure 1.3.1. A simple example of irreducibility	13
Figure 1.3.2. System diagram and transition probability matrix	16
Figure 1.3.3. Computation of integrated information for a single mechanism	18
Figure 1.3.4. Computation of integrated information for a system of mechanisms	20
Figure 2.2.1. From fMRI acquisition to binarized time-series.	28
Figure 2.2.2. TPMs for the time-series of a single subject and 17 subjects	30
Figure 2.2.3. Computing average integrated information for a time-series	34
Figure 2.3.1. Modulation of $\mu[\Phi^{\max}]$ by propofol in individual RSNs.....	39
Figure 2.3.2. Modulation of $\mu[\rho]$ by propofol in individual RSNs.....	40
Figure 2.3.3. Comparison of $\mu[\Phi^{\max}]$ between RSNs and spatial control networks	41
Figure 2.3.4. Comparison of $\mu[\rho]$ between RSNs and spatial control networks.....	42
Figure 2.3.5. Distributions of $\mu[\Phi^{\max}]$ and $\mu[\rho]$ obtained for temporal controls	43
Figure 3.2.1. Individual and grouped permutations.....	50
Figure 3.2.2. Exponential plateau function used to fit grouped permutation data	51
Figure 3.2.3. TPMs for the original and permuted time-series of the frontoparietal network	52
Figure 3.2.4. Diagonally selective rearrangements of time-series	53

Figure 3.2.5. The effects of diagonally selective reordering of a TPM.....	54
Figure 3.3.1. Grouped permutations in higher-order RSNs	55
Figure 3.3.2. Grouped permutations in sensorimotor and sensory RSNs.....	56
Figure 3.3.3. Modulation of λ and its inverse relationship with $\mu[\Phi^{\max}]$	58
Figure 3.3.4. Anticorrelation between λ and $\mu[\Phi^{\max}]$	59
Figure 3.3.5. Relationship between $\mu[\Phi^{\max}]$ and TPM diagonal strength for grouped permutations	61
Figure 3.3.6. Relationship between $\mu[\Phi^{\max}]$ and TPM diagonal strength for grouped permutations	62
Figure 4.2.1. Simulating functional connectivity using the GIM.....	69
Figure 4.2.2. Empirical and simulated functional connectivity matrices	70
Figure 4.3.1. Variation of $\mu[\Phi^{\max}]$ with respect to the number of subjects concatenated	71
Figure 4.3.2. Variation of $\mu[\Phi^{\max}]$ in empirical and simulated time-series.....	72
Figure 6.1.1. Obtaining five representative regions for an RSN	83
Figure 6.1.2. Representative cortical regions of the 11 RSNs from the parcellation scheme.	84

Chapter 1 : Introduction and Background

1.1. Overview

Consciousness is a remarkably complex concept that lacks a concrete and agreed-upon definition [1]. Colloquially, an individual would be described as “conscious” if they demonstrate wakefulness and awareness [2]. However, neuroscientific literature continues to show that conscious phenomena transcend the mere presence or absence of these behavioral traits [3–5].

One technology that has been particularly useful in this respect is functional magnetic resonance imaging (fMRI), which allows for observation of cortical activity in spatial and temporal domains [6]. Activity measured with fMRI corresponds to blood flow patterns that reflect the metabolic activities of different brain regions [7,8]. In studies involving healthy individuals and patients with traumatic brain injuries, fMRI has played a critical role in advancing our knowledge on the neurophysiology that underlies cognition, perception, and conscious phenomena

Owing to developments made with neuroimaging technologies, several theories have been proposed to explain consciousness and how it may arise. In this respect, integrated information theory (IIT) has emerged as a leading framework [9–11]. IIT proposes that consciousness arises when a system, such as the brain, generates a high degree of information through the causal interactions that exist between its elements. It is important to emphasize that in IIT, “information” is derived from a system’s intrinsic causal properties and is hence different from the definition used in traditional information theory (i.e., Shannon entropy) [10]. The metric describing the extent to which this information is integrated within the system is called integrated information, which is denoted by Φ . In this work, we attempt to quantify Φ for resting-state fMRI (rs-fMRI) data, which involves acquisition of spontaneous brain activity without overt stimulation [12]. Even at rest, the brain maintains a sophisticated functional organization that has been the subject of extensive study in recent years. To our knowledge, this is the first published study that attempts to implement IIT to this form of neuroimaging.

Although IIT has had important implications to research on consciousness, its computational scheme poses several requirements that limit applicability to empirical data [13]. Thus, the first part of our study is dedicated to analyzing the time-dependent properties of fMRI data and the extent to which they satisfy the current requirements of IIT. For example, IIT requires that time-series satisfy a temporal feature known as a Markov property, and we employ an extensive statistical analysis to test this.

To determine whether Φ presents a valid means of analyzing “consciousness”, we obtained fMRI data from subjects who underwent sedation with propofol, a short-acting anesthetic [14]. In the central analysis of Chapter 2, we evaluate how Φ varies as subjects experience the different states of awareness induced by this sedative. In Chapter 3, we focus on the relationship between certain temporal properties of fMRI data and how they may relate to integrated information. In Chapter 4, we discuss potential solutions to certain limitations intrinsic to fMRI, where we also aim to provide several guidelines for future implementations of IIT to resting-state fMRI. Ultimately, we hope that our approach will be applied in future studies that focus on other altered states of awareness, such as subjects who are engaged in a task, or patients suffering from disorders of consciousness that arise from traumatic brain injuries.

The following pages of this chapter are organized into two main sections that provide readers with substantial background on our research. First, we provide a brief description of the human nervous system. We then discuss fMRI and how this technique works to produce measurements of brain activity. This is followed by a summary of the literature on propofol and its previously observed effects on resting-state brain activity. Once we cover all relevant concepts in neurophysiology and neuroimaging, we provide a comprehensive overview of IIT. We begin with the qualitative aspects of this framework, and then explain how integrated information is computed at two levels: individual mechanisms and a system of mechanisms.

1.2. Background on Neuroscience and Neuroimaging

1.2.1. The Nervous System

The human brain has always been a subject of great interest to the scientific community. For centuries, physicians, scientists, and philosophers studied the brain and attempted to understand how it gives rise to cognition, perception, and conscious processing. The brain is the central structure of the nervous system, and it has many crucial physiological roles related to control and regulation in the body [15,16].

To briefly discuss the fundamental properties of the nervous system, we begin at the cellular level with neurons. Neurons are the primary cells making up nervous tissue, and their function is to transmit electrical signals known as action potentials. Action potentials allow for communication with other neurons through connections called synapses. We present a diagram of the neuron and its principal parts in Figure 1.2.1a [16]. The part of the neuron responsible for transmitting an action potential is called the axon (also known as a nerve fiber), which projects from the cell body towards other neurons that are connected at nerve terminals.

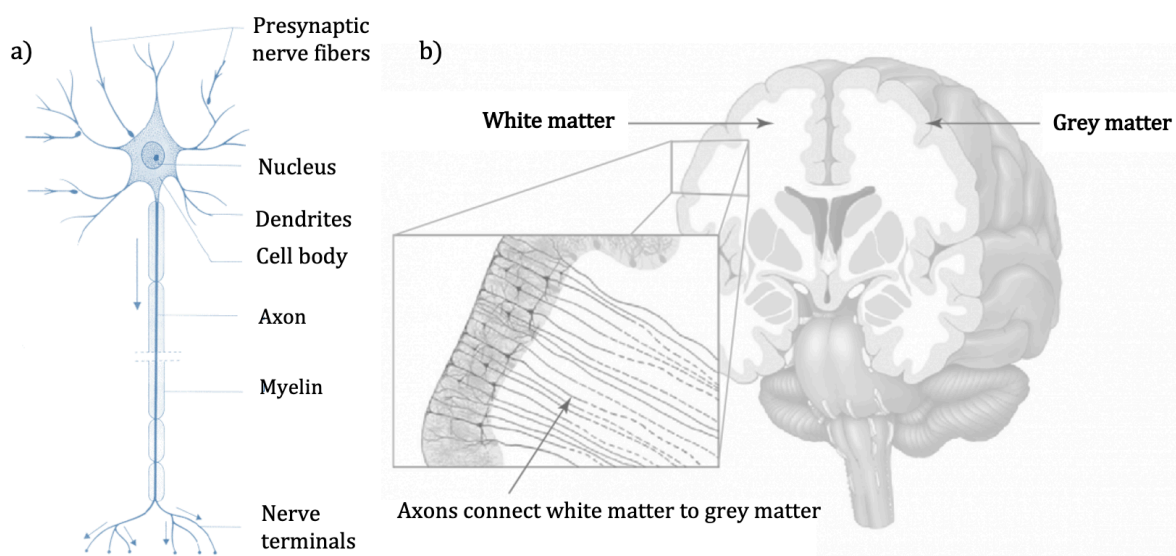


Figure 1.2.1. Fundamental structures of the nervous system. a) Diagram of the neuron, the cells responsible for transmitting signals within the nervous system; b) Diagram of the brain showing white and grey matter (inner and outer layers, respectively), as well as their connections through axons.

While we do not discuss anatomical details, it is important to understand the brain's division into two fundamental structures known as white and grey matter, which we demonstrate in Figure 1.2.1b [15]. White matter is located deep within the brain and is primarily composed of axons. Accordingly, white matter is responsible for the connections between different parts of grey matter, which constitutes the brain's outer layer and is primarily composed of neuronal cell bodies. The outermost layer of the grey matter is called the cerebral cortex, which is responsible for many of the brain's "higher-order" functions, such as awareness, attention, thought, memory, and language. Using neuroimaging, different regions of the cerebral cortex (referred to as cortices) have been associated with specific functions.

1.2.2. Functional Magnetic Resonance Imaging

The advent of advanced neuroimaging techniques greatly advanced our understanding of the brain's structure and function. As a result, contemporary neuroscience almost completely relies on these technologies. In this respect, magnetic resonance imaging (MRI) has been an invaluable tool. MRI scanners produce images of organs through their strong magnetic field [17]. This field interacts with molecules in the body to generate excitatory signals that are captured by the scanner and processed to a form that can be interpreted visually. Although MRI can be used to image several different organs, it is particularly prevalent in neuroscience as it can produce higher resolution images of the brain compared to other forms of imaging [18]. Beyond generating images that allow for a structural perspective of the brain, MRI can also be used to analyze brain activity in the temporal domain. This is accomplished with a technique known as functional MRI (fMRI) [7], which is the principal form of neuroimaging used in our study.

The advent of fMRI was inspired by a series of fundamental discoveries about the magnetic properties of blood, which change depending on its level of oxygenation. More specifically, molecules in deoxygenated blood are paramagnetic and those in oxygenated blood are diamagnetic, meaning that the former has stronger magnetic properties than the latter [19]. As regions of the brain become more active, their consumption of oxygen increases, causing more oxygenated blood to flow towards them. Accordingly, activity-based changes to blood flow correspond to spatial variations of the brain's magnetic

properties. These changes are captured by the scanner in the form of the blood-oxygen-level-dependent (BOLD) signal. Since blood flow patterns are dynamic and change over time, the continuous acquisition of images allows for a time-dependent analysis of the brain's activity patterns.

Consecutive images are separated by the amount of time it takes for the scanner to generate excitatory signals and detect them, which is known as the repetition time (TR). In most cases, TR ranges between 2-3 seconds [7,12]. For each voxel (a three-dimensional pixel of the image), a time-series is obtained, which shows changes to the BOLD signal's strength after each TR. From here, a parcellation scheme is typically applied. Parcellation involves dividing the brain into non-overlapping regions made up of many voxels that are close in proximity. The regions of a parcellation scheme obtain their own time-series based on the signals of their constituent voxels, which allows for a better organized representation of brain structure and function.

The parcellated time-series are then used to construct an important entity known as functional connectivity (FC). Through correlations, FC describes the degree of synchrony that different cortical regions present through their time-series; two regions are described as functionally connected if there is a high correlation between the evolution of their BOLD signals [20]. These statistical relationships are traditionally seen as indicators that regions are coupled or belong to the same functional network. For any pair of regions (x and y), FC is calculated using the Pearson correlation coefficient ρ :

$$\rho_{xy} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \quad 1.2.1$$

where x_i and y_i represent the BOLD signal strength at a particular time-point i . To describe the functional connectivity of N regions (which may include the whole brain or a smaller network of cortical areas), it is common to construct a functional connectivity matrix with dimensions $N \times N$. Each row and column represent a certain region, and the entries are populated with the correlation coefficient of the corresponding pair (e.g., the entry in row 1 and column 3 corresponds to ρ_{13}). The overall process of obtaining FC, starting from image acquisition, is summarized below in Figure 1.2.2.

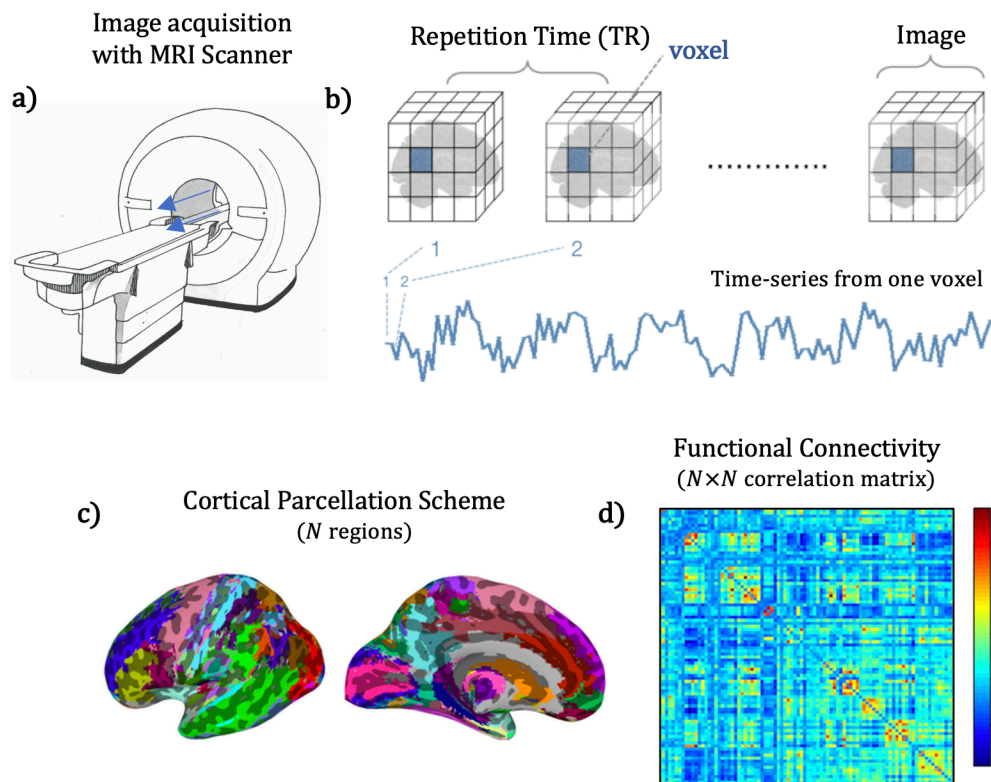


Figure 1.2.2. Overview of functional MRI. a) Image of a typical MRI scanner, which produces a strong magnetic field that induces molecular excitations and picks up their signals to generate images of organs; b) Obtaining a time-series for each voxel based on changes to the BOLD signal over time, with each time point corresponding to one TR; c) Mapping of the cortex with a parcellation scheme, which yields well-defined regions and their corresponding time-series; d) Example of a functional connectivity (FC) matrix, a symmetric array that contains the Pearson correlation coefficient between every pair of cortical regions in the parcellation scheme.

Traditionally, fMRI studies require subjects to undergo stimulation or participate in a certain task, with the goal of detecting which brain regions are activated in response. While such procedures have been essential to our understanding of cortical function, fMRI can also be used to analyze “resting-state” activity, which describes the brain’s intrinsic behavior in the absence of external stimuli or tasks [21]. Even at rest, the brain demonstrates a sophisticated functional organization that is worthy of analysis.

Acquisition of subjects in conditions of rest is called resting-state fMRI (rs-fMRI, also known as task-negative fMRI), and studies employing it have identified a series of resting-state networks (RSNs) [22]. RSNs comprise of several cortical regions that are temporally correlated, while also being functionally segregated from regions of other networks (i.e., they either show no correlation or a negative correlation with external regions). One of the first RSNs to be described was the default mode network (DMN), which gained a lot of attention for its high resting-state activity [23]. Although its functions are not completely understood, it is attributed to introspective thought, memory recollection, and mind-wandering. While these functions correspond to the sort of cognition that would occur in a task-negative state, even networks that are involved in specific tasks, such as the auditory and visual networks, remain synchronized at rest, and are hence also examples of RSNs [22].

The parcellation scheme adopted for this study was used to extract 11 RSNs [24]. In Tables 1.2.1 and 1.2.2, we briefly introduce all these networks and describe their principal roles [24–28]. Seven are characterized as “higher-order” networks, meaning their functions are responsible for advanced cognitive and behavioral processes. The other four networks are associated with sensorimotor functions and the processing of stimuli; although these are not necessarily “lower-order” RSNs, their functions are related to more specific tasks than those in the former group.

RSN	Functions
Frontoparietal (FPN)	<ul style="list-style-type: none"> • Executive control and goal-oriented behavior (task-related) • Integration of external stimuli with working memory
Default Mode (DMN)	<ul style="list-style-type: none"> • Introspective thought, mind wandering, daydreaming • Activity is most pronounced in the resting state
Retrosplenial (RS)	<ul style="list-style-type: none"> • Spatial learning, visual processing, and navigation • Episodic memory of past personal experiences
Dorsal Attention (DAN)	<ul style="list-style-type: none"> • Externally directed and voluntarily driven attention • Goal-orienting and visuospatial processing
Ventral Attention (VAN)	<ul style="list-style-type: none"> • Complements the dorsal attention network • Directs attention towards unexpected stimuli (involuntary)

Cingulo-Opercular (CO)	<ul style="list-style-type: none"> Monitoring and coordination of control networks (such as FPN and DAN)
Cingulo-Parietal (CP)	<ul style="list-style-type: none"> High activity during resting-state, complements the activity of the DMN

Table 1.2.1. Descriptions of higher-order cortical networks.

RSN	Functions
Sensorimotor Hand	<ul style="list-style-type: none"> Coordination and execution of movements in the hand
Sensorimotor Mouth	<ul style="list-style-type: none"> Coordination and execution of mouth and lip movements
Auditory	<ul style="list-style-type: none"> Processing of auditory stimuli
Visual	<ul style="list-style-type: none"> Processing of visual stimuli

Table 1.2.2. Descriptions of sensorimotor, auditory, and visual networks.

1.2.3. Anesthesia with Propofol

In attempts to understand the cortical pathways responsible for awareness, fMRI has been used to study anesthesia. Anesthetics are drugs that suppress the body's normal autonomic functions, such as cardiovascular activity, respiration, and brain function, hence resulting in a state of unconsciousness [29]. One of the most common anesthetics used in clinical and research settings is propofol, which is administered intravenously to rapidly induce an unconscious state [30].

Despite the common use of propofol and all the extensive research conducted on it, much remains unknown about its mechanisms of action. Previous findings indicate that propofol preferentially suppresses higher-order processing regions in the frontal and prefrontal cortices, which are commonly attributed to awareness and perception [14,31,32] (the four lobes of the cortex are presented in Figure 1.2.3). One important RSN that is affected is the frontoparietal network (FPN), whose principal function is related to executive control and the integration of external stimuli [26,33]. The FPN includes regions of the prefrontal cortex and the parietal lobe. Previous fMRI studies found that under anesthesia, the functional connections between these regions break

down [31,34,35]. At the same time, propofol was not found to diminish the reactivity of sensory cortices (i.e., visual, auditory networks) to external stimuli, which also emphasizes the preferential suppression of higher-order cortices [36].

Since this study is based on fMRI data of subjects who underwent sedation with propofol, we will return to these principles when discussing how propofol affects the conscious state of subjects in Chapter 2.

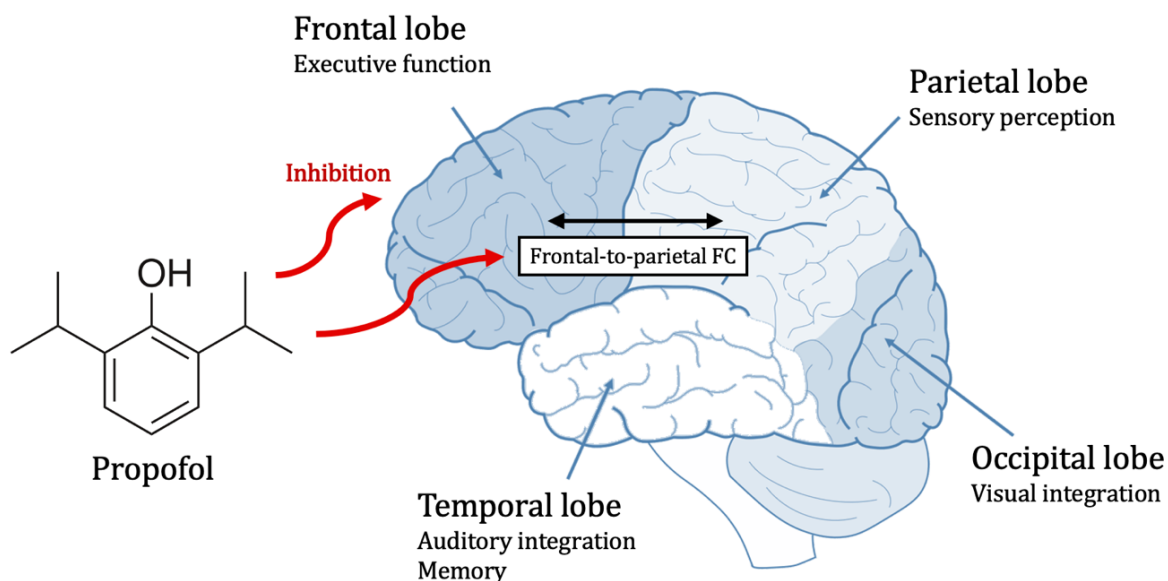


Figure 1.2.3. Propofol and the cerebral cortex. The molecular structure of propofol is given on the left. The red arrows indicate propofol's preferential suppression of the frontal lobe (cerebral blood flow to this region is inhibited under anesthesia), as well as the breakdown of functional connectivity (FC) it induces between frontal and parietal regions. We also indicate the four major lobes of the cortex and briefly describe their functions.

1.3. Integrated Information Theory

1.3.1. Defining Consciousness

In recent years, advances in neuroimaging emphasized the intricate nature of consciousness and changed the neuroscientific community's perspective on this concept. Traditionally, consciousness has been described as a behavioral phenomenon; an individual would be described as "conscious" if they are awake and aware, and "unconscious" if they are asleep, anaesthetized, or comatose.

Over the last two decades, several important studies were published on patients suffering from disorders of consciousness (DOCs) [37]. DOCs typically arise from traumatic brain injuries and pertain to patients who demonstrate insufficient motor skills and an inability to communicate. One example is a condition known as the vegetative state, in which patients suffer from an almost complete lack of behavioral abilities [38]. Even so, fMRI studies have shown that individuals in this condition are more conscious than their physical state suggests. In a series of experiments, patients in a vegetative state participated in an imaginary task where they were instructed to imagine playing a sport or to navigating a familiar environment. Remarkably, some of these patients demonstrated cortical activation patterns that were similar to those observed for healthy controls performing the same task, indicating a level of awareness that cannot be observed externally. These findings were highly consequential, and it is now well-established that individuals diagnosed with DOC may maintain a "covert" level of consciousness that is independent of their behavioral ability [39–41].

Over the last decade, these discoveries have been complemented by several attempts to explain conscious phenomena and how they may arise. By using neuroimaging techniques, one important goal has been to identify neural correlates of consciousness (NCCs), which are measurable features of brain activity that correlate with specific conscious percepts (perceptions) [42,43]. In analysis involving NCCs, a recurring concept of great importance is causality: the relationship between cause and effect [44–46]. Considering what it means to be conscious in day-to-day life, it is expected that some event (cause) will produce a behavioral, emotional, or cognitive

response (effect). From a neurological perspective, the analysis of causal interactions between different regions of the brain can identify the functional pathways that underly conscious phenomena.

While NCCs have been useful in understanding the neurological foundation for specific conscious percepts, recent attempts have been made to develop a theory that can explain consciousness more generally. As such, integrated information theory (IIT) was introduced as a framework that aims to account for all types of conscious experiences. This thesis focuses on IIT and the use of its principal metric, Φ , which aims to quantify the extent to which a system of interacting elements, such as the brain, is conscious [9–11]. IIT was initially introduced by Giulio Tononi in 2004 [9] and has since undergone extensive developments. The qualitative and quantitative material we present in this section is primarily based on the latest version of IIT (version 3.0), which was published by Oizumi et al. in 2014 [11]. Although we provide a considerable amount of detail in this section, we emphasize that IIT is an exceptionally sophisticated framework, and our explanation is only a summary of its fundamental ideas.

First, IIT defines the fundamental properties of consciousness based on its phenomenological aspects. The conscious nature of an experience is explained with a series of five axioms, or self-evident truths, about the experience [47]:

- 1) **Existence:** Consciousness exists and is real; this reality is intrinsic to the subject undergoing the conscious experience, who views their present experience as actual, regardless of what external observers experience.
- 2) **Composition:** Each conscious experience is structured and composed of multiple distinguishable features, such as colors, shapes, and sizes.
- 3) **Information:** A conscious experience is specific and informative. This is not only because of the distinguishable features it contains, but also in how these features differ from those of other experiences.
- 4) **Integration:** The features making up a conscious experience are interdependent and inseparable; they cannot be experienced alone. (i.e., when looking at a blue square, color and shape are experienced together as interwoven entities.

- 5) **Exclusion:** A conscious experience is limited; its composition and information exclude external features that are not part of the experience.

Using these axioms as a foundation, IIT proposes that consciousness arises when a system, such as the brain, generates more information as an integrated whole than that which is generated by its elements acting alone. In other words, “consciousness” is tantamount to “irreducibility”; the interactions between a system’s elements, governed by its mechanisms, give rise to information that would not be produced if the system was reduced to its individual components.

Although this may seem to contradict traditional notions of information theory, such as the principle of additivity (i.e., a composite system’s total entropy must equal to the sum of the entropies produced by its subsystems), it is important to emphasize exactly what type of “information” is being discussed. In accordance with the axiom of existence, the system must exist *intrinsically*, regardless of how it may be viewed externally. Fundamentally, the idea of intrinsic existence means that the system must be able to make a difference to itself; accordingly, information, as it is defined in IIT, quantifies the extent of this difference [47]. As we shall see, this intrinsic level of information is computed through the causal properties that govern the interactions between the system’s elements. On the other hand, measures from information theory, such as Shannon entropy, are *extrinsic*, meaning they are assessed from the perspective of an external observer [9,10]. This is not the type of information used in IIT, as a system can exist extrinsically but not from its own perspective (i.e., it cannot make a difference to itself).

Once information is defined, IIT introduces the concept of integrated information, which describes the extent to which the information of the type described above pertains to a system as an inseparable whole. The measure for integrated information is Φ , which is obtained by assessing the system’s irreducibility to its individual components.

The idea of irreducibility and how it relates to consciousness is summarized with a simple example in Figure 1.3.1. When we look at a smiley face, our conscious interpretation pertains to the symbol as an inseparable whole; although we may recognize that it consists of several features, such as an outer circle, eyes, and a mouth, we do not

experience these features separately. Instead, our experience is an integration that contains more meaning than the individual features themselves. The same could be said for the shapes and colors of these features (i.e., when we observe a blue circle, shape and color are inseparable – they cannot be experienced separately).

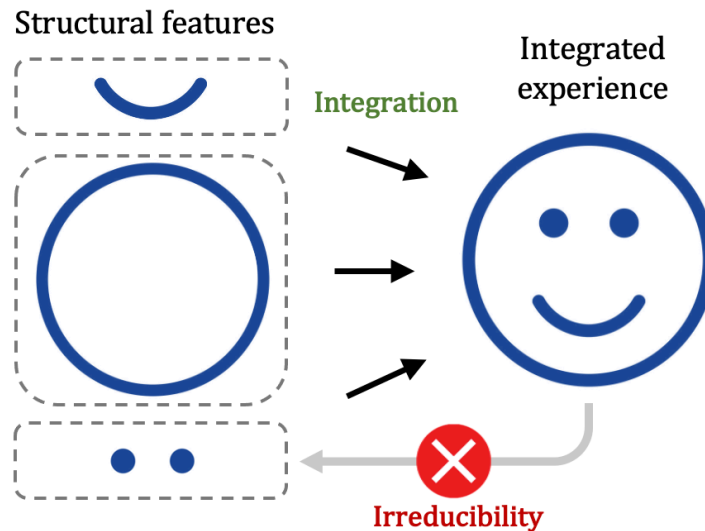


Figure 1.3.1. A simple example of irreducibility. When observing a smiley face, our interpretation is an integration of its distinguishable features, and we ultimately observe the symbol as an irreducible whole that has more meaning than the individual features themselves; this conscious experience cannot be reduced to its individual elements.

As another example, consider the following comparison between a camera and the human visual system [9]. Although digital photographs contain millions of pixels, the camera sensor that produces these images is not considered to be “conscious”, as it does not integrate the elements of the images in any meaningful way; the information stored by the camera is no more than the color grade of each pixel. When we observe an image with our eyes, light hits the cells of our retinas as it does in a camera sensor. However, these inputs undergo processing in the visual cortex, which integrates the different aspects of the stimulus. The ultimate result is a meaningful interpretation of what we see, which can be described as a “conscious” experience; the information produced by the visual system is more than just the input received by each retinal cell.

Finally, an example of neuroscientific evidence that is commonly used to support IIT is the distinction between two essential structures of the human brain: the cerebellum and cerebral cortex [47]. Although the cerebellum contains four times as many neurons, functions associated with consciousness are primarily attributed to the cerebral cortex [48,49]. This is because its neurons are much more interconnected than those in the cerebellum, which results in the complex interactions and pathways that underly this structure's higher-order cognitive functions [47,50]. Previous studies on patients with damage to their cerebellum (and remarkably, even those born without one) emphasized that they still demonstrate overt consciousness through their behavior [51].

1.3.2. Quantifying Integrated Information for a Single Mechanism

When computing integrated information, one must begin with a system (or network) consisting of several elements that interact through well-defined mechanisms. An element is any component of the system that can take on a certain state, and a mechanism is any subsystem of elements that can change the system's state (i.e., mechanisms have a causal role in the system). A mechanism can be elementary, meaning it consists of one element, or composite, meaning it consists of two or more elements. In the current computational scheme of IIT [13], each element in the system is restricted to be a binary node, meaning its possible states are 1 and 0 ("ON" and "OFF", respectively). Integrated information is first computed for an individual mechanism, with which we begin our computational overview. Following the example presented by Oizumi et al., [11] we focus on a simple bidirectional network consisting of three elements A , B , C , all of which interact as logic gates (OR, AND, XOR, respectively). These three logic gates were chosen for this specific example, but this analysis could be applied to any other collection of logic gates or system with well-defined interactions.

For a single mechanism, integrated information is quantified using "small phi", or ϕ . The system on which we focus is presented in Figure 1.3.2a, and the elementary mechanism for which we aim to compute ϕ is A^1 , where A is in the "ON" state. When thinking about the causal properties of A^1 , we aim to determine how this mechanism restricts the past and future states of the system. In other words, we are asking the

following question: what does A being ON tell us about the system's past and future states?

The mathematical entity that defines each mechanism's causal properties is the system's transition probability matrix (TPM). In Figure 1.3.2b, the system's TPM is given in a state-by-state form, where each row and column correspond to the system's overall state. Since each node is binary, the system can be in one of 8 (2^3) possible states, so the dimensions of this TPM are 8×8 . First, each state is assigned a binary representation ($A = 1, B = 0, C = 0 \rightarrow 100$), which is converted to an integer for indexing purposes. In PyPhi, this conversion is accomplished using a little-endian byte order (i.e., bytes are flipped from left to right, meaning the least significant bytes are encoded first: state 100 \rightarrow state index 1). The indices corresponding to each row's state are presented on the right side of the TPM in Figure 1.3.2b.

An entry of the TPM describes the probability that the state corresponding to the row transitions to the state corresponding to the column. For instance, the state 100 (state index 1) has a 100% chance of transitioning to the state 001 (state index 4), so the entry in row 1 and column 4 is assigned a 1. Since C is an XOR gate and is only activated when both inputs are different, it will be activated as a result of A being ON and B being OFF in the current state. At the same time, neither A nor B will be activated (both inputs are the same for A and different for B), hence the transition to state 001.

The mechanisms defined in this example leave no chance of transitioning to more than one state, so all other entries in row 1 correspond are assigned a zero (the sum of each row must be normalized to 1). Therefore, this TPM is said to be *deterministic*, as the mechanisms of the system are well-defined, and each state is associated with one specific transition. On the other hand, a TPM can also be *probabilistic*; instead of having definite transitions for each state (i.e., 100% or 0% probability for each entry as in the deterministic case), a state can transition to more than one other state with a specified probability for each transition. The entries of a probabilistic TPM can hence be less than 1. In either case, each row must be normalized so that all its elements add to 1.

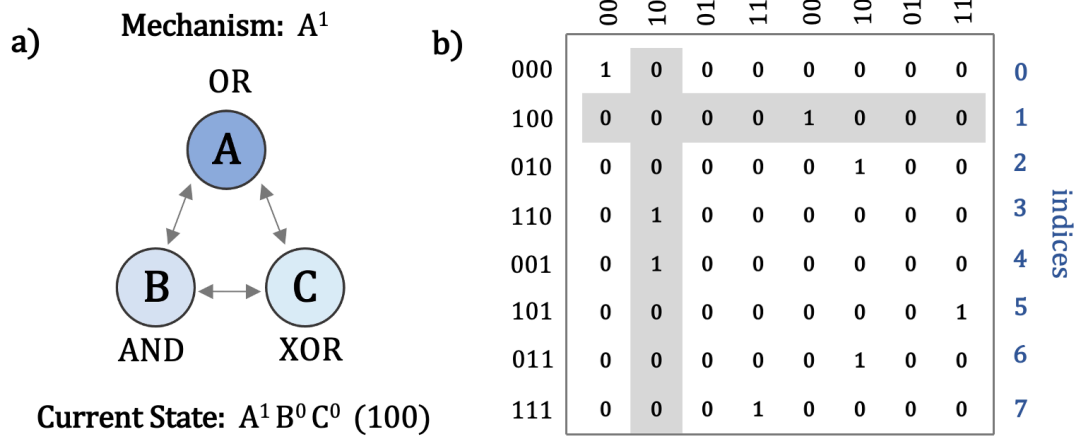


Figure 1.3.2. System diagram and transition probability matrix. a) The system being considered, which has three elements acting as logic gates; b) A state-by-state transition probability matrix, where each entry specifies the probability that the state corresponding to the row transitions to the state corresponding to the column. Since each state can only transition to one other state and all non-zero entries are 1, this TPM is said to be deterministic.

The transition probability matrix is then used to construct a set of probability mass functions for the mechanism of interest, which are known as cause-effect repertoires. These are presented as bar graphs in Figure 1.3.3a. The horizontal and vertical axes contain the system's possible states and their corresponding probabilities, respectively. The cause repertoire, $p(ABC^p | A^1)$, specifies the probability that a certain past state (ABC^p) preceded the mechanism, whereas the effect repertoire, $p(ABC^f | A^1)$, specifies the probability of each future state (ABC^f) that may follow A^1 . The cause-effect repertoires are defined for time points that *immediately* precede or follow the mechanism.

These distributions are then used to compute the information generated by the mechanism. Information quantifies the degree to which a mechanism's causes and effects are selective, or constrained; the more selective these causal properties are, the more "relevant" the mechanism is to the evolution of the system, which corresponds to high information. On the other hand, an unselective mechanism has a weaker causal role in the system and is hence less relevant to the system's evolution; its causes and effects are less specific, which yields little to no information. This is quantified by measuring the distance between the cause-effect repertoires and their "unconstrained" counterparts,

which are shown underneath the original cause-effect repertoires in Figure 1.3.3a. The unconstrained cause repertoire, $p^{uc}(ABC^p)$, is a uniform distribution where all states are equally possible, while the unconstrained effect repertoire, $p^{uc}(ABC^f)$, is obtained by noising the inputs to each element (A^1 is not included in the unconstrained probability mass functions as they do not depend on the mechanism in question). The distance between the cause repertoire and its unconstrained counterpart is defined as the cause information (ci), while the same metric applied to the effect probability mass functions is defined as the effect information (ei). The mechanism's overall information is defined as the minimum of ci and ei , or cause-effect information (cei). The minimum implies that for a mechanism to generate information, it must do so through *both* its causes and effects. In other words, if one of ei or ci is 0, then cei is also 0.

Since integrated information is a measure of irreducibility, it is computed by partitioning the system (breaking connections between its elements) and measuring the loss of information that follows. If the mechanism's information in the partitioned system is not reduced, it is an indication that the system is poorly integrated and can be reduced to its individual components. This process is demonstrated in Figure 1.3.3b, which begins with all possible ways of partitioning the system.

For each partition, cei is computed and compared to that of the unpartitioned mechanism. The partition that induces the smallest difference to information (Δcei), or the minimum information partition (MIP), is the one used to measure the mechanism's integrated information, ϕ . The MIP is used because it provides a definite constraint on how strongly integrated the system can be. To compute ϕ , the distance between the unpartitioned cause-effect repertoires and those of the MIP are obtained. Finally, this is repeated for every possible subset of the system's elements (see methods for more details on subset analysis). The degree to which the system is integrated for a single mechanism is ϕ^{\max} , which corresponds to the subset that maximizes ϕ .

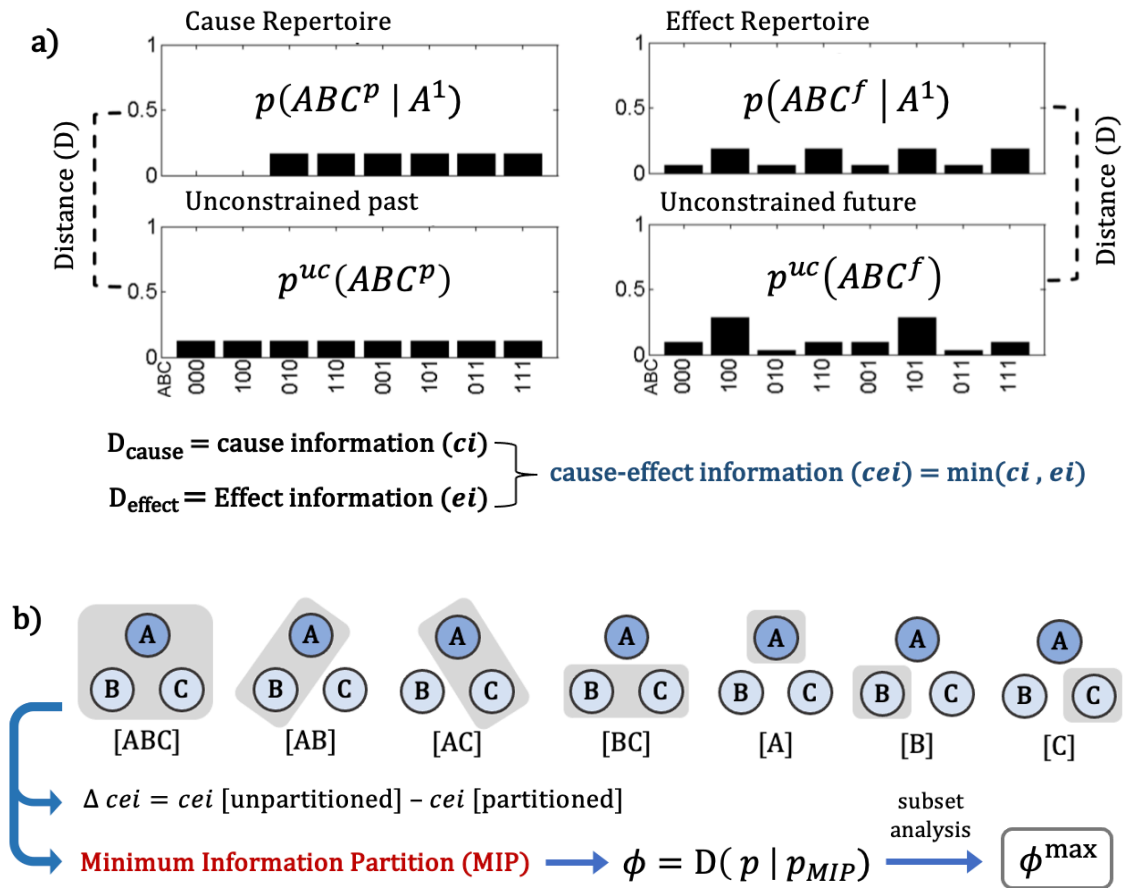


Figure 1.3.3. Computation of integrated information for a single mechanism.

a) The distance between the cause-effect repertoires and their unconstrained counterparts is computed to obtain the mechanism's cause and effect information. The minimum of these defines the overall information produced by the mechanism;

b) The system is partitioned in all possible ways, and the MIP is used to compute the integrated information of the mechanism, ϕ . This is repeated over all possible subsets, and the one that maximizes is used to obtain the maximum integrated information, ϕ^{\max} .

1.3.3. Quantifying Integrated Information for a System of Mechanisms

When a system takes on a specific state, there is more than one mechanism that can impact its evolution. In the next part of the formulation, the analysis for a single mechanism is extended to all the possible mechanisms within a system. The ultimate result is an intricate measure of irreducibility called the maximally integrated conceptual information, or Φ^{\max} , which is the principal metric of integrated information theory. Although this computation is significantly more sophisticated, we briefly describe how it is analogous to the procedure used for a single mechanism.

First, the ϕ^{\max} values of all possible mechanisms are computed. At this point, each mechanism, its cause-effect repertoires, and its maximum integrated information are referred to as concepts. This is used to construct a high-dimension probability distribution called a conceptual structure, which is populated with the ϕ^{\max} values of each concept (like the cause-effect repertoires, a conceptual structure is obtained for both past and future states). The dimensions of this space are 2×2^N , N being the number of nodes, with one dimension for each system state in both past and future. This time, the system is also partitioned, but the loss of information is evaluated for the conceptual space, and a version of the MIP is found using all possible partitions. The integrated information for the system of mechanisms, Φ , is defined as the distance between the unpartitioned conceptual structure and that of the MIP. Finally, a subset analysis is performed to obtain Φ^{\max} , which corresponds to the conceptual structure that maximizes Φ , or the maximally irreducible conceptual structure. This process is summarized in Figure 1.3.4, where the initial conceptual structure is compared to that of the MIP.

In essence, the maximally integrated conceptual structure is a quantitative representation of a conscious experience. As explained in the beginning of this section, a conscious experience integrates the distinguishable features that make up its composition. In like manner, the maximally irreducible conceptual structure integrates the individual concepts that make it up. The concepts themselves can be seen as distinct and distinguishable features of the experience, and we discussed how these are both informative and integrated through their causal roles in the system.

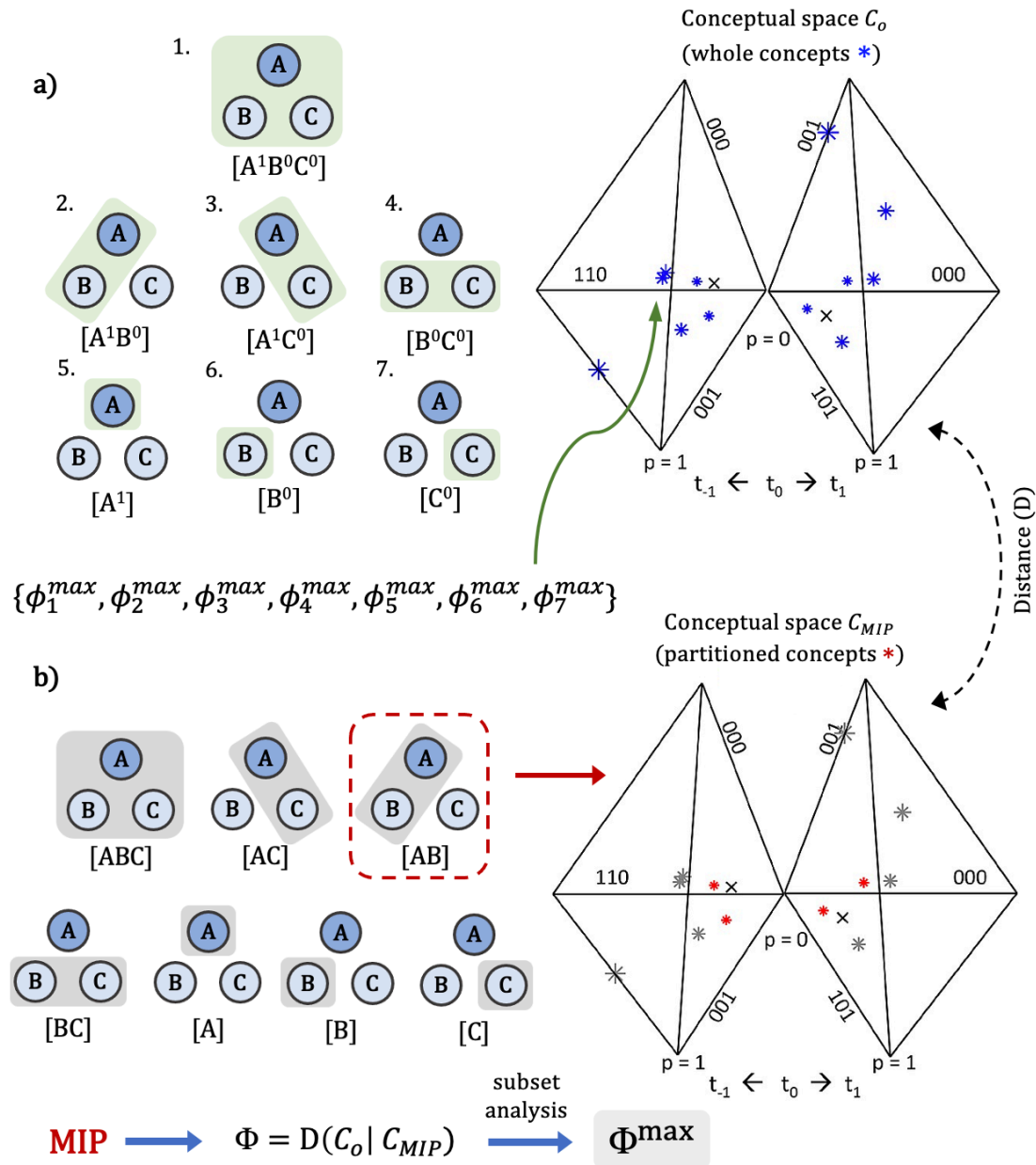


Figure 1.3.4. Computation of integrated information for a system of mechanisms. a) Each mechanism of the system when it is in the state 100 is highlighted. Each mechanism and its ϕ^{max} values specify a “concept”, which is used to populate the system’s conceptual space; b) The system is partitioned, and the subsequent loss of information is evaluated for the conceptual space. The MIP is used to compute Φ , which is repeated for every possible subset. The subset that maximizes Φ is termed the maximally integrated conceptual structure, and its Φ^{max} value (maximally integrated conceptual information) specifies the degree to which the system is “conscious”.

In conclusion, extracting the principal metric of IIT is a computationally rigorous process. Implementing this analysis to neuroimaging data has proven to be a difficult challenge, and as a result, most developments have been theoretical and speculative. In the next chapter, we outline a methodology for how Φ^{\max} could be extracted from empirical fMRI data. Although this section focused on a simple network with three elements and a deterministic evolution, calculations of Φ^{\max} can be extended to larger systems as well as those with probabilistic mechanisms.

Although the present work is primarily focused on IIT, we acknowledge that this framework continues to be debated in the neuroscientific community and that other theories have been proposed to explain consciousnesses [43–45]. Therefore, we emphasize that our work is mainly an experimentation with this theory; while we attempt to relate our results to existing neuroscientific literature, much remains to be investigated when it comes to NCCs like integrated information. Ultimately, we hope that our approach will motivate future studies on how IIT pertains to real-life neurological systems and, most importantly, whether it can account for conscious phenomena.

1.4. References

- [1] Zeman A. Consciousness. *Brain* 2001;124:1263–89. <https://doi.org/10.1093/brain/124.7.1263>.
- [2] Sanders RD, Tononi G, Laureys S, Sleigh JW, Warner DS. Unresponsiveness \neq Unconsciousness. *Anesthesiology* 2012;116:946–59. <https://doi.org/10.1097/ALN.0b013e318249d0a7>.
- [3] Bayne T, Hohwy J, Owen AM. Are There Levels of Consciousness? *Trends in Cognitive Sciences* 2016;20:405–13. <https://doi.org/10.1016/j.tics.2016.03.009>.
- [4] Owen AM. The Search for Consciousness. *Neuron* 2019;102:526–8. <https://doi.org/10.1016/j.neuron.2019.03.024>.
- [5] Bayne T, Seth AK, Massimini M. Are There Islands of Awareness? *Trends in Neurosciences* 2020;43:6–16. <https://doi.org/10.1016/j.tins.2019.11.003>.
- [6] Snider SB, Edlow BL. MRI in disorders of consciousness. *Current Opinion in Neurology* 2020;33:676–83. <https://doi.org/10.1097/WCO.0000000000000873>.
- [7] Glover GH. Overview of Functional Magnetic Resonance Imaging. *Neurosurgery Clinics of North America* 2011;22:133–9. <https://doi.org/10.1016/j.nec.2010.11.001>.
- [8] Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 2008.
- [9] Tononi G. An information integration theory of consciousness. *BMC Neuroscience* 2004;5:42. <https://doi.org/10.1186/1471-2202-5-42>.
- [10] Tononi G. Integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie* 2012;150:56–90. <https://doi.org/10.4449/aib.v149i5.1388>.
- [11] Oizumi M, Albantakis L, Tononi G. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology* 2014;10:e1003588. <https://doi.org/10.1371/journal.pcbi.1003588>.
- [12] Smitha K, Akhil Raja K, Arun K, Rajesh P, Thomas B, Kapilamoorthy T, et al. Resting state fMRI: A review on methods in resting state connectivity analysis and resting state networks. *The Neuroradiology Journal* 2017;30:305–17. <https://doi.org/10.1177/1971400917697342>.
- [13] Mayner WGP, Marshall W, Albantakis L, Findlay G, Marchman R, Tononi G. PyPhi: A toolbox for integrated information theory. *PLOS Computational Biology* 2018;14:e1006343. <https://doi.org/10.1371/journal.pcbi.1006343>.
- [14] Song X, Yu B. Anesthetic effects of propofol in the healthy human brain: functional imaging evidence. *Journal of Anesthesia* 2015;29:279–88. <https://doi.org/10.1007/s00540-014-1889-4>.

- [15] Kandal E, Schwartz JH, Jessel T, Siegelbaum S, Hudspeth A. Principles of Neural Science. vol. 1. 3rd ed. New York: McGraw-Hill; 2012.
- [16] Nichols JG, Martin RA, Fuchs PA, Brown DA, Diamond M, Weisblat DA. From Neuron to Brain. vol. 2015. 5th ed. Sunderland, MA: Sinauer Associates; n.d.
- [17] Dale BM, Brown MA, Semelka RC. MRI: Basic Principles and Applications. 5th ed. Hoboken, NJ: Wiley Blackwell ; 2015.
- [18] Alvarez-Linera J. 3T MRI: Advances in brain imaging. *European Journal of Radiology* 2008;67:412–26.
- [19] Logothetis NK, Wandell BA. Interpreting the BOLD Signal. *Annual Review of Physiology* 2004;66:735–69. <https://doi.org/10.1146/annurev.physiol.66.082602.092845>.
- [20] Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine* 1995;34:537–41. <https://doi.org/10.1002/mrm.1910340409>.
- [21] Lee MH, Smyser CD, Shimony JS. Resting-State fMRI: A Review of Methods and Clinical Applications. *American Journal of Neuroradiology* 2013;34:1866–72. <https://doi.org/10.3174/ajnr.A3263>.
- [22] Damoiseaux JS, Rombouts SARB, Barkhof F, Scheltens P, Stam CJ, Smith SM, et al. Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences* 2006;103:13848–53. <https://doi.org/10.1073/pnas.0601417103>.
- [23] Raichle ME. The Brain’s Default Mode Network. *Annual Review of Neuroscience* 2015;38:433–47. <https://doi.org/10.1146/annurev-neuro-071013-014030>.
- [24] Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cerebral Cortex* 2016;26:288–303. <https://doi.org/10.1093/cercor/bhu239>.
- [25] Vossel S, Geng JJ, Fink GR. Dorsal and Ventral Attention Systems. *The Neuroscientist* 2014;20:150–9. <https://doi.org/10.1177/1073858413494269>.
- [26] Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL. Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. *Journal of Neurophysiology* 2008;100:3328–42. <https://doi.org/10.1152/jn.90355.2008>.
- [27] Toro R, Fox PT, Paus T. Functional coactivation map of the human brain. *Cerebral Cortex* (New York, NY : 1991) 2008;18:2553–9. <https://doi.org/10.1093/cercor/bhn014>.
- [28] Vann SD, Aggleton JP, Maguire EA. What does the retrosplenial cortex do? *Nature Reviews Neuroscience* 2009;10:792–802. <https://doi.org/10.1038/nrn2733>.
- [29] Fox AJ, Rowbotham DJ. Anaesthesia. *BMJ (Clinical Research Ed)* 1999;319:557–60. <https://doi.org/10.1136/bmj.319.7209.557>.

- [30] Lundström S, Twycross R, Mihalyo M, Wilcock A. Propofol. *Journal of Pain and Symptom Management* 2010;40:466–70. <https://doi.org/10.1016/j.jpainsymman.2010.07.001>.
- [31] Alkire MT, Hudetz AG, Tononi G. Consciousness and Anesthesia. *Science* 2008;322:876–80. <https://doi.org/10.1126/science.1149213>.
- [32] Hudetz AG. General Anesthesia and Human Brain Connectivity. *Brain Connectivity* 2012;2:291–302. <https://doi.org/10.1089/brain.2012.0107>.
- [33] Szczepanski SM, Pinsk MA, Douglas MM, Kastner S, Saalman YB. Functional and structural architecture of the human dorsal frontoparietal attention network. *Proceedings of the National Academy of Sciences* 2013;110:15806–11. <https://doi.org/10.1073/pnas.1313903110>.
- [34] Ku S-W, Lee U, Noh G-J, Jun I-G, Mashour GA. Preferential Inhibition of Frontal-to-Parietal Feedback Connectivity Is a Neurophysiologic Correlate of General Anesthesia in Surgical Patients. *PLoS ONE* 2011;6:e25155. <https://doi.org/10.1371/journal.pone.0025155>.
- [35] Brown EN, Lydic R, Schiff ND. General anesthesia, sleep, and coma. *The New England Journal of Medicine* 2010;363:2638–50. <https://doi.org/10.1056/NEJMra0808281>.
- [36] Liu X, Lauer KK, Douglas Ward B, Roberts C, Liu S, Gollapudy S, et al. Propofol attenuates low-frequency fluctuations of resting-state fMRI BOLD signal in the anterior frontal cortex upon loss of consciousness. *NeuroImage* 2017;147:295–301. <https://doi.org/10.1016/j.neuroimage.2016.12.043>.
- [37] Giacino JT, Fins JJ, Laureys S, Schiff ND. Disorders of consciousness after acquired brain injury: the state of the science. *Nature Reviews Neurology* 2014;10:99–114. <https://doi.org/10.1038/nrneurol.2013.279>.
- [38] Laureys S, Owen AM, Schiff ND. Brain function in coma, vegetative state, and related disorders. *The Lancet Neurology* 2004;3:537–46. [https://doi.org/10.1016/S1474-4422\(04\)00852-X](https://doi.org/10.1016/S1474-4422(04)00852-X).
- [39] Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD. Detecting awareness in the vegetative state. *Science (New York, NY)* 2006;313:1402. <https://doi.org/10.1126/science.1130197>.
- [40] Goldfine AM, Victor JD, Conte MM, Bardin JC, Schiff ND. Bedside detection of awareness in the vegetative state. *Lancet (London, England)* 2012;379:1701–2; author reply 1702. [https://doi.org/10.1016/S0140-6736\(12\)60714-4](https://doi.org/10.1016/S0140-6736(12)60714-4).
- [41] Edlow BL, Claassen J, Schiff ND, Greer DM. Recovery from disorders of consciousness: mechanisms, prognosis and emerging therapies. *Nature Reviews Neurology* 2021;17:135–56. <https://doi.org/10.1038/s41582-020-00428-x>.
- [42] Rees G. Neural correlates of consciousness. *Annals of the New York Academy of Sciences* 2013;1296:4–10. <https://doi.org/10.1111/nyas.12257>.

- [43] Koch C, Massimini M, Boly M, Tononi G. Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience* 2016;17:307–21. <https://doi.org/10.1038/nrn.2016.22>.
- [44] Seth AK, Izhikevich E, Reeke GN, Edelman GM. Theories and measures of consciousness: An extended framework. *Proceedings of the National Academy of Sciences* 2006;103:10799–804. <https://doi.org/10.1073/pnas.0604347103>.
- [45] Seth AK, Barrett AB, Barnett L. Granger Causality Analysis in Neuroscience and Neuroimaging. *Journal of Neuroscience* 2015;35:3293–7. <https://doi.org/10.1523/JNEUROSCI.4399-14.2015>.
- [46] Seth AK, Barrett AB, Barnett L. Causal density and integrated information as measures of conscious level. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 2011;369:3748–67. <https://doi.org/10.1098/rsta.2011.0079>.
- [47] Tononi G, Koch C. Consciousness: here, there and everywhere? *Philosophical Transactions of the Royal Society B: Biological Sciences* 2015;370:20140167. <https://doi.org/10.1098/rstb.2014.0167>.
- [48] Glickstein M, Doron K. Cerebellum: Connections and Functions. *The Cerebellum* 2008;7:589–94. <https://doi.org/10.1007/s12311-008-0074-4>.
- [49] Van Essen DC, Donahue CJ, Glasser MF. Development and Evolution of Cerebral and Cerebellar Cortex. *Brain, Behavior and Evolution* 2018;91:158–69. <https://doi.org/10.1159/000489943>.
- [50] Bower JM. The organization of cerebellar cortical circuitry revisited: implications for function. *Annals of the New York Academy of Sciences* 2002;978:135–55. <https://doi.org/10.1111/j.1749-6632.2002.tb07562.x>.
- [51] Boyd CAR. Cerebellar agenesis revisited. *Brain* 2010;133:941–4. <https://doi.org/10.1093/brain/awp265>.

Chapter 2 : Integrated Information in Resting-State Networks

2.1. Overview

In this chapter, we focus on quantifying the integrated information generated by individual resting-state networks (RSNs). We treat these as isolated systems, meaning no interactions between different RSNs are considered. Details about the acquisition procedure are provided in Chapter 6. In short, resting-state fMRI data was acquired for 17 healthy subjects who underwent sedation with propofol, which allowed for measurements over four states of awareness: 1) awake (prior to propofol administration), 2) mild sedation (low propofol concentration), 3) deep sedation (high propofol concentration), and 4) recovery (subjects regain consciousness). We obtained 11 RSNs, which were processed to include five regions and their corresponding time-series.

Our methodology begins with the processing and conversion of fMRI signals to a binary form that is suitable for calculations of integrated information. The software used to compute Φ was PyPhi, a publicly available Python software package developed in accordance with IIT 3.0 [1]. PyPhi assumes that networks obey two important features, which are the Markov property and conditional independence. After testing how well the time-series obey these requirements, we used PyPhi to compute Φ^{\max} for each state in an RSN's time-series. Based on each network's transitions between different states, we obtained a weighted average of Φ^{\max} , which we refer to as $\mu[\Phi^{\max}]$. This is the primary metric we use when describing each RSN's overall integrated information.

In our central analysis, we compute $\mu[\Phi^{\max}]$ for all networks and states of awareness. Our goal is to determine whether integrated information can meaningfully reflect sedative-induced changes to the condition of subjects. If this metric can serve as a valid NCC, it is expected that its value will gradually decrease as subjects transition from wakefulness to deep sedation and increase as the anesthetic wears off. As a reference point, we also computed each RSN's average interregional Pearson correlations, with which we hoped to point out any results of $\mu[\Phi^{\max}]$ that go beyond traditional descriptions of functional connectivity. In the discussion, will relate our findings to existing neuroscientific literature on RSNs and propofol.

2.2. Methods

2.2.1. Obtaining binary signals and transition probability matrices

The process used to obtain discrete networks of binary elements, starting with fMRI acquisition, is presented in Figure 2.2.1. For each subject, RSN, and conscious condition, we obtained a time-series array with dimensions of 245×5 (number of time points \times number of regions). We chose five regions to optimize balance between spatial resolution and computational efficiency; PyPhi has a scalability of $O(n53^n)$, where n is the number of regions, meaning that a sixth region would have increased computation time by over 60-fold. Once time-series were denoised and filtered (see Chapter 6.1.2), they were standardized to have zero mean, which was accomplished by obtaining each region's mean signal strength and subtracting it from the region's time-series. For each time-point, regions with a positive z-score (above-baseline activity) were set to equal 1, and those with a negative z-score (below-baseline activity) were set to equal 0.

With five regions and two possible states for each one, this allowed for a network to be in one of 32 (2^5) possible states a certain point in time. For instance, the state (1, 1, 1, 1, 1) represents above-baseline activity in all of a network's regions. Each state was then assigned an index between 0 and 31 using little-endian byte ordering (see Chapter 1.3.2) as shown to the right of the binarized time-series in Figure 2.2.1d.

Indexing was essential for the next step, which was to obtain a state-by-state transition probability matrix (TPM) for each RSN and subject condition. The TPM is the principal input necessary for calculating Φ , and there are two variants of it in PyPhi [1]:

1. **State-by-state TPM:** Each entry describes the probability that a certain state of the network transitions to another state in the subsequent time point (dimensions 32×32). This form is obtained directly from the time-series.
2. **State-by-node TPM:** Each entry describes the probability of a node flipping from (0 to 1 or 1 to 0) when the system is in a certain state (dimensions 32×5). In our procedure, the state-by-state TPM is obtained first and then converted to this form.

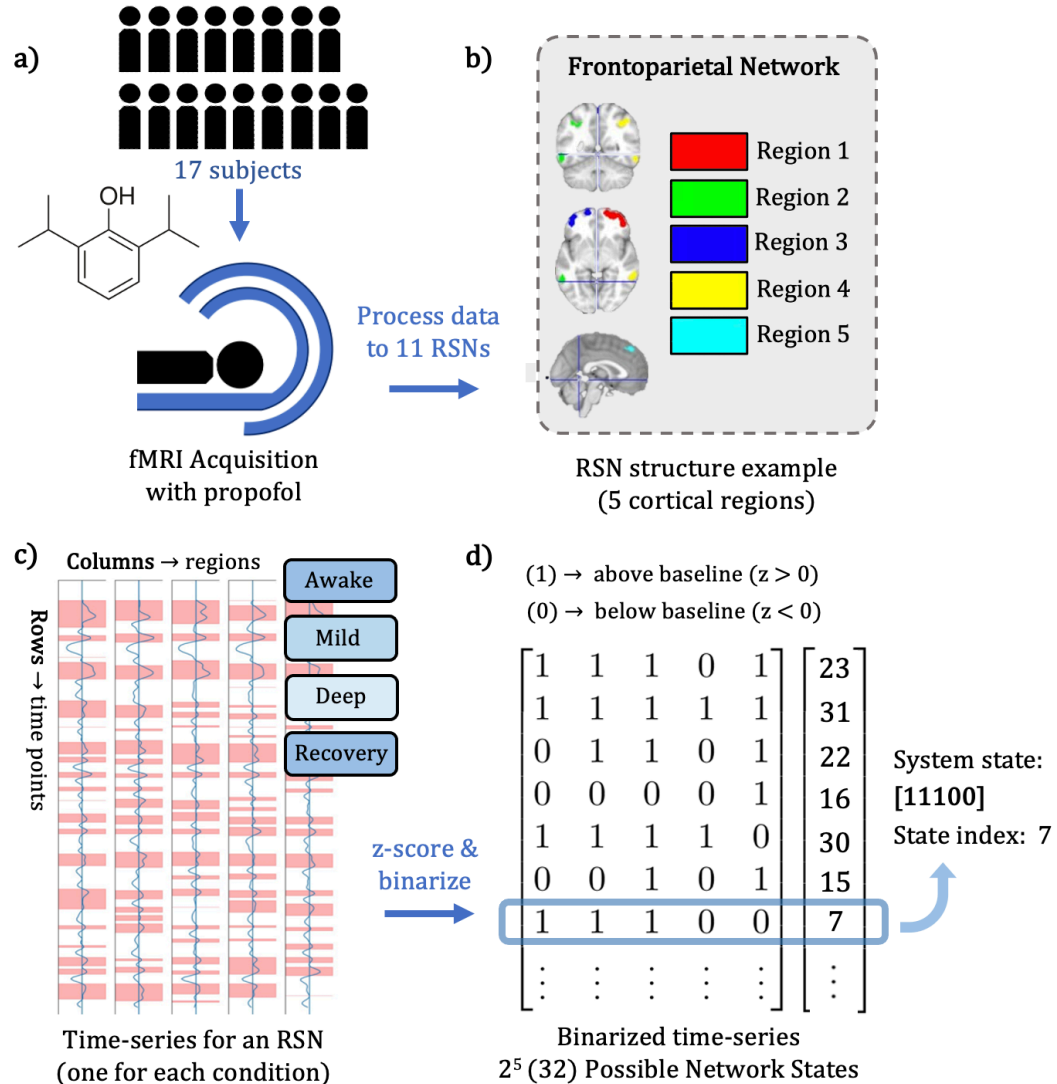


Figure 2.2.1. From fMRI acquisition to binarized time-series. a) Resting-state fMRI data was acquired for 17 subjects who underwent sedation with propofol; b) Data was processed to obtain 11 RSNs, each consisting of five regions (see Chapter 6.1.2 for image processing procedure). The frontoparietal network and its five constituent regions are shown as an example; c) A time-series array: each column corresponds to one of an RSN's five regions, and the rows correspond to time points (i.e., moving down corresponds to moving in time). For each subject and RSN, four time-series arrays were acquired, with one for each of the four conditions of awareness (awake, mild sedation, deep sedation, recovery). Time points highlighted in pink represent activity above the baseline, while those unhighlighted show activity below the baseline; d) The time-series were z-scored and binarized to obtain a binary representation for each region at a certain time point, leading to 32 possible states. Each time point's state was assigned an index as indicated to the right of the time-series array.

To generate a state-by-state TPM directly from a time-series, we counted the number of times each of the 32 possible states transitioned to any other state. The count of transitions was used to populate a 32×32 square matrix, in which row and column indices corresponded to little-endian state integers. To normalize the matrix, each row was divided by the number of times the state corresponding to its index appeared in the time series. For example, if 5 transitions occurred from state (1, 0, 0, 0, 0) (index = 1) to state (1, 1, 0, 0, 0) (index = 3), then the entry in row 1 and column 3 was assigned a value of 5. The row was then normalized with respect to the number of times the state occurred in the time-series; if the first state occurred 20 times, the entry would be set to equal 0.25, representing a 25% chance that the transition from state 1 to 3 occurs.

2.2.2. Concatenating time-series

Recall that the time-series for an individual subject in one of the four conditions consisted of 245 time points, each separated by the repetition time (TR). Considering the 32 possible states of the system, these time-series were relatively short, which was also problematic as it yielded sparse TPMs with significant inter-subject variation, resulting in spurious calculations of Φ . To address this issue, our solution was to concatenate, or link, the time-series of the entire subject population for each RSN. This resulted in longer time series, which we hoped would produce denser TPMs that provide a more complete description of the mechanisms involved in each network. The difference between single subject and concatenated TPMs is demonstrated in Figure 2.2.2.

Inevitably, there were limitations to this approach. In chapter 4, we conduct a thorough analysis on how concatenations affected calculations of Φ . As we will show, linking time-series from different subjects resulted in decreased Φ values compared to those obtained at the single subject level. Moreover, differences between subjects resulted in discontinuities within the elongated time-series that would not exist if a signal of the same length was acquired for a single subject. Ultimately, these limitations result from the low temporal resolution of fMRI (2-3 seconds between time points), which limits the length of the signal that can be obtained in a single scan. Nevertheless, we hoped that concatenating would produce TPMs where the predominant mechanisms

contributing to Φ are those likely to be found across several subjects and are hence intrinsic to a particular network. With the exception of our analysis on the modulation of propofol (see section 2.2.6), we concatenated all 17 subjects for each RSN and subject condition. With 11 RSNs and four conditions, we obtained a total of 44 time-series.

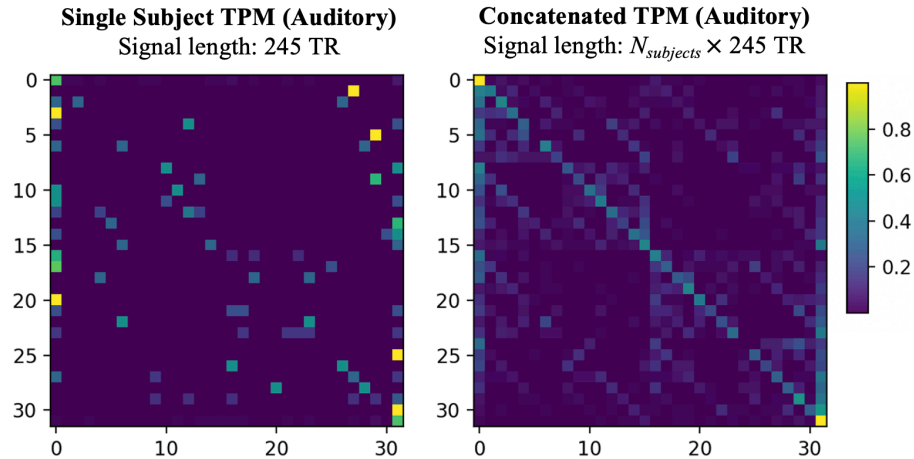


Figure 2.2.2. TPMs for the time-series of a single subject and 17 subjects. The TPM on the left, which was obtained from the time-series of the auditory network in the awake condition from a single subject, is sparse and varies significantly between subjects. The TPM on the right was obtained by concatenating the time-series of the same network and condition from all 17 subjects.

2.2.3. Statistical test for the Markov property

PyPhi was programmed to work for discrete Markovian systems consisting of binary elements, and its algorithms were written with the assumptions that a network's time series and TPM satisfy the following conditions [1,2]:

1. **The Markov property:** The subsequent state of the whole system depends only on its state in the current point in time, and not on any previous state (i.e., only the state at time point t has an impact on the state at $t+1$; the states at $t-1$, $t-2$, ... do not).
2. **Conditional independence:** The subsequent state of each individual element depends only on the current state of the other elements, meaning that mechanisms are not allowed to act instantaneously (i.e., only the state at t has an impact on elements at $t+1$; the state of other elements at $t+1$ do not).

To evaluate whether a time-series obeys the Markov property, we employed a statistical test that scans through the sequences of states appearing in each time-series to detect a potential violation [3–6]. For each time-series, we recorded the occurrence of the most commonly occurring 3-state sequence (states a, b, c ; $b \neq c$) as well as other 3-state sequences that begin with the same two states but end with a different one (i.e., a 3-state sequence where the first and second states are a, b). Note that these are the states of the whole network, meaning that each of a, b , and c correspond to a state index between 0 and 31. A violation of the Markov property is more likely if a particular 3-state sequence occurs at a greater frequency than other sequences that start with the same two states; if this is the case, it indicates that the third state in the sequence depends on two previous time points, which is not allowed for a Markovian system.

We assigned a label h ($h = 1, 2, 3, \dots, H$) to each of the three-state sequences mentioned above, and the following quantities were obtained for each one:

1. **Sequence count, SC :** The occurrence of the 3-state sequence in the time-series.
2. **Non-sequence count, NSC :** The count of all 2-state sequences that match the first two states of h sequence but are followed by a different third state ($a, b, e, e \neq c$).
3. **Total sequence count, TSC :** The count of all 2-state sequences that correspond to the first two states of the 3-state sequence, including those followed by the third state in the sequence of interest ($a, b, c; a, b, e; TSC = SC + NSC$).

These quantities were then organized into a contingency table of the type shown below. Each recorded count is referred to as N_{hk} , where h and k are row and column indices, respectively. As described above, the index h ranges between 1 and the total number of 3-state sequences recorded in the table, H , while the index k can equal 1 or 2 for the SC and NSC columns, respectively. N_{hT} represents a row total, N_{Tk} represents a column total, and n represents the total number of counts in the table.

The table was used to compute a χ^2 value (equation 2.2.1), which indicates the extent to which the network's temporal behavior deviates from the Markov property. This can be seen as a comparison between the state-sequence distribution of the time-series

and that of a state sequence distribution that satisfies the Markov property (i.e., this is essentially a “goodness of fit” test for the Markov property).

$$\chi^2 = \sum_{h,k} \frac{[N_{hk} - n(N_{hT}/n)(N_{Tk}/n)]^2}{n(N_{hT}/n)(N_{Tk}/n)} \quad 2.2.1$$

This χ^2 value was then used to obtain a two-tailed p -value with the SciPy statistics module in Python (<https://docs.scipy.org/doc/scipy/reference/stats.html>). Significance was set to 0.05 and the number of degrees of freedom ν was $H - 1$. Statistical significance indicated a significant deviation from the Markov property. Finally, we also obtained a reduced χ^2 statistic, or χ^2_ν , to analyze the quality of each fit. The results obtained for each RSN’s time-series (with 17 subjects concatenated) are presented in Chapter 6.2.1.

<i>h</i>	<i>SC</i>	<i>NSC</i>	<i>TSC</i>
1	N_{11}	N_{12}	N_{1T}
2	N_{21}	N_{21}	N_{2T}
3	N_{31}	N_{32}	N_{3T}
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
<i>H</i>	N_{H1}	N_{H2}	N_{HT}
Total	N_{T1}	N_{T2}	<i>n</i>

Table 2.2.1. Contingency table for counting the sequences of states in a time-series. A contingency table of this form was constructed to compare the most common three-state sequence to other sequences that share the last two states (i.e., 1-0-1, 2-0-1, 3-0-1, ..., 31-0-1). Rows correspond to a specific three-state sequence (*h*). The middle columns correspond to counts of the three-state sequences and the occurrences of two state-sequences that start with the same two states (*SC* and *NSC*). The third column tallies the totals of each row. The table was used to compute a χ^2 statistic, which indicates the extent to which a time-series deviates from the Markov property.

2.2.4. Test for conditional independence test

Next, we tested whether the networks satisfy the conditional independence property. PyPhi has built-in functions to check whether a state-by-state TPM satisfies conditional independence. For every state-by-node TPM, there is one unique state-by-state TPM that is conditionally independent, which is outputted by PyPhi when converting a TPM from state-by-node to state-by-state form [1]. We checked the extent to which our data met this property by taking the original state-by-state TPM, converting it to state-by-node form, and then back to state-by-state to obtain the conditionally independent version. The relative distance D between the original TPM (\mathbf{A}) and the conditionally independent TPM (\mathbf{B}) was computed by subtracting the two arrays and calculating the Frobenius norm (\mathcal{F}) of the residual, which was divided by the Frobenius norm of the conditionally independent variant:

$$\mathcal{F}(\mathbf{A}) = \sqrt{\sum_i \sum_j |A_{i,j}|^2} \quad 2.2.2$$

$$D(\mathbf{A}, \mathbf{B}) = \frac{\mathcal{F}(\mathbf{A} - \mathbf{B})}{\mathcal{F}(\mathbf{B})} \quad 2.2.3$$

2.2.5. Calculating $\mu[\Phi^{\max}]$ for a time-series

To calculate Φ , the TPM of the concatenated signals was converted to a state-by-node form and submitted to PyPhi to generate a network class, which defines a system's mechanisms and elements. Φ is calculated for a particular subsystem S of the network class, which comprises of 1) the state of the system at a given time, which sets the necessary background conditions for each element, and 2) the subset of nodes to be included in the subsystem.

In a complete analysis, Φ is computed for every possible subset of a network's regions, and the system's Φ^{\max} at a certain state is defined as the maximum value obtained from the Φ values of all subsets. A subset must consist of at least two elements and can contain as many as five (the whole network). As the state of each RSN varies throughout its time series, we calculated Φ^{\max} for every state and obtained a weighted

average, which we refer to as $\mu[\Phi^{\max}]$. The contribution, or weight, of each state was based on the frequency of its occurrence in the time-series. Calculating $\mu[\Phi^{\max}]$ allowed us to use a single metric to compare the integrated information of all networks and states of awareness. The process used to compute $\mu[\Phi^{\max}]$ is presented in the schematic below.

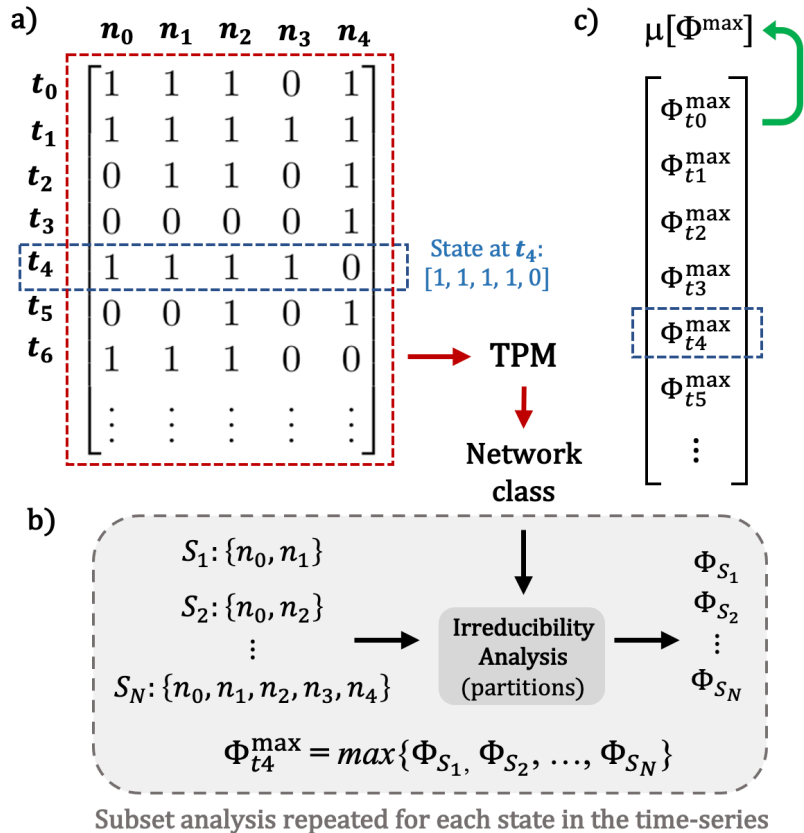


Figure 2.2.3. Computing average integrated information for a time-series. a) The TPM is obtained from a time-series and used to generate a network class in PyPhi (nodes/regions are labeled as n_0, n_1, \dots, n_4); b) For a specific state (in this case, the state at time point t_4), a subset analysis is performed. The process begins with defining each subset consisting of at least 2 elements leading up to the entire system (S_1, S_2, \dots, S_N), with N being the total number of subsets. The irreducibility analysis outlined in Chapter 1.3, where the system is partitioned to obtain integrated information Φ , is performed for each subset, leading to a collection of Φ values ($\Phi_1, \Phi_2, \dots, \Phi_N$). The maximum value of Φ , Φ^{\max} , is then obtained from this set of values; c) The subset analysis is repeated for each state in the time-series to obtain Φ^{\max} for every time point. Finally, a weighted average of Φ^{\max} is computed based on the frequency of each state in the time-series.

2.2.6. Analysis of modulation by propofol

In our central analysis, we measured sedative-induced changes to integrated information within each individual RSN. To allow for a statistical analysis of these changes, we did not use the 44 time-series obtained by concatenating 17 subjects, as those only had one measurement for each RSN and condition. Instead, we created sampling distributions of $\mu[\Phi^{\max}]$ as follows: for each RSN and condition, 17 time-series were generated by concatenating signals from 16 subjects in random order, leaving a different subject out each time. A sample of $\mu[\Phi^{\max}]$ values was then obtained, and the mean of this sample, m , was taken to be the mean of the population. The standard deviation of the sample was taken to be the standard error, which was multiplied by the square root of the sample size ($\sqrt{17}$) to obtain the standard deviation of the population, σ [7].

We tested for statistically significant differences between the four conditions using Welch's t -test [8], which is used for two samples with different variances and sizes. Although the sample size N was the same for each condition, the variance of $\mu[\Phi^{\max}]$ was assumed to be different. Within the same network, t values were computed for every possible pair of conditions i and j (e.g., awake vs. mild sedation):

$$t = \frac{m_i - m_j}{\sqrt{N^{-1}(\sigma_i^2 + \sigma_j^2)}} \quad 2.2.4$$

We then computed the number of degrees of freedom ν using the Welch-Satterthwaite equation (simplified for two samples of the same size) [8]:

$$\nu = (N - 1) \frac{(\sigma_i^2 + \sigma_j^2)^2}{\sigma_i^4 + \sigma_j^4} \quad 2.2.5$$

Finally, we computed a two-tailed p -value using the t -distribution obtained for each comparison. Through these comparisons, our goal was to see which networks demonstrate a modulation pattern that corresponds to the conscious state of subjects:

$$\mu[\Phi^{\max}]_{\text{awake}} > \mu[\Phi^{\max}]_{\text{mild}} > \mu[\Phi^{\max}]_{\text{deep}} < \mu[\Phi^{\max}]_{\text{recovery}}$$

2.2.7. Interregional correlations

Throughout our analysis, we introduced the average Pearson correlation of each RSN's time-series as a reference point for our results with $\mu[\Phi^{\max}]$. For every pair of regions within an RSN (regions a and b), the Pearson correlation coefficient ρ was computed between their time-series (ts_a and ts_b) using the SciPy statistics module. The average of all correlations, $\mu[\rho]$, was then obtained by adding up all the correlations and dividing by 10 (total number of possible pairs of regions out of 5 regions):

$$\mu[\rho] = \frac{1}{10} \sum_{a \neq b} \rho(ts_a, ts_b) \quad 2.2.6$$

2.2.8. Spatial and temporal control procedures

As a control procedure, we tested whether the RSNs in question generate more integrated information than a random collection of five regions. To do this, we took the concatenated time-series of the 11 RSNs and randomly distributed the individual nodes (columns of the time-series array) into 100 control networks, ensuring all nodes within the new networks came from different RSNs. We then calculated $\mu[\Phi^{\max}]$ for each new time-series. This was repeated for every conscious condition to obtain four control distributions, to which we then compared the $\mu[\Phi^{\max}]$ values of the original RSNs.

In the second control procedure, the original grouping of each RSN's cortical regions remained unchanged, but the order of each signal's time points was randomly permuted. All time points were shuffled individually, which completely reordered each time-series' sequences of states and introduced a high level of disorder to the signals. This was achieved by switching the positions of two randomly selected time points 10^5 times, which we repeated 50 times for each network and condition to obtain temporal control distributions of $\mu[\Phi^{\max}]$. Ultimately, the goal of these procedures was to point out any intrinsic dependencies of an RSN's integrated information on its spatial and temporal structure.

2.3. Results

2.3.1. Markov property and conditional independence

For almost all time-series acquired by concatenating the entire subject population, the test for the Markov property did not show statistically significant differences between the observed sequence distribution and the reference (Markovian) sequence distribution. Out of the 44 time-series tested, only three yielded $p < 0.05$: the retrosplenial network in the recovery condition, the ventral attention network in the awake condition, and the cingulo-parietal network in awake condition. Out of the remaining 41 time-series, the reduced chi-squared values ranged between 0.26 and 1.50, with most values between 0.7 and 1.3, indicating a good quality fit of the Markovian model to the state-sequence distributions of our data. The chi-squared values, degrees of freedom, and significance tests obtained for each time-series are provided in Chapter 6 (supplementary material).

For the conditional independence property, we found that the relative difference between the raw (original) TPM and its conditionally independent counterpart ranged between 22% to 75% depending on the network and condition in question. The networks with the lowest deviations were the default mode, frontoparietal, and dorsal attention networks, while those with the highest deviations were the visual, sensorimotor mouth, and cingulo-parietal networks. All relative differences are shown in Table 2.3.1.

	Awake	Mild	Deep	Recovery
Frontoparietal	36.97	27.48	21.81	39.22
Default Mode	43.49	33.75	34.46	51.52
Retrosplenial	57.36	50.75	49.34	59.69
Dorsal Attention	41.56	37.85	34.53	43.9
Ventral Attention	42.94	33.44	35.98	40.5
Cingulo-Opercular	42.27	33.93	31.14	41.9
Cingulo-Parietal	67.85	68	56.85	65.12
SM Hand	51.53	62.68	60.34	56.74
SM Mouth	68.59	61.28	51.44	64.93
Auditory	62.24	53.76	54.73	65.36
Visual	74.88	62.24	60	71.37

Table 2.3.1. Relative distance between original TPMs and their conditionally independent versions. The relative residual between the TPMs generated directly from the time-series and their conditionally independent counterparts are given as percent differences for each RSN (row) and condition (column).

2.3.2. Modulation of $\mu[\Phi^{\max}]$ by propofol

Our central analysis focused on how the administration of propofol impacts the integrated information generated within each RSN. The $\mu[\Phi^{\max}]$ distributions we obtained by concatenating 16 subjects (with a different subject excluded for each sample) are shown in Figure 2.3.1. Each network is presented in a subplot that includes four bars for each condition. The RSNs are arranged according to their association with “higher-order” cortical functions and conscious processing. For visual simplicity, statistically significant differences are shown only for comparisons that include the awake condition (all t and p -values are provided in Chapter 6.2.1).

The networks that most clearly demonstrated the expected propofol-induced modulation pattern were the frontoparietal (FPN) and dorsal attention networks (DAN) ($p < 0.001$ for awake vs. deep & awake vs. recovery in both RSNs). While both demonstrated a significant increase moving from deep sedation to recovery, the FPN underwent a remarkably large jump in this transition. A similar behavior was seen for the default mode network (DMN), albeit at significantly lower $\mu[\Phi^{\max}]$ values. In each of these instances, no significant difference was observed between the awake and mild sedation conditions.

Several RSNs demonstrated small and statistically insignificant fluctuations across the four conditions. Out of the higher-order networks, these included the retrosplenial, ventral-attention, and cingulo-opercular networks. Although the cingulo-parietal network may appear to fall in this category, its relative margins of error are very large, and it is hence not possible to accurately describe its modulation. Turning to the four motor and sensory cortices, the sensorimotor hand network (SM hand) came closest to matching the conscious evolution of subjects, as it demonstrated a substantial drop in deep sedation followed by an increase in recovery. While the auditory network underwent significant changes to $\mu[\Phi^{\max}]$, its modulation is inconsistent with any of the other behaviors described. On the other hand, the sensorimotor mouth and visual cortices did not demonstrate any meaningful sedative-induced changes.

There are also notable differences in the magnitudes of $\mu[\Phi^{\max}]$ across different networks. Out of all RSNs, the two sensorimotor networks generated the greatest amount of integrated information. In the higher-order group, the RSNs with the greatest $\mu[\Phi^{\max}]$ were the frontoparietal and retrosplenial networks, while the ones with the lowest values were the DMN and cingulo-opercular network. Altogether, there was substantial heterogeneity in the magnitude and variation of $\mu[\Phi^{\max}]$ among the networks.

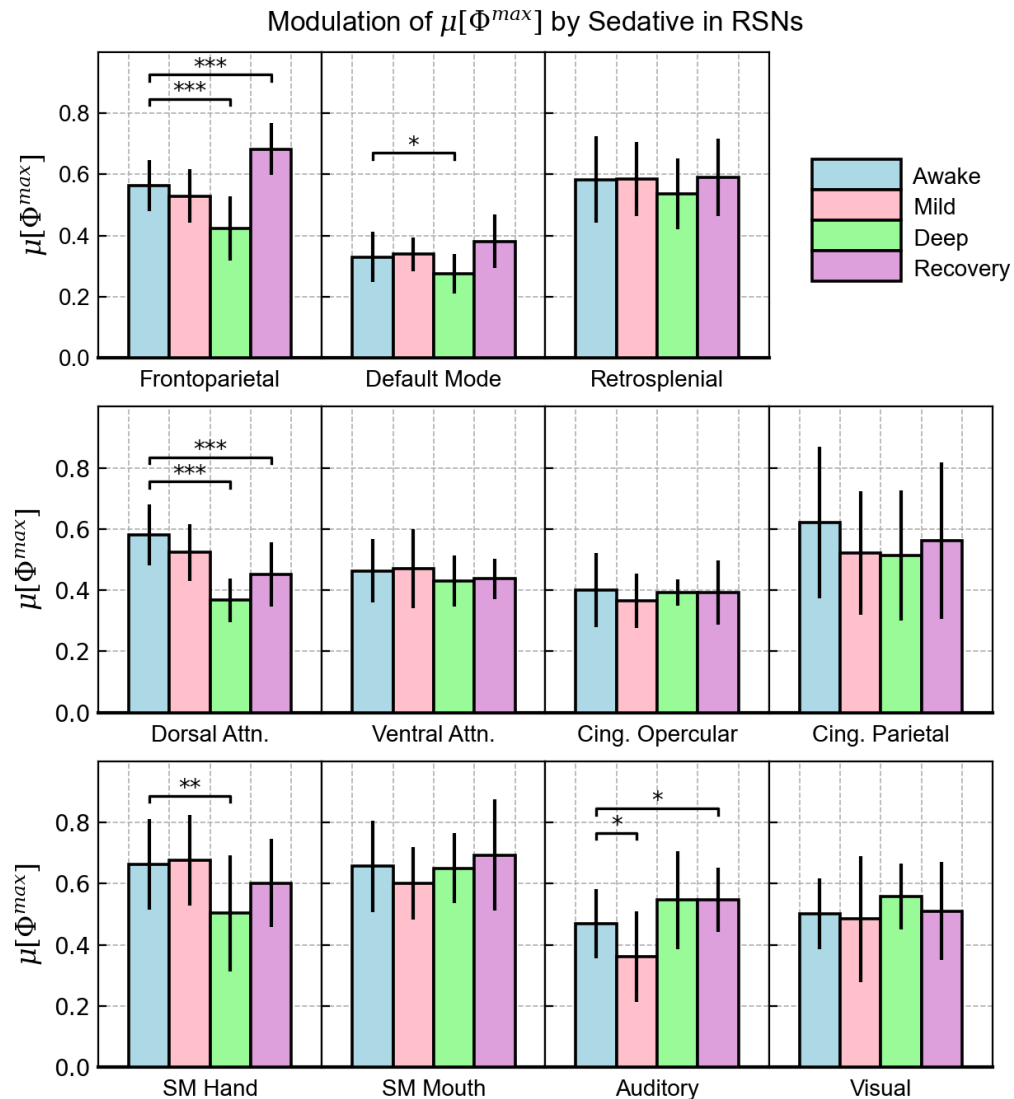


Figure 2.3.1. Modulation of $\mu[\Phi^{\max}]$ by propofol in individual RSNs. For each network, $\mu[\Phi^{\max}]$ is compared for the four conditions. Bar heights represent the mean of the $\mu[\Phi^{\max}]$ distribution and error bars represent standard deviations. Statistically significant differences are shown for all comparisons that involve the awake condition ($0.01 < p < 0.05$: *, $0.001 < p < 0.01$: **, $p < 0.001$: ***).

We also computed average interregional correlations, $\mu[\rho]$, for the same set of time-series (16 subjects concatenated, 1 left out), which we present in Figure 2.3.2. In contrast to $\mu[\Phi^{\max}]$, the sedative-induced modulation of this measure was much more consistent among the RSNs, nearly all of which demonstrate a significant decrease moving from awake to mild sedation and mild to deep sedation, followed by an increase in recovery. Accordingly, the heterogeneity observed for $\mu[\Phi^{\max}]$ across the RSNs, which can be seen as a feature differentiating the networks from each other, is not demonstrated by variations in $\mu[\rho]$.

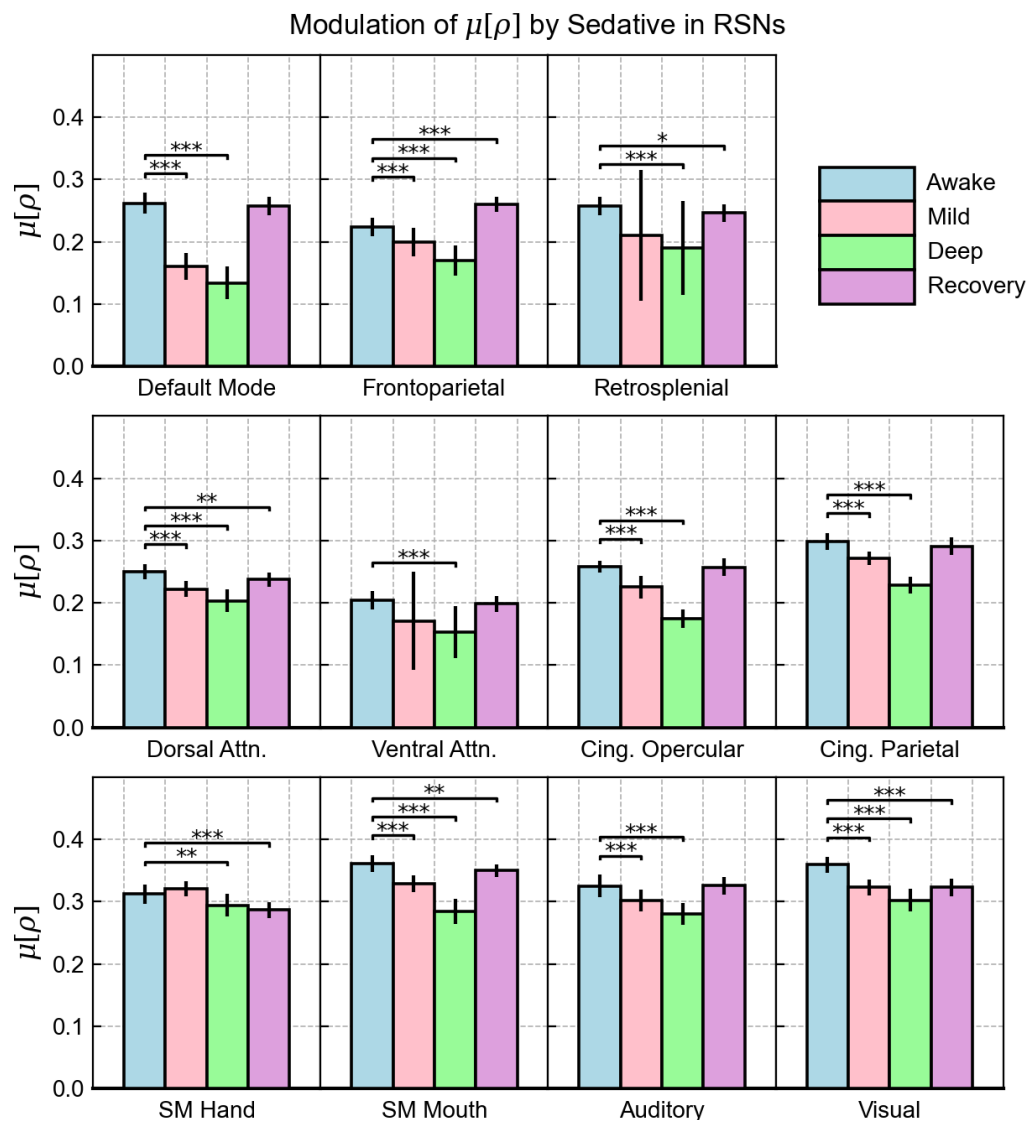


Figure 2.3.2. Modulation of $\mu[\rho]$ by propofol in individual RSNs. The analysis in Figure 2.3.1 was repeated for each RSN's average interregional correlations.

The only exception to this general trend was the sensorimotor hand network, whose correlations increased during mild sedation and decreased during recovery. Another abnormality is that the retrosplenial and ventral attention networks had an unusually large variance in the mild and deep sedation conditions. While these RSNs followed the universal behavior with their average correlation, this large margin of error may reflect significant inter-subject variation in the acquisition procedure.

2.3.3. Spatial and temporal control procedures

The following control procedures were based on the time-series acquired by concatenating all 17 subjects. Figure 2.3.3 presents a comparison between the $\mu[\Phi^{\max}]$ distributions of the control networks, which consisted of random regions from different RSNs, and the $\mu[\Phi^{\max}]$ values of the 11 concatenated RSNs for the four conditions.

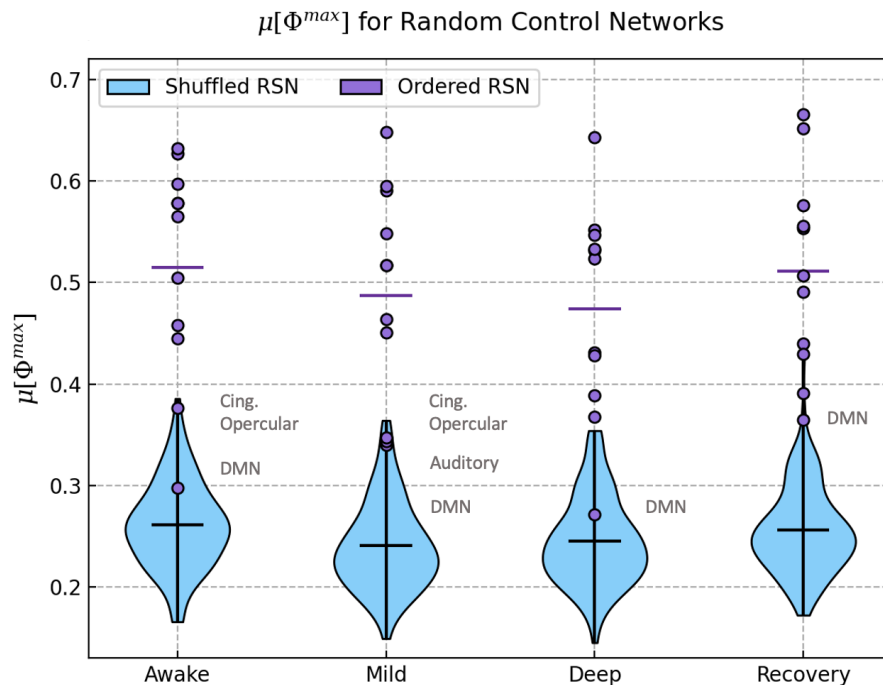


Figure 2.3.3. Comparison of $\mu[\Phi^{\max}]$ between RSNs and spatial control networks. For each of the four conditions, the $\mu[\Phi^{\max}]$ distribution of networks generated by shuffling nodes (blue violins) is shown together with the $\mu[\Phi^{\max}]$ values of the original RSNs (purple dots). The means of both distributions are indicated with horizontal bars. Text adjacent to the violins indicates networks that fall within the distribution of random networks, which are those with lower $\mu[\Phi^{\max}]$ relative to other RSNs (default mode, cingulo-opercular, auditory).

Most of the original RSNs fell well above the mean $\mu[\Phi^{\max}]$ of the control networks. Three RSNs, whose $\mu[\Phi^{\max}]$ were relatively low compared to the rest, fell within the control distributions. These were the default mode (awake, mild, deep, recovery), cingulo-opercular (awake, mild), and auditory (mild) networks. Otherwise, this indicates that the majority of RSNs, through their intrinsic composition, produce more integrated information than networks generated in random.

A similar comparison is shown for the reference correlation metric, $\mu[\rho]$, in Figure 2.2.4. Similarly, randomly grouping cortical regions significantly reduced the correlational strengths of the original RSNs. This result is not surprising, as the networks were constructed using correlation-based functional connectivities in the first place.

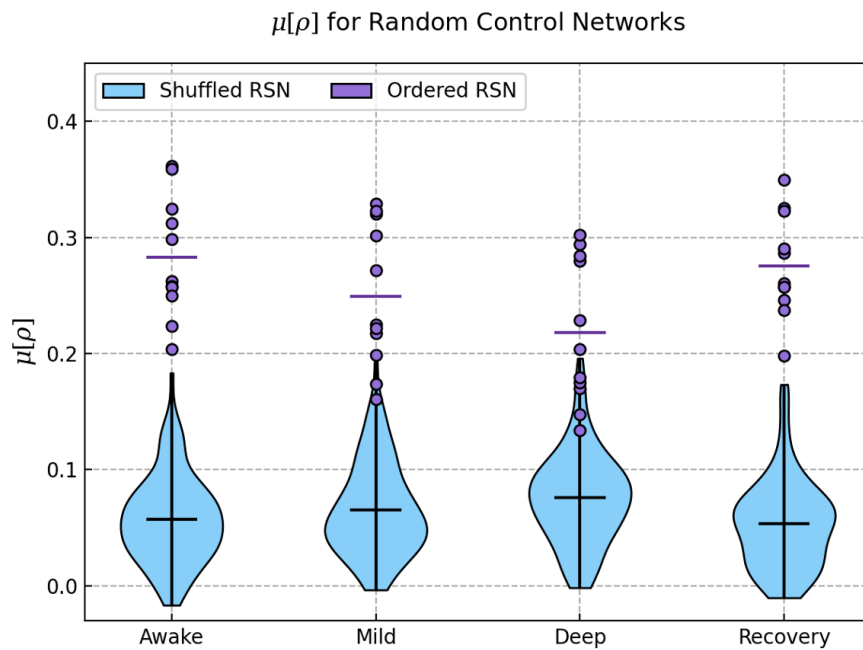


Figure 2.3.4. Comparison of $\mu[\rho]$ between RSNs and spatial control networks. The analysis in Figure 2.3.3 was repeated for average interregional correlations.

In the temporal control procedure, we permuted each signal's time points to generate a collection of new time-series with completely different state transitions. The $\mu[\Phi^{\max}]$ distributions we obtained are presented in the bottom panel of Figure 2.3.5. While $\mu[\Phi^{\max}]$ ranged between 0.3 and 0.7 for the original (unpermuted) signals, the values obtained for the controls ranged between 0.03 and 0.1. This drastic decrease

underscores the importance of the temporal structure in the fMRI data we obtained. We also present the correlations obtained in top panel of Figure 2.3.4. These values correspond to both the original and shuffled signals; unlike $\mu[\Phi^{\max}]$, $\mu[\rho]$ remained unchanged when the signals were permuted.

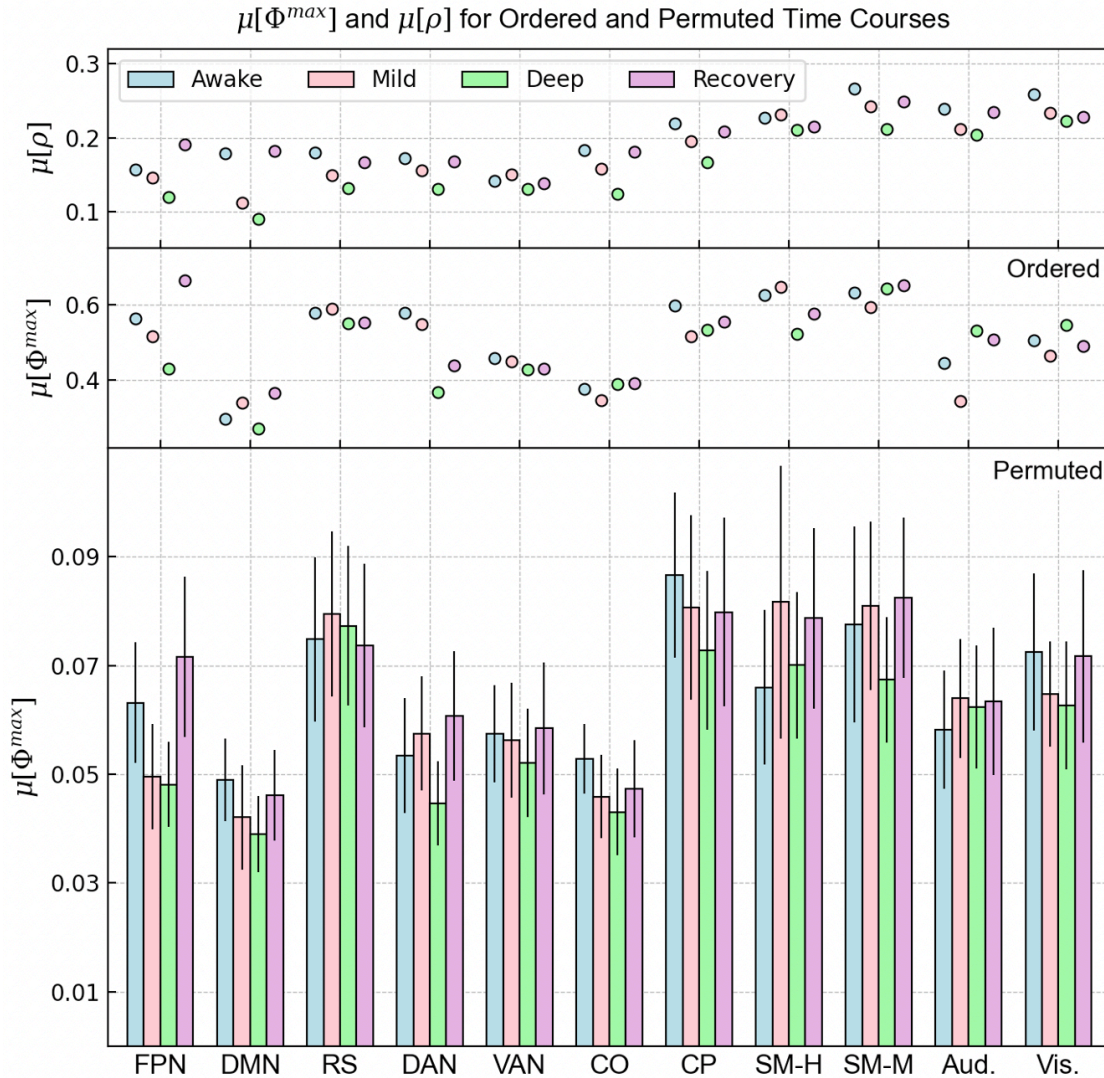


Figure 2.3.5. Distributions of $\mu[\Phi^{\max}]$ and $\mu[\rho]$ obtained for temporal controls. Each RSN is presented with abbreviations across the horizontal axis. The bottom panel presents the $\mu[\Phi^{\max}]$ distributions of the permuted time-series (50 measurements for each RSN and condition); bar heights correspond to the mean of the controls, and error bars represent their standard deviation. The middle panel shows the $\mu[\Phi^{\max}]$ of the unpermuted time-series. The top panel presents the $\mu[\rho]$ values for *both* the ordered and permuted time-series; while $\mu[\Phi^{\max}]$ dropped drastically for the controls, correlations remained unchanged.

2.4. Discussion

We started our analysis with tests for the Markov property and conditional independence, which are assumed to be satisfied in the mathematics underlying PyPhi's algorithms. While the first test showed that the Markov property was mostly obeyed, we saw that the TPMs obtained directly from the fMRI signals were not conditionally independent. In other words, while the entire system exhibited a Markovian behavior, the individual elements themselves had instantaneous causal properties. Calculations of Φ^{\max} are based on the conditionally independent variant of the TPM, and as a consequence of converting to this form, some of the features contained in the original signals were lost. Nevertheless, we believe that the network properties preserved in the conditionally independent variants could still be used for meaningful calculations of integrated information, especially for networks with lower deviations from this property.

Our central investigation focused on each network's behavior as subjects transitioned throughout the four states of sedation. In general, the effects of anesthesia differed among the RSNs. The variation of $\mu[\Phi^{\max}]$ that corresponds to the conscious state of subjects is a decrease moving from wakefulness to deep sedation, followed by an increase in the recovery condition. Such modulation was most clearly observed in the frontoparietal and dorsal attention networks (FPN and DAN, respectively). The FPN, also known as the central executive network, is a crucial hub for cognitive control and goal-oriented behavior [9,10]. It is said to connect external stimuli with "stored internal representations", and hence plays an important role in integrating inputs from other cortical regions [11]. Accordingly, its drop in $\mu[\Phi^{\max}]$ coincides with the loss of these functions as subjects become deeply sedated. Interestingly, its integrated information peaked during the recovery stage, which could indicate an important integrative role in the recovery of consciousness.

The DAN demonstrated a very similar behavior to the FPN, but without a drastic increase in the recovery condition. This network facilitates top-down, voluntary mediation of external attention that is primarily associated with visuospatial orienting [12,13]. As with the FPN, its reduced $\mu[\Phi^{\max}]$ coincides with the loss of these voluntary

functions as subjects become unresponsive. Considering the related functions and previously observed interactions between the FPN and DAN [10], the similarities they demonstrate through integrated information is a promising result.

It is also important to discuss the default mode network (DMN), which is widely investigated in neuroscientific literature for its role in introspective cognition [14]. Interestingly, the DMN presented the lowest average integrated information for all conditions. This may indicate that despite a high metabolic activity at rest, the DMN's internal integrated information as an isolated system is intrinsically low. Perhaps, an analysis of its interactions with other cortical networks may be necessary to better understand its integrative mechanisms [15,16]. Like the FPN and DAN, it demonstrated a significant drop in $\mu[\Phi^{\max}]$ during deep sedation, but with no decrease moving from wakefulness to mild sedation. Previous studies found that some regions of the DMN resist a suppression of activity during mild propofol sedation but not under deep anesthesia [17,18], which may be reflected in the modulation we observed.

On the other hand, no meaningful sedative-induced changes were observed in the ventral attention network. Like the DAN, this counterpart has an important role in sensory orienting, although its function is associated with involuntary shifts of attention directed at unexpected stimuli [13]. The lack of modulation may therefore be appropriate, as the sort of stimulation necessary for its integration of external stimuli was absent in our resting-state acquisition procedure. Minor and insignificant fluctuations of $\mu[\Phi^{\max}]$ were also observed in the retrosplenial and cingulo-opercular networks. Although some significant changes were demonstrated by the sensory RSNs, we did not find any consistent relationship between their integrated information and the condition of subjects.

As explained in Chapter 1, propofol is known to preferentially suppress higher-order processing regions in the frontal and prefrontal cortices, which are commonly attributed to awareness [19]. In our attempt to quantify “consciousness” using integrated information, $\mu[\Phi^{\max}]$ most closely reflected the conscious evolution of subjects in the FPN and DAN, both of which include frontal and prefrontal regions. Our results are also supported by findings that propofol inhibits functional connections between the frontal

and parietal cortices, which are essential to the FPN's integrative pathways [20]. At the same time, propofol was not found to diminish the reactivity of sensory cortices to stimulation, which may explain the lack of modulation in $\mu[\Phi^{\max}]$ for these RSNs [19].

We also note that while the RSNs demonstrated a variety of modulations to their integrated information, the behavior of their correlations was much more homogenous. In almost all cases, the average Pearson correlation $\mu[\rho]$ decreased during deep sedation and increased in recovery. Although this is consistent with the conscious evolution of subjects, it suggests that all networks equally reflect conscious experiences. This cannot be the case, as it is well-established that certain RSNs have more important roles than others in consciousness. Instead, the homogeneity of correlations across networks may reflect a global neurophysiological effect that is induced by the sedative [21].

In the control procedures, we showed that the BOLD signal of each RSN generates an intrinsic level of integrated information that arises from each network's spatial composition and time-dependent behavior. While $\mu[\Phi^{\max}]$ was subject to a drastic decrease whenever an RSN's time-series was permuted, doing the same had no effect on $\mu[\rho]$. This shows that our measure of integrated information can capture an RSN's causal structures and differentiate its time-dependent behavior from that of a time-series generated in random. On the other hand, correlational descriptions of activity appear to be insensitive to a network's causal properties. The insensitivity of $\mu[\rho]$ to permutations, as well as the global behavior of this metric in nearly all networks, supports the argument that traditional measures of functional connectivity may not be substantial to effectively analyze fMRI data [22].

2.5. References

- [1] Mayner WGP, Marshall W, Albantakis L, Findlay G, Marchman R, Tononi G. PyPhi: A toolbox for integrated information theory. *PLOS Computational Biology* 2018;14:e1006343. <https://doi.org/10.1371/journal.pcbi.1006343>.
- [2] Krohn S, Oswald D. Computing integrated information. *Neuroscience of Consciousness* 2017;2017. <https://doi.org/10.1093/nc/nix017>.
- [3] Anderson TW, Goodman LA. Statistical Inference about Markov Chains. *The Annals of Mathematical Statistics* 1957;28:89–110. <https://doi.org/10.1214/aoms/1177707039>.
- [4] Billingsley P. Statistical Methods in Markov Chains. *The Annals of Mathematical Statistics* 1961;32:12–40. <https://doi.org/10.1214/aoms/1177705136>.
- [5] Hiscott RN. Chi-square tests for markov chain analysis. *Journal of the International Association for Mathematical Geology* 1981;13:69–80. <https://doi.org/10.1007/BF01032010>.
- [6] Monika Skuriat-Olenchnowska. Statistical inference and hypothesis testing for Markov chains with Interval Censoring. 2005.
- [7] Altman DG, Bland JM. Standard deviations and standard errors. *BMJ* 2005;331:903. <https://doi.org/10.1136/bmj.331.7521.903>.
- [8] Ruxton GD. The unequal variance t-test is an underused alternative to Student’s t-test and the Mann–Whitney U test. *Behavioral Ecology* 2006;17:688–90. <https://doi.org/10.1093/beheco/ark016>.
- [9] Miller EK, Cohen JD. An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience* 2001;24:167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>.
- [10] Dixon ML, de La Vega A, Mills C, Andrews-Hanna J, Spreng RN, Cole MW, et al. Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proceedings of the National Academy of Sciences* 2018;115:E1598–607. <https://doi.org/10.1073/pnas.1715766115>.
- [11] Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL. Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. *Journal of Neurophysiology* 2008;100:3328–42. <https://doi.org/10.1152/jn.90355.2008>.
- [12] Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience* 2002;3:201–15. <https://doi.org/10.1038/nrn755>.
- [13] Vossel S, Geng JJ, Fink GR. Dorsal and Ventral Attention Systems. *The Neuroscientist* 2014;20:150–9. <https://doi.org/10.1177/1073858413494269>.
- [14] Raichle ME. The Brain’s Default Mode Network. *Annual Review of Neuroscience* 2015;38:433–47. <https://doi.org/10.1146/annurev-neuro-071013-014030>.

- [15] Di X, Biswal BB. Modulatory interactions between the default mode network and task positive networks in resting-state. *PeerJ* 2014;2:e367. <https://doi.org/10.7717/peerj.367>.
- [16] Gordon EM, Laumann TO, Marek S, Raut R v., Gratton C, Newbold DJ, et al. Default-mode network streams for coupling to language and control systems. *Proceedings of the National Academy of Sciences* 2020;117:17308–19. <https://doi.org/10.1073/pnas.2005238117>.
- [17] Qiu M, Scheinost D, Ramani R, Constable RT. Multi-modal analysis of functional connectivity and cerebral blood flow reveals shared and unique effects of propofol in large-scale brain networks. *NeuroImage* 2017;148:130–40. <https://doi.org/10.1016/j.neuroimage.2016.12.080>.
- [18] Saxena N, Gili T, Diukova A, Huckle D, Hall JE, Wise RG. Mild Propofol Sedation Reduces Frontal Lobe and Thalamic Cerebral Blood Flow: An Arterial Spin Labeling Study. *Frontiers in Physiology* 2019;10. <https://doi.org/10.3389/fphys.2019.01541>.
- [19] Liu X, Lauer KK, Douglas Ward B, Roberts C, Liu S, Gollapudy S, et al. Propofol attenuates low-frequency fluctuations of resting-state fMRI BOLD signal in the anterior frontal cortex upon loss of consciousness. *NeuroImage* 2017;147:295–301. <https://doi.org/10.1016/j.neuroimage.2016.12.043>.
- [20] Hudetz AG. General Anesthesia and Human Brain Connectivity. *Brain Connectivity* 2012;2:291–302. <https://doi.org/10.1089/brain.2012.0107>.
- [21] Brown EN, Lydic R, Schiff ND. General anesthesia, sleep, and coma. *The New England Journal of Medicine* 2010;363:2638–50. <https://doi.org/10.1056/NEJMra0808281>.
- [22] Mohanty R, Sethares WA, Nair VA, Prabhakaran V. Rethinking Measures of Functional Connectivity via Feature Extraction. *Scientific Reports* 2020;10:1298. <https://doi.org/10.1038/s41598-020-57915-w>.

Chapter 3 : Resting State Networks and their Temporal Structure

3.1. Overview

At its root, integrated information theory (IIT) is a framework based on causality; in Chapter 1, we saw how the “information” aspect of IIT is measured directly from the causes and effects of a mechanism. It has been previously argued that due to its low temporal resolution, causality-based metrics may be unsuitable in fMRI because many neuronal processes happen at a much faster rate than the 2-3 second repetition time (TR) separating images [1]. While fMRI may be unable to capture high frequency processes, research on low-frequency activity is a subject of continuing research. There is substantial evidence to suggest that a neural basis exists for BOLD fluctuations captured in the 0.008 - 0.1 Hz range, for which the temporal resolution of fMRI is adequate [2,3].

We began to address the importance of the BOLD signal’s temporal structure in the control procedure of Chapter 2. When permuted, the $\mu[\Phi^{\max}]$ value of each RSN’s concatenated time-series was subject to a ten-fold decrease. This suggested an inherent causality in the signals we measured, which we were unable to capture with interregional correlations. In this chapter, we extend this procedure and apply a series of permutations that vary in their extent of disruption to the original time-series. After generating new signals and measuring their $\mu[\Phi^{\max}]$ values, we aim to characterize how rapidly integrated information drops with respect to the “strength” of the permutation performed (i.e., the extent to which the signal’s original order of states is disrupted). To quantify this rate, we introduce λ , a fit parameter that represents the exponential decay of $\mu[\Phi^{\max}]$.

Using these permuted signals, we also analyze how $\mu[\Phi^{\max}]$ relates to the fluctuation rate of a time-series. Since permutations reorder a signal randomly, the permuted signals can be seen as having a higher rate of random fluctuations than the original time-series. We measure the effects of each permutation using a property referred to as the “diagonal strength” of the signal’s corresponding TPM. By analyzing the relationship between diagonal strength and $\mu[\Phi^{\max}]$, we aim to point out which temporal properties of the BOLD signal are associated with high levels of integrated information.

3.2. Methods

3.2.1. Grouped permutations

For each RSN and condition, we started with the concatenated time-series containing all 17 subjects. While the control procedure consisted of shuffling time points individually, the following analysis extended this idea by permuting each network's time-series in grouped blocks of time points. The order of states within these blocks was unchanged, but their position within the time-series was shuffled in random. By varying the sizes of these blocks, we aimed to analyze how $\mu[\Phi^{\max}]$ varies with different degrees of disruption to the original time-series.

The block sizes used, which we refer to as n , were 2, 3, 5, 7, 15, and 35 time points. We also included the individual time point permutations from the control procedure, for which $n = 1$. An illustration for $n = 1, 2$, and 3 is given in Figure 3.2.1.

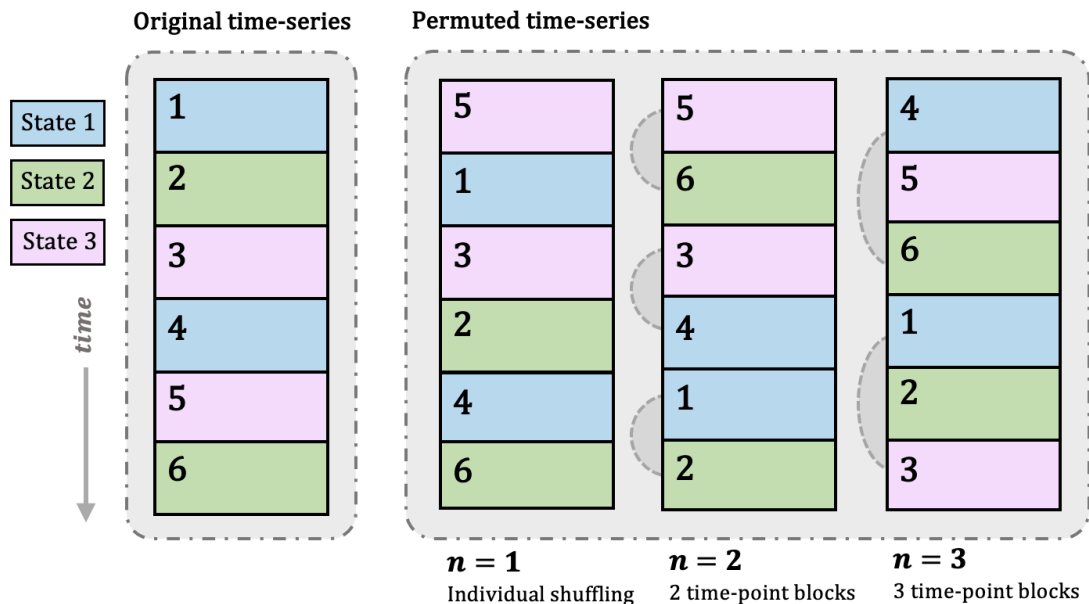


Figure 3.2.1. Individual and grouped permutations. A simple version of the original time-series is shown on the left, with 6 time points and 3 possible states indicated by different colors. On the right, we present permutations for $n = 1, 2$, and 3. For $n = 2$ and $n = 3$, grouped time points are highlighted with the grey “links” on the left of the permuted time-series. The arrangement of blocks was shuffled, but the order of time points within them remained unchanged (i.e., for $n = 2$: 12-34-56 \rightarrow 56-34-12).

As an intermediate step between the $n = 1$ and $n = 2$ permutations, we also generated an $n = 1.5$ group, which was achieved by grouping two consecutive time points and permuting the third one separately. We obtained 20 new time-series for each n and computed $\mu[\Phi^{\max}]$ for each one. Although these time-series had a different arrangement of states, this procedure did not change the overall content of states in the new signals (i.e., the number of times each state appeared was unchanged).

In terms of how these permutations vary in extent, larger n corresponds to a greater degree of signal preservation; by keeping more time points together, the total number of blocks to shuffle is smaller, and permutations result in less changed transitions. On the other hand, permuting with smaller n resulted in a greater degree of disruption to the original time-series, as there are more blocks to shuffle and hence more transitions to change. After obtaining the mean $\mu[\Phi^{\max}]$ of each group, these values were plotted against n on a logarithmic scale. As our results will show, the relationship resembled an exponential plateau function. We applied a fit of the form in equation 3.2.1 using a non-linear least squares method in the SciPy statistics module. The behavior of this function and its parameters is illustrated in Figure 3.2.2.

$$\mu[\Phi^{\max}] = \Gamma[1 - e^{-\lambda n}] \quad 3.2.1$$

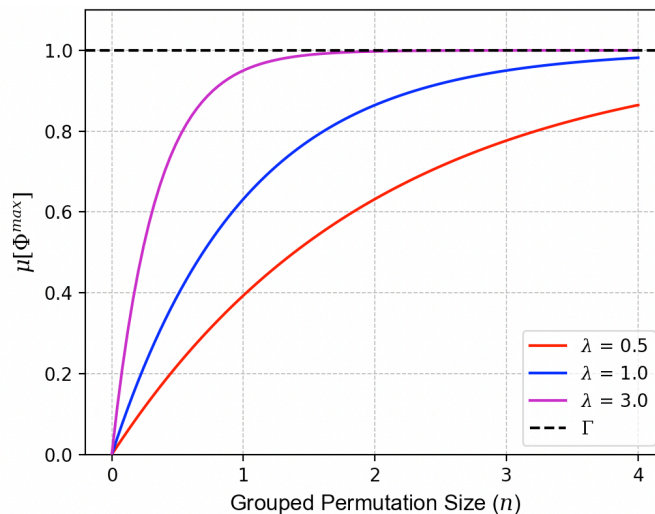


Figure 3.2.2. Exponential plateau function used to fit grouped permutation data. The means of $\mu[\Phi^{\max}]$ were plotted against n and an exponential plateau function of the type shown above was used to fit the data points.

The first parameter, Γ , corresponds to the maximum value reached by the curve and is effectively an approximation of the network's unpermuted $\mu[\Phi^{\max}]$, while the exponential parameter, λ , is proportional to the rate at which the plateau is approached. To allow for a fit of this form and to avoid a third parameter, we used the average $\mu[\Phi^{\max}]$ of the $n = 1$ (control) distribution as a baseline (i.e., this value was subtracted from all data points), and the first point in the logarithmically scaled graph was $\ln(n=1) = 0$, $\mu[\Phi^{\max}] = 0$.

3.2.2. Diagonal strength and diagonally selective permutations

To gain better insight into how permutations affected each RSN's time-series and integrated information, we sought to obtain a metric that represents a network's overall rate of fluctuation. Since permutations randomly disrupted each time-series' transitions, we expected to see an increased fluctuation rate in the permuted signals. To underscore this effect, Figure 3.2.3 compares the state-by-state TPMs of the original and permuted time-series for the frontoparietal network. When shown side-by-side, the most apparent difference is that the original TPM contains much larger probabilities along its diagonal entries, whereas the entries of the permuted TPM are smaller and more dispersed.

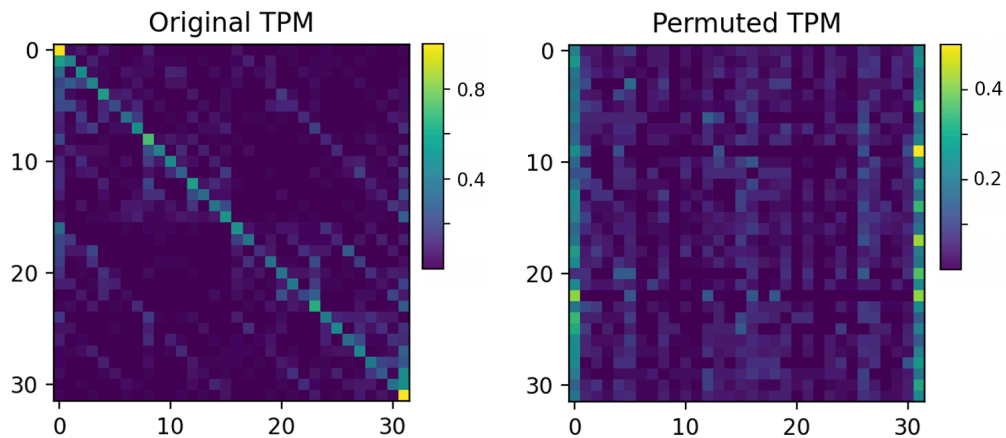


Figure 3.2.3. TPMs for the original and permuted time-series of the frontoparietal network. A noticeable difference between the original TPM and its permuted counterpart is that the former has many more diagonal entries with high probabilities.

The diagonal entries of a TPM represent the probabilities of states transitioning to themselves. Accordingly, a TPM with larger values along its diagonal represents a network that spends more time in a certain state before transitioning to another one, and

hence fluctuates more slowly. For each permuted time-series and their respective TPMs, we computed the proportion of probabilities in the diagonal entries to those in the off-diagonal entries, which we referred to as diagonal strength (DS):

$$DS = \frac{\sum_{i=j} TPM_{ij}}{\sum_{i \neq j} TPM_{ij}} \quad 3.2.2$$

In the analysis that followed, we aimed to determine whether any relationship exists between DS and $\mu[\Phi^{\max}]$. In the unpermuted signals, DS ranged between 0.2 and 0.3. Considering the random nature of the permutations performed, we expected DS to drop for the permuted time-series. Therefore, if we only used these permuted signals, our analysis would have been limited to DS values between 0 and 0.3. Theoretically, the upper limit of DS is 1, which corresponds to a matrix where all entries are along the diagonal. To increase the range of values included in our analysis, we performed another set of rearrangements to each RSN's time-series. Instead of permuting time points randomly, we grouped time points belonging to the same state so that they appear consecutively within the time-series. We outline this procedure and the type of time-series it produced in Figure 3.2.4.

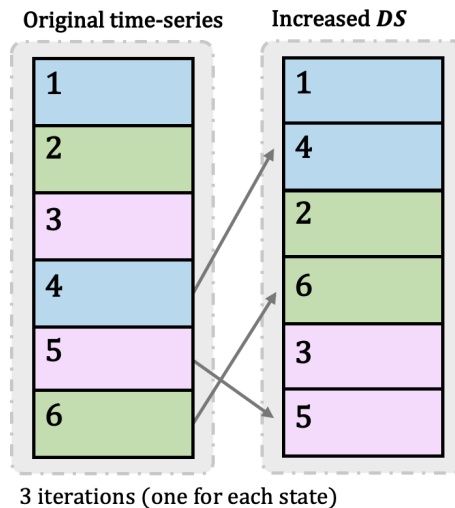


Figure 3.2.4. Diagonally selective rearrangements of time-series. To rearrange a time-series for increased diagonal strength, we changed the positions of time points sharing the same state (indicated by colors as in Figure 3.2.1) so that they appear consecutively in the permuted time-series. Each iteration consisted of doing this to a single state chosen at random (3 iterations are indicated by the arrows).

We wrote an iterative algorithm that scans through the signals, picks a state at random, and locates two non-consecutive instances of the chosen state within the time-series. The second instance of the state was then repositioned to follow the first; note that this repositioning could have been applied to one or several time points, depending on whether the chosen instance contained a consecutive repetition of the same state. The more iterations performed, the greater the increase in diagonal strength. The effects this had on the original TPM are presented in Figure 3.2.5. This algorithm was repeated to obtain a set of 140 new reordered time-series for each RSN in the awake condition.

We started with 100 iterations and increased this number by 35 each time, which allowed for an incremental increase of DS for each new signal. We found that the upper limit of DS could be reached with 5000 iterations. The upper limit was usually very close to 1, but always slightly less. Since the original time-series cycled through each state at least once and the frequency of each state's appearance in the time-series was unchanged, there had to be off-diagonal entries representing these inevitable transitions.

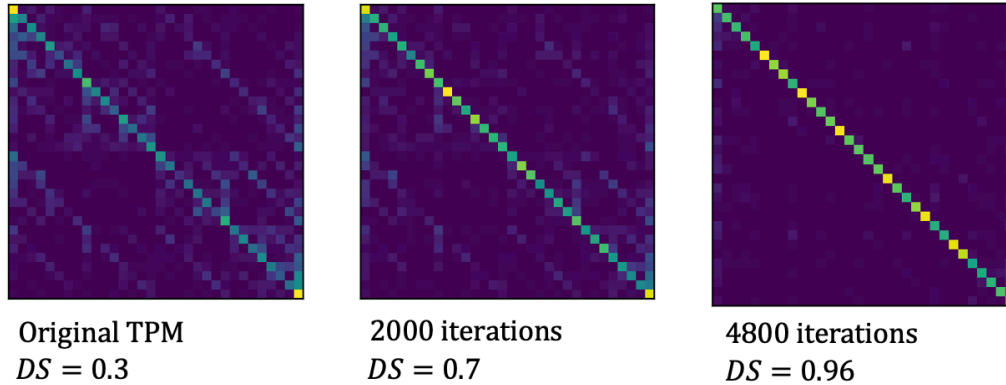


Figure 3.2.5. The effects of diagonally selective reordering of a TPM. The algorithm used to increase diagonal strength was repeated for increasing iterations until the upper limit of DS was reached.

3.3. Results

3.3.1. Grouped permutations

The results of the grouped permutation procedure are presented in Figures 3.3.1 and 3.3.2, where $\mu[\Phi^{\max}]$ values are plotted against n , the number of grouped time points.

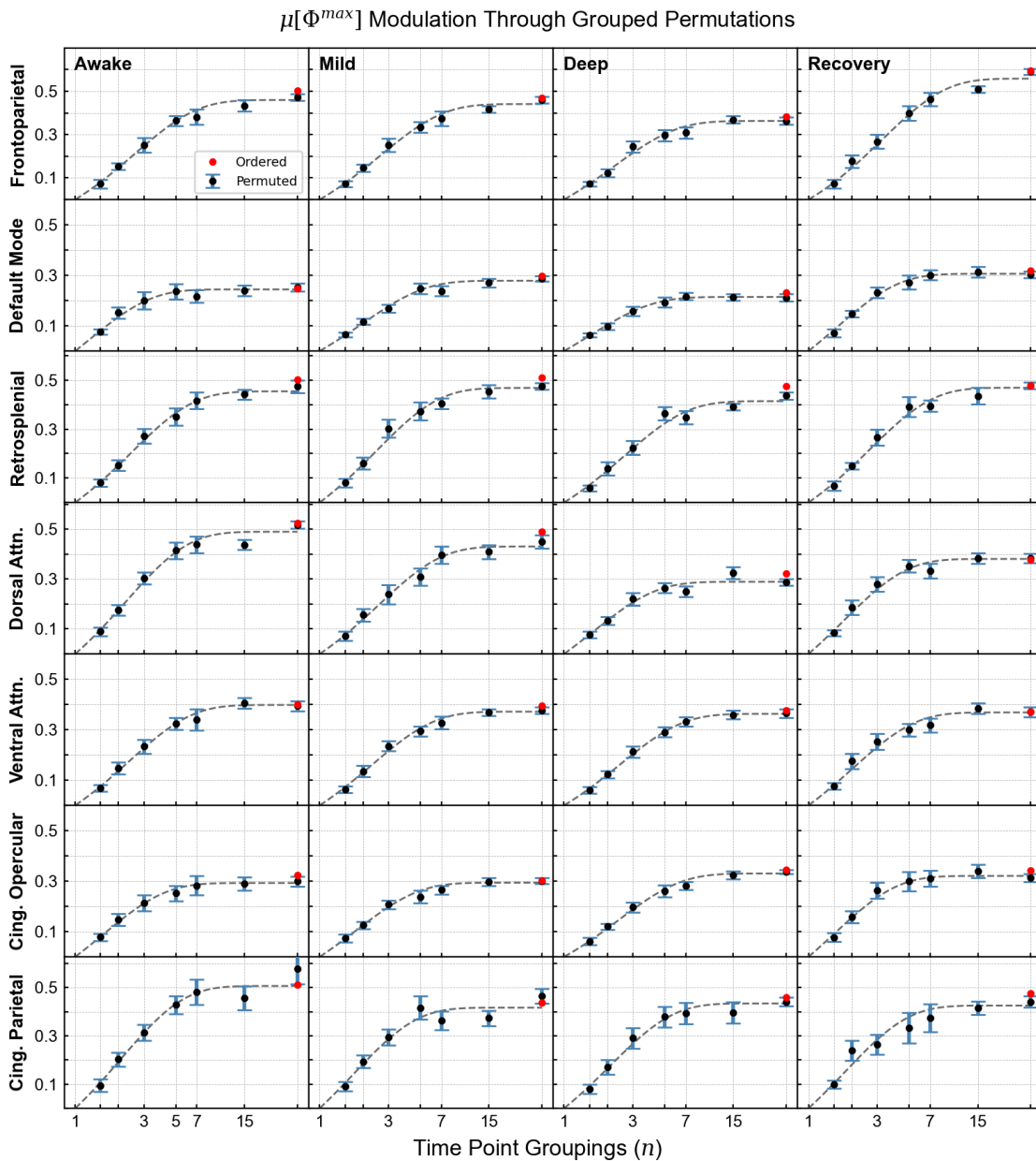


Figure 3.3.1. Grouped permutations in higher-order RSNs. For several group sizes containing n time points, permutations were performed by randomly rearranging the positions of grouped blocks within the time-series. The mean and

standard deviation of $\mu[\Phi^{\max}]$ are plotted against each n (logarithmically scaled) for RSNs that are associated with higher-order functions. Different networks are given along rows and the four conditions are presented across columns. While all RSNs plateau towards the $\mu[\Phi^{\max}]$ value of the original (unpermuted) time-series as n increases, the rate of approach differs across networks and states of sedation.

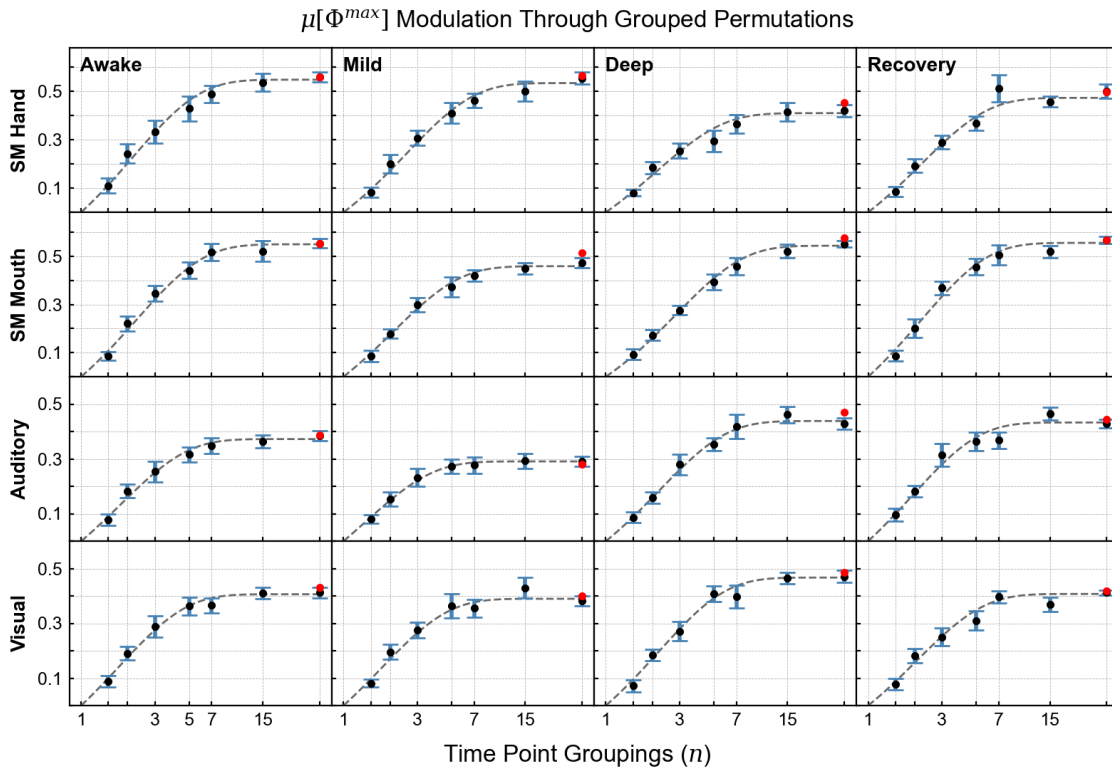


Figure 3.3.2. Grouped permutations in sensorimotor and sensory RSNs. The results for the same procedure are shown but for RSNs associated with sensorimotor and sensory functions.

In almost all cases, the grouped permutations decreased $\mu[\Phi^{\max}]$. Any increase of this measure was minor in magnitude and mostly observed for permutations with large n . In the two figures above, the same general behaviour is demonstrated across RSNs and conditions: as n increases, $\mu[\Phi^{\max}]$ converges towards the values of the original time-series, from which the $n = 35$ distributions do not deviate significantly. This result is not surprising considering our description of these permutations; with smaller n , there is a larger number of blocks to be shuffled, which induces a greater degree of sequential disruption and reduces $\mu[\Phi^{\max}]$ more drastically. On the other hand, permutations with larger n swap fewer transitions and hence preserve more of the original signal.

It appears that the span of each plateau, and hence the rate at which the unpermuted $\mu[\Phi^{\max}]$ values are approached, varies across networks and conditions. The means in these figures were fitted using the exponential plateau function given in equation 3.2.1. As demonstrated in Figure 3.2.2 of the Methods section, the exponential parameter λ is proportional to the rate at which the plateau is reached: a high λ corresponds to a faster rate of approach as the number of grouped time points increases, which can be observed with a longer plateau (i.e., $\mu[\Phi^{\max}]$ is near the maximum at $n = 5$). In contrast, networks with low λ demonstrate a steadier increase that continues with larger n values, and their plateau is visibly shorter. The λ values obtained for each RSN and condition are presented in the bar plots of Figure 3.3.3. For reference purposes, we also plotted each network's $\mu[\Phi^{\max}]$ values.

Upon closer inspection, it is noticeable that higher λ values pertain to networks lower in $\mu[\Phi^{\max}]$. For example, consider the default mode (DMN) and frontoparietal (FPN) networks in the awake condition (Figure 6, rows 1 & 2): the latter generates an unpermuted $\mu[\Phi^{\max}]$ value nearly twice as great as the former, while its λ values are substantially lower (DMN: $\mu[\Phi^{\max}] = 0.30$, $\lambda = 0.81$; FPN: $\mu[\Phi^{\max}] = 0.56$, $\lambda = 0.39$). The same observation holds true for other RSNs with higher $\mu[\Phi^{\max}]$, such as the retrosplenial, ventral-attention, and sensorimotor mouth networks. Recall that some higher-order RSNs, such as the ventral attention and retrosplenial networks, presented no significant changes in $\mu[\Phi^{\max}]$ throughout sedation. This is also the case for their λ values, which do not fluctuate significantly over the four conditions. Furthermore, there appears to be an exception to this trend in the DMN; although λ is significantly different for the four conditions, it presents no clear relationship with $\mu[\Phi^{\max}]$. The sensorimotor hand network is another exception, as there are no significant changes in λ despite its substantial variations in $\mu[\Phi^{\max}]$. Otherwise, the sensorimotor mouth and sensory networks also demonstrate an inverse relationship between the two parameters.

This relationship is also apparent for sedative-induced changes within a network. In the FPN and dorsal attention network (DAN), $\mu[\Phi^{\max}]$ decreased moving towards deep sedation and increased in recovery; in terms of λ , both networks demonstrate a significant

increase in deep sedation followed by a significant decrease in recovery for the FPN. Taking all networks and conditions together, the Pearson correlation coefficient between $\mu[\Phi^{\max}]$ and λ was -0.68 ($p < 0.001$, with 44 samples for each parameter). The relationship between these two parameters is shown more clearly in Figure 3.3.4, where we plot each of the 44 values obtained for λ against $\mu[\Phi^{\max}]$.

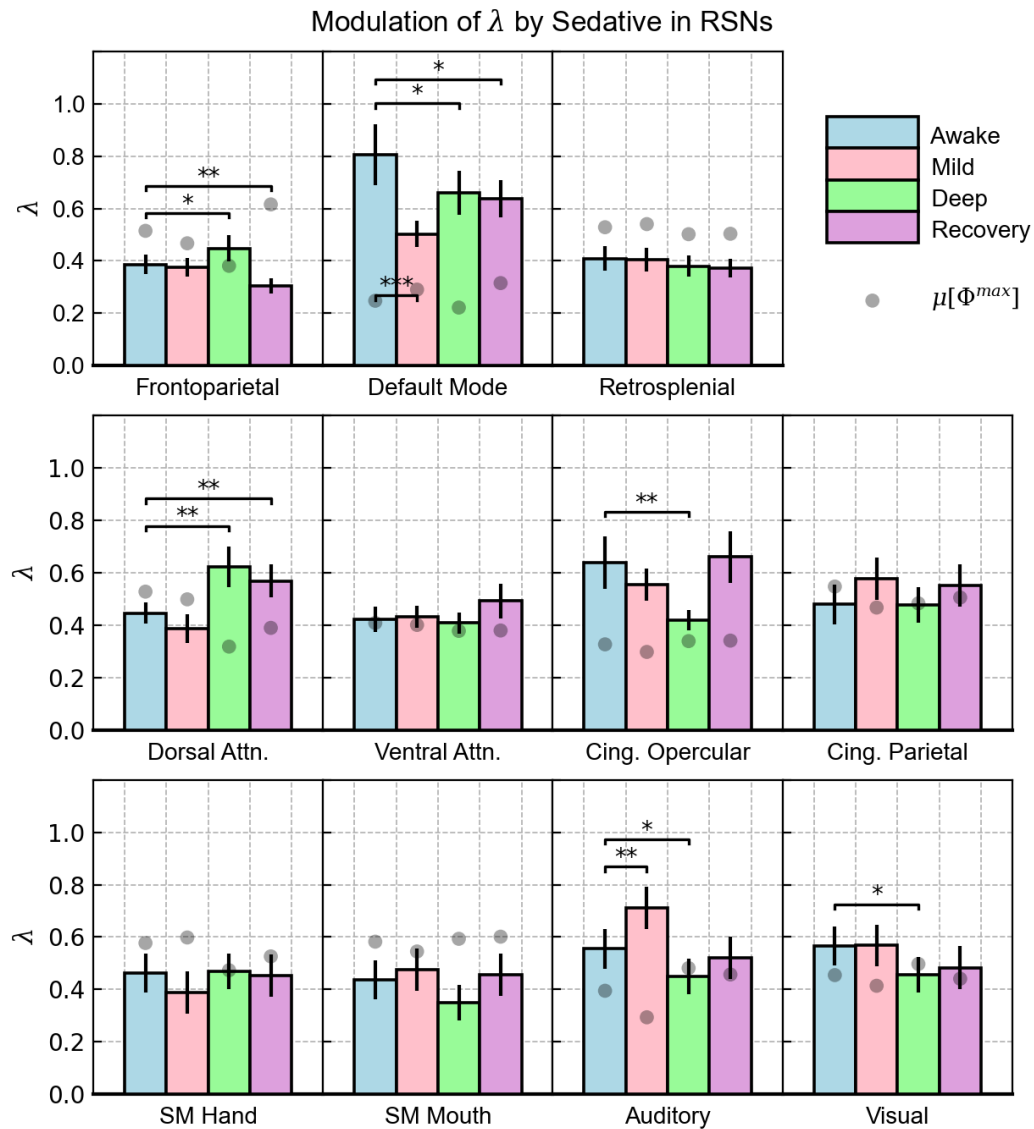


Figure 3.3.3. Modulation of λ and its inverse relationship with $\mu[\Phi^{\max}]$. The data in Figures 3.3.1 and 3.3.2 was fitted using an exponential plateau function, from which we obtained the exponential parameter λ to quantify the rate of approach with respect to increasing n . As a reference, we also plot the corresponding $\mu[\Phi^{\max}]$ values. An anticorrelation was found between λ and

$\mu[\Phi^{\max}]$, which can be observed across different networks and the four conditions. Statistically significant differences are shown only for comparisons that involve λ in the awake condition ($0.01 < p < 0.05$: *, $0.001 < p < 0.01$: **, $p < 0.001$: ***).

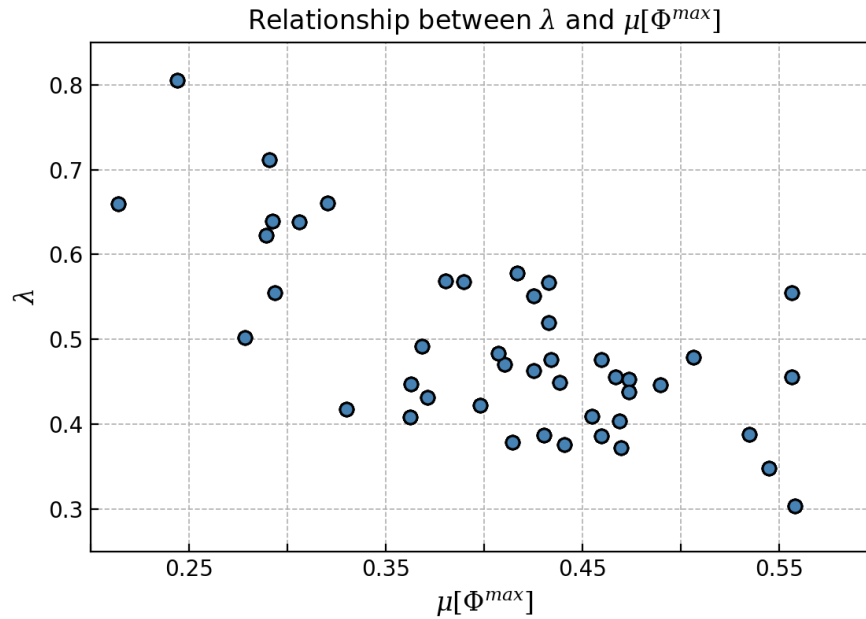


Figure 3.3.4. Anticorrelation between λ and $\mu[\Phi^{\max}]$. All values obtained for λ and $\mu[\Phi^{\max}]$ (44 time-series, obtained by concatenating all 17 subjects). There is an inverse correlation between the two parameters, with an increase in $\mu[\Phi^{\max}]$ corresponding to a decrease in λ .

3.3.2. Diagonal strength and diagonally selective permutations

Although we analyzed the relationship between $\mu[\Phi^{\max}]$ and n for grouped permutations, we also wanted to confirm that a decrease in diagonal strength would correspond to a decrease in $\mu[\Phi^{\max}]$. In Figure 3.3.5, we plot $\mu[\Phi^{\max}]$ against the range of DS values corresponding to the permuted signals of each RSN in the awake condition. While these plots verify the expected relationship in a general sense, it is important to comment on the vertical clusters of data points. These clusters, which correspond to the same n value, present a significant variance in $\mu[\Phi^{\max}]$. At the same time, the DS value does not vary as significantly for the points within them. Therefore, the resolution of this relationship is limited, and DS does not appear to provide any more information than the group size used for each permutation.

Nevertheless, we continued with the second part of this analysis, where $\mu[\Phi^{\max}]$ was analyzed for signals with DS ranging between 0.3 and 1.0. The results for each RSN in the awake condition are presented in Figure 3.3.6. The points on the left of the unpermuted signal are those presented in Figure 3.3.5, while those to the right of it were obtained from the diagonally selective permutations. Over the entire range of DS values, $\mu[\Phi^{\max}]$ demonstrates a bimodal behavior. The first peak corresponds to the value obtained for the unpermuted time-series. A further increase in DS leads to a second peak between 0.6 and 0.7. In all RSNs except for the cingulo-parietal network, the value of $\mu[\Phi^{\max}]$ at the second peak is greater than the first. From this point, $\mu[\Phi^{\max}]$ drops rapidly and approaches 0 as DS nears its maximum value. $\mu[\Phi^{\max}]$ also vanishes at the lower limit of DS ($DS = 0$). As seen with the clusters corresponding to grouped permutations, $\mu[\Phi^{\max}]$ continues to have large variance over a small range of DS values, so the resolution of this measure remains limited. Although there is variation in values of $\mu[\Phi^{\max}]$ and DS at the second peak, we could not find any meaningful relationships in how these properties vary between different networks and states of awareness.

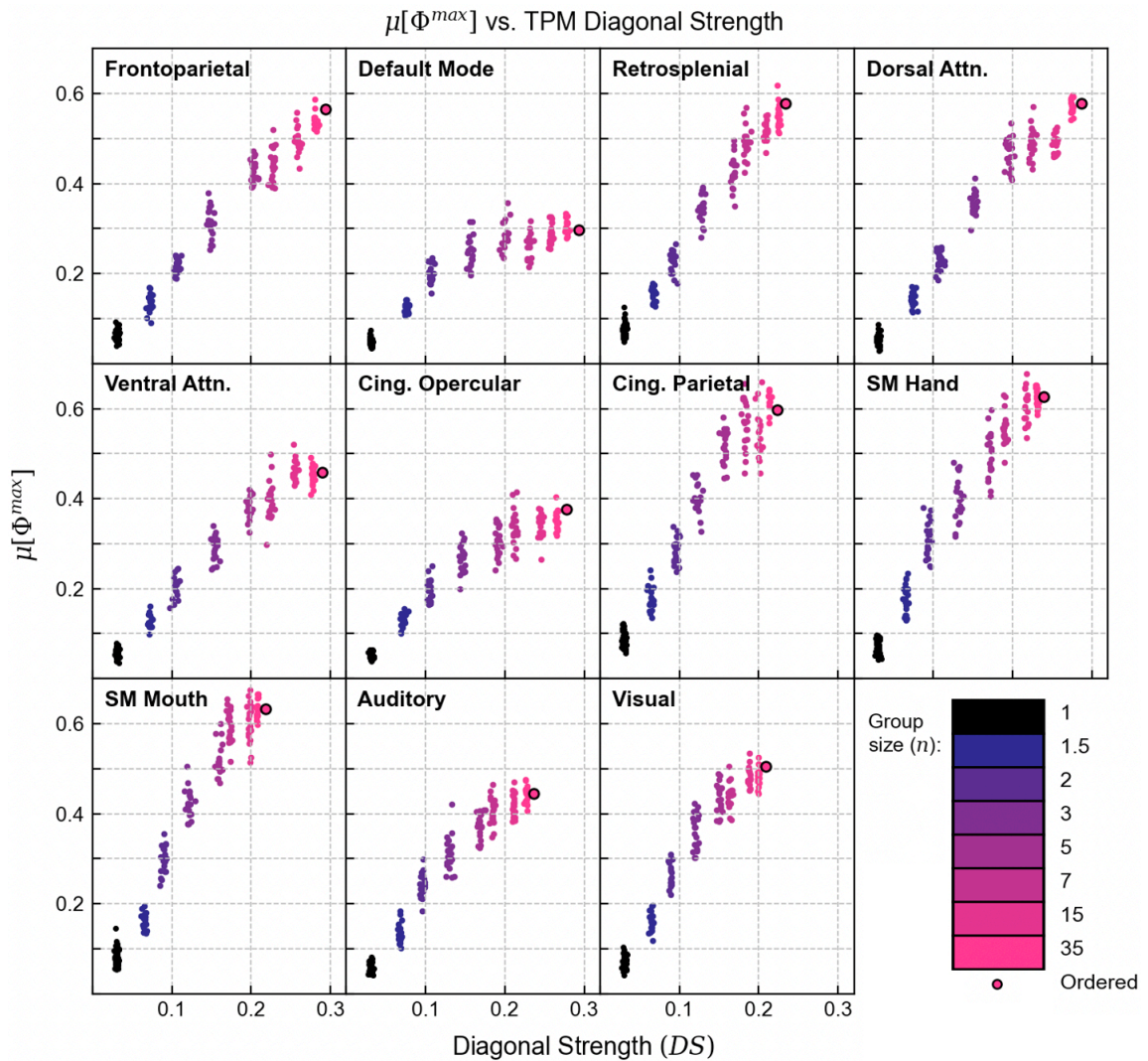


Figure 3.3.5. Relationship between $\mu[\Phi^{max}]$ and TPM diagonal strength for grouped permutations. The first part of our attempt to relate $\mu[\Phi^{max}]$ and DS involved the grouped permutations ($0 < DS < 0.3$). The $\mu[\Phi^{max}]$ values obtained by randomly permuting each RSN (awake condition) in groups are plotted against the diagonal strength of the TPM corresponding to the permuted time-series. The colors of points reflect the n value of the permutation performed, which are shown using the legend on the right.

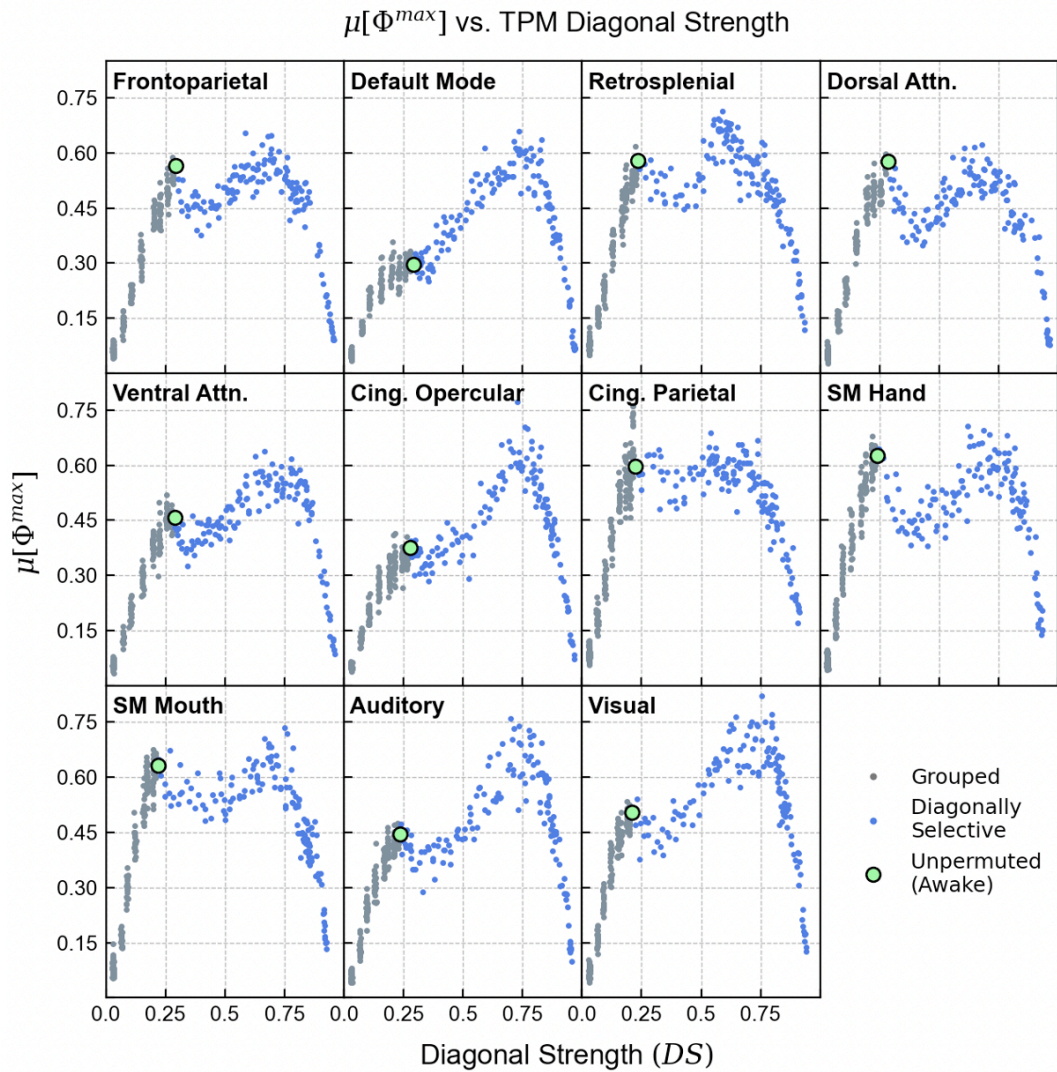


Figure 3.3.6. Relationship between $\mu[\Phi^{max}]$ and TPM diagonal strength for grouped permutations. This figure builds on the relationship shown in Figure 3.3.5 and extends to the upper limit of diagonal strength ($DS = 1$). Grey scatters indicate grouped permutations (same data as Figure 3.3.5), whereas blue scatters represent diagonally selective permutations. The overall behavior is bimodal, with two peaks of $\mu[\Phi^{max}]$ at different values of DS . At the lower and upper limits, $\mu[\Phi^{max}]$ approaches 0.

3.4. Discussion

In this chapter, we sought to relate $\mu[\Phi^{\max}]$ to several temporal characteristics of the fMRI data we obtained. First and foremost, the permutation procedure we applied (as well as the related control procedure in Chapter 2) showed that when random changes are introduced to each RSN's time-series, integrated information is almost always subject to a decrease. Since causality is the foundation for computing integrated information, this finding underscores the existence of causal structures that govern the time-dependent behavior of each RSN.

By applying a series of permutations that varied in their extent of disruption to the original time-series, we found that each RSN's integrated information can persist until a substantial portion of its temporal structure is broken. While all networks demonstrated a similar behavior as the degree of permutation was increased (which corresponded to smaller group sizes n), there was variation in how rapidly each RSN's $\mu[\Phi^{\max}]$ value changed. To describe this rate, we introduced a fit parameter λ . Based on our results, this can be seen as a measure of a signal's capacity to maintain its integrated information despite disruptions to its temporal structure; the higher λ is for a network, the more it must be randomly permuted for $\mu[\Phi^{\max}]$ to drop significantly.

The network with the highest value for this parameter was the DMN; when its time-series was permuted in groups containing as few as five time points, $\mu[\Phi^{\max}]$ did not drop significantly. Fundamentally, this means that the dynamics giving rise to the DMN's integrated information are relatively simple, as they can be captured over a short sequence of states. On the other hand, significantly lower λ values were observed in the FPN, where the slightest permutation significantly reduced $\mu[\Phi^{\max}]$. In contrast to the DMN, this means that the dynamics giving rise to its integrated information span a larger group of time points and are hence more sophisticated.

While the DMN had the highest λ value, we also saw that its time-series generated the lowest integrated information. By comparing λ and $\mu[\Phi^{\max}]$ in all RSNs and conditions, we found a moderate to strong negative correlation between these two

measures. In time-series with high λ and low $\mu[\Phi^{\max}]$, such as the DMN, grouped permutations did not drastically impact $\mu[\Phi^{\max}]$ because these signals contained weaker causal structures in the first place. In other words, one can describe the transitions in their time-series as being more “random” or “inconsequential. On the other hand, highly integrated networks with low λ were more susceptible to permutations; this reflects a more intricate temporal behavior that contains a higher frequency of “consequential” transitions between states, which ultimately accounts for their high $\mu[\Phi^{\max}]$.

These differences were also observed for different states of awareness within the same network. In the FPN and DAN, λ reached its maximum value during deep sedation. At the same time, $\mu[\Phi^{\max}]$ was minimized for this condition. By the same reasoning, this could indicate that the time-series of the FPN and DAN exhibited more sophisticated causal properties when subjects are awake than when they were deeply sedated. This is in line with our discussion about these networks in Chapter 2, where we referred to neuroscientific literature to explain why their integrated information most closely corresponded to the conscious state of subjects.

In contrast to the use of random permutations, the second procedure in this chapter relied on rearrangements that were specifically designed to increase the consecutive appearance of a states within the time-series. When random permutations were applied, we noticed that the probabilities in the diagonal entries of the TPMs were significantly reduced (see Figure 3.2.3); originally, each RSN’s time-series tended to remain in a certain state for several time points before transitioning to another one, which was characterized as having a high diagonal strength (DS) in its TPM. However, permutations introduced random transitions that resulted in more random fluctuations between different states, and consequently, the initial level of “inertia” was lost.

While the first analysis showed that random permutations reduce $\mu[\Phi^{\max}]$, we aimed to determine how integrated information would behave for time-series where DS was increased. After generating such signals, we found that $\mu[\Phi^{\max}]$ had a bimodal behavior with respect to this TPM property. At both lower and upper limits of DS (0 and 1, respectively), $\mu[\Phi^{\max}]$ approached zero. In keeping with what integrated information

theory aims to do, we can consider these two extremes to be two separate cases of impaired consciousness. The case where $DS \rightarrow 0$ may correspond to a time-series that is highly disordered and fluctuates too rapidly for integration to occur. This may describe the cortical activity of patients diagnosed with disorders of consciousness (DOC). For example, the connectivity of the DMN was found to be diminished in patients diagnosed with a vegetative state [4,5]. While some DMN activity in these patients persists, its impairment can result in temporal fluctuations that lack coherence. On the other hand, the case where $DS \rightarrow 1$ represents a network that is frozen in its state and hence lacks the fluctuations necessary to generate causal structures. This would correspond to a condition known as brain death; whereas DOC patients lose behavioral abilities but maintain a range of brain functions, brain death involves a complete loss of brain activity and, inevitably, consciousness [6].

While we were unable to interpret both peaks in the $\mu[\Phi^{\max}]$ vs DS curves, it is important to mention that a healthy and wakeful brain (as measured with the unpermuted signal) operates somewhere between these two limits. This is consistent with an important theory known as the “critical brain hypothesis”, which argues that the brain operates at a “critical point” that maintains balance between states of order and disorder [7–9]. In most of the networks shown in Figure 3.3.5, $\mu[\Phi^{\max}]$ was greater at the second peak. This peak may represent the time-series of a network in a “hyper-conscious state”, such as involvement in a task or hyper arousal [10]. Although we found some general relationships between integrated information and DS , these were limited in precision, as a certain value of DS corresponded to a large range of $\mu[\Phi^{\max}]$ values. Ultimately, this underscores that $\mu[\Phi^{\max}]$ is a much more intricate measure than a TPM property that is relatively easy to compute

3.5. References

- [1] Glover GH. Overview of Functional Magnetic Resonance Imaging. *Neurosurgery Clinics of North America* 2011;22:133–9. <https://doi.org/10.1016/j.nec.2010.11.001>.
- [2] Kyathanahally SP, Wang Y, Calhoun VD, Deshpande G. Investigation of True High Frequency Electrical Substrates of fMRI-Based Resting State Networks Using Parallel Independent Component Analysis of Simultaneous EEG/fMRI Data. *Frontiers in Neuroinformatics* 2017;11. <https://doi.org/10.3389/fninf.2017.00074>.
- [3] Lewis LD, Setsompop K, Rosen BR, Polimeni JR. Fast fMRI can detect oscillatory neural activity in humans. *Proceedings of the National Academy of Sciences* 2016;113. <https://doi.org/10.1073/pnas.1608117113>.
- [4] Vanhaudenhuyse A, Noirhomme Q, Tshibanda LJ-F, Bruno M-A, Boveroux P, Schnakers C, et al. Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. *Brain* 2010;133:161–71. <https://doi.org/10.1093/brain/awp313>.
- [5] Hannawi Y, Lindquist MA, Caffo BS, Sair HI, Stevens RD. Resting brain activity in disorders of consciousness: A systematic review and meta-analysis. *Neurology* 2015;84:1272–80. <https://doi.org/10.1212/WNL.0000000000001404>.
- [6] Wijdicks EFM. The Diagnosis of Brain Death. *New England Journal of Medicine* 2001;344:1215–21. <https://doi.org/10.1056/NEJM200104193441606>.
- [7] Beggs JM. The criticality hypothesis: how local cortical networks might optimize information processing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 2008;366:329–43. <https://doi.org/10.1098/rsta.2007.2092>.
- [8] Hesse J, Gross T. Self-organized criticality as a fundamental property of neural systems. *Frontiers in Systems Neuroscience* 2014;8. <https://doi.org/10.3389/fnsys.2014.00166>.
- [9] Crutchfield JP. Between order and chaos. *Nature Physics* 2012;8:17–24. <https://doi.org/10.1038/nphys2190>.
- [10] Weston CSE. Posttraumatic Stress Disorder: A Theoretical Model of the Hyperarousal Subtype. *Frontiers in Psychiatry* 2014;5. <https://doi.org/10.3389/fpsy.2014.00037>.

Chapter 4 : Integrated Information and the Length of a Time-Series

4.1. Overview

Throughout this thesis, we computed the integrated information of each RSN over the entire population of subjects. To do this, we concatenated the time-series of individual subjects and obtained elongated signals for each network and condition. The initial motivation for this was the inconsistency of outputs obtained at the single subject level, which resulted from short time-series and sparse TPMs. In the data we obtained, subjects were scanned over four different states of awareness, which was a limiting factor in the number of time points that could be acquired for each condition. Although concatenations did not preclude our analyses from yielding meaningful results, there were inevitable limitations to this approach (see section 2.2.2).

In future studies, we hope to see our procedure applied to the fMRI data of individual subjects. This will be an important advancement, especially if IIT is to be implemented as a diagnostic measure for conditions such as disorders of consciousness. To reliably measure $\mu[\Phi^{\max}]$, this would require a longer acquisition time that yields a substantial number of time points for a single subject. In this chapter, we analyze the variation of $\mu[\Phi^{\max}]$ with respect to the number of time points used to compute it. We begin this analysis by varying the number of concatenated subjects. In doing so, we hope to provide a general guideline of how long a scan would have to be to obtain consistent measurements of integrated information.

Even so, it is important to keep in mind that the time used for fMRI acquisition is a matter of concern to clinicians and subjects receiving the scan [1,2]. This is because the process can cause discomfort and distress, as subjects are required to lie still in the scanner's narrow space. As a potential solution to this problem, we briefly discuss a theoretical framework known as the Generalized Ising Model (GIM), which our group previously implemented to simulate resting-state brain activity using structural data [3]. Implementing this model may overcome the time constraint of fMRI measurements, as there is no limit on the number of time points that can be obtained from a simulation.

4.2. Methods

4.2.1. Concatenations and their effects on $\mu[\Phi^{\max}]$

To analyze the dependence of an RSN's $\mu[\Phi^{\max}]$ on the number of time points in its signal, we obtained a range of concatenated time-series that differed in length. Starting with the time-series of individual subjects, we incrementally increased the number of time points by adding the time-series of one subject at a time until all 17 subjects were included. For each number of subjects (N), a sample of 16 time-series was obtained by randomly choosing the subjects to include in the concatenation. For a single subject in the awake condition, each time-series contained 245 time points. Therefore, the lengths of these time-series were 245 ($N = 1$), 490 ($N = 2$), 735 ($N = 3$), ..., and 4165 ($N = 17$).

After computing $\mu[\Phi^{\max}]$ for each sample, we applied Welch's t-test (see section 2.2.6) to analyze statistically significant differences that result from different numbers of subjects. The standard deviation of the sample was multiplied by the square root of the sample size to obtain the population's standard deviation.

4.2.2. The Generalized Ising Model

Although the first analysis may indicate the number of time points needed to reliably compute $\mu[\Phi^{\max}]$, it still contained data from different subjects. As previously described, inter-subject variations resulted in time-series with discontinuities. The ideal case would be to obtain a time-series of comparable length but from a single subject. Although we did not have longer empirical time-series for individual subjects, we approached this problem by simulating an RSN's activity with the Generalized Ising Model (GIM) [4–6].

The GIM, which originates from thermodynamics, simulates an RSN's time dependent behavior using structural data obtained with diffusion tensor imaging (DTI). DTI provides a structural connectivity matrix J that specifies the strength of white matter connections between the different regions of an RSN. An entry of this matrix represents the coupling strength between any two regions (i and j). We provide a brief demonstration of how this simulation works in Figure 4.2.1.

Each region in an RSN is assigned a “spin” S , which can be +1 or -1 in accordance with the two binary states we have been using (+1 for high activity, 1- for low activity). An important parameter in this simulation is the temperature T , which describes the network’s external conditions. The energy E of the system at a certain state s is obtained by summing over the product of each pair of spins and their coupling strengths. The probability of the system being in a certain state, P_s , is defined by using the Boltzmann probability distribution.

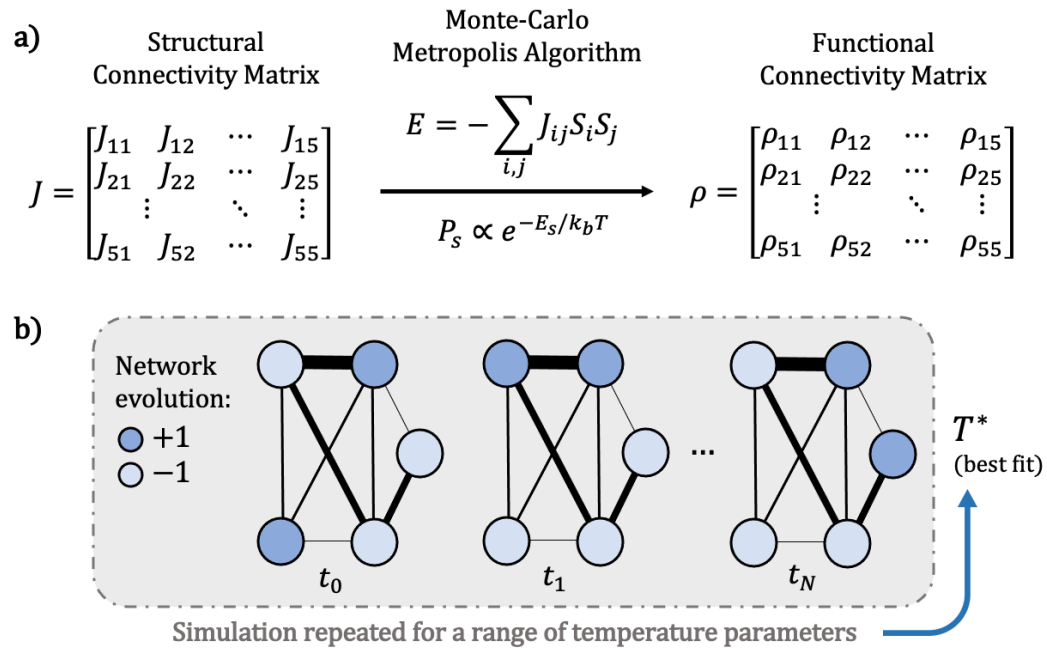


Figure 4.2.1. Simulating functional connectivity using the GIM. a) The process begins with a structural connectivity matrix J , which is fed into a Monte-Carlo Metropolis algorithm based on the Boltzmann probability distribution. The time-series obtained is used to construct a functional connectivity matrix ρ ; b) the networks consist of five regions (circles) that are connected through different coupling strengths (represented by lines of varying widths). At each time step, every spin is given the chance to flip. This process is repeated for a range of temperatures to find the one that best fits empirical data (T^*).

To simulate the network’s evolution over time, these principles are implemented in a Monte-Carlo Metropolis simulation. Starting with a random state, each region is given the chance to flip. If the flip is energetically favorable (i.e., $\Delta E < 0$), it is allowed to occur. If it is unfavorable, a random number is generated; if this random number falls

between 0 and the probability of the new configuration, the flip is allowed to occur. Once this has been repeated for every spin, the state of the system at the next time point is obtained. This is repeated for a specified number of time points, which constitute the simulated time-series of the network. The time-series is then used to construct a functional connectivity matrix that describes the network's temporal correlations.

Finally, this simulation is repeated for a range of temperature parameters. A correlation matrix is obtained for each iteration, and the temperature that produces the best fit between the simulated and empirical functional connectivity matrices is denoted T^* (i.e., this is the temperature where the simulation most closely resembles the correlation matrix that would be observed with real fMRI data).

4.2.3. $\mu[\Phi^{\max}]$ and the length of a simulated time-series

To test the relationship between time-series length and $\mu[\Phi^{\max}]$ in a single subject, we ran the GIM for the retrosplenial network at its T^* in the awake condition. DTI data corresponding to a single subject was obtained from the human connectome project database, which was processed to match the representative regions used in this study. We chose this RSN based on simulations conducted for previous studies, where we found that its simulated activity adequately matched empirical fMRI data at $T^* = 1.12$ (see Figure 4.2.2.) We ran the GIM and obtained simulated time-series for the same set of lengths used in the first procedure (245, 490, ..., 4165, or the equivalent of 1, 2, ..., 17 subjects). 16 time-series were generated for each length and $\mu[\Phi^{\max}]$ was computed for each one.

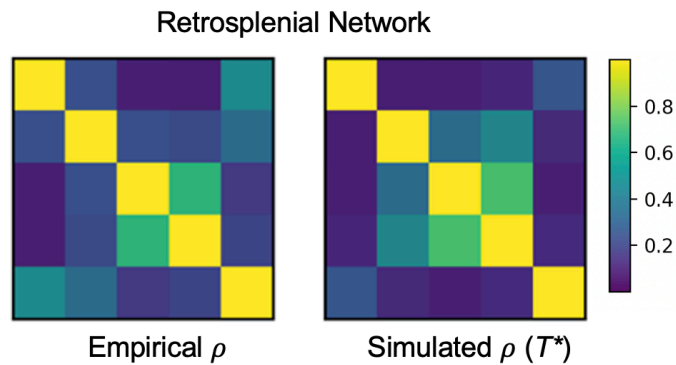


Figure 4.2.2. Empirical and simulated functional connectivity matrices. The GIM was found to accurately model ρ for the retrosplenial network at T^*

4.3. Results

4.3.1. Concatenations and their effects on $\mu[\Phi^{\max}]$

The effects of concatenating time-series from different subjects are presented in Figure 4.3.1, where we focus on the DMN and FPN in the awake condition. First, the mean and standard deviation of $\mu[\Phi^{\max}]$ were highest for the time-series of individual subjects. As the number of concatenated subjects increased, $\mu[\Phi^{\max}]$ dropped steadily and leveled off with time-series consisting of 8-10 subjects. Using Welch's t-test, we measured a significant difference from the single subject sample ($N = 1$) for all $N \geq 8$ time-series.

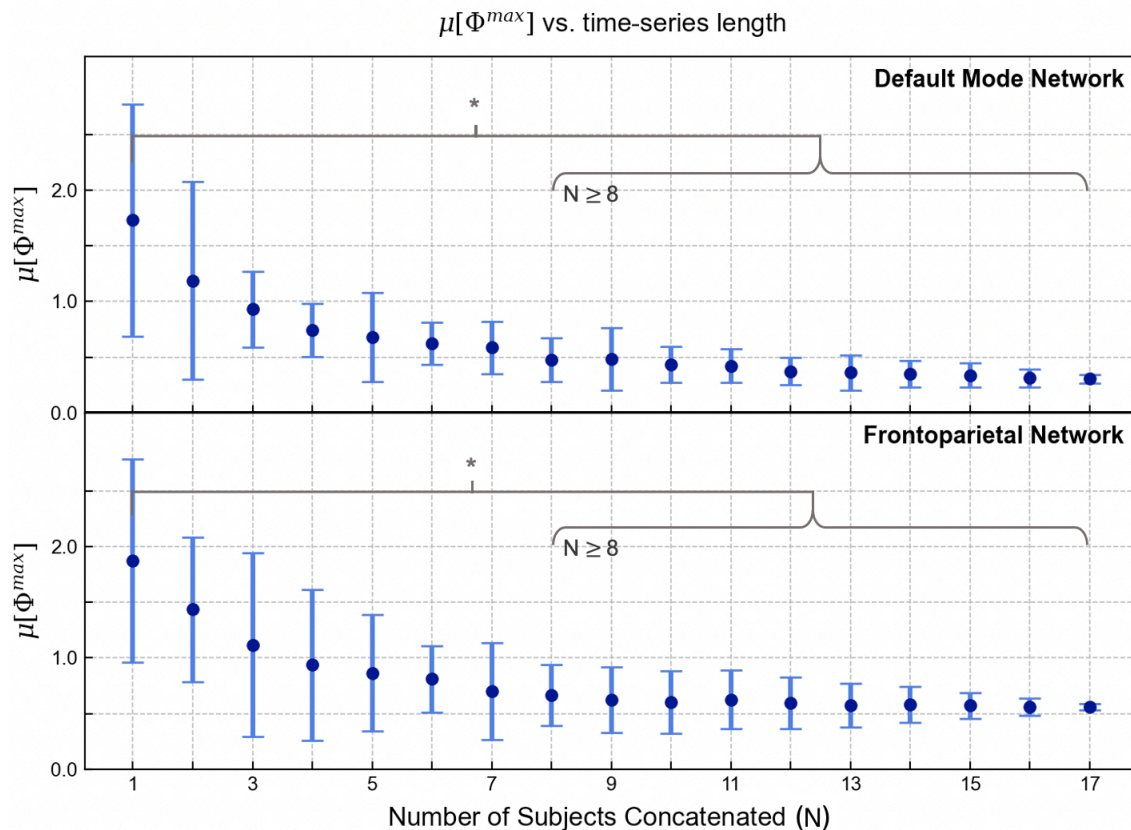


Figure 4.3.1. Variation of $\mu[\Phi^{\max}]$ with respect to the number of subjects concatenated. The means and standard deviations are presented for each sample of time-series obtained by concatenating a different number of subjects (N). $\mu[\Phi^{\max}]$ decreases for time-series with increased length. A significant difference from $N = 1$ was measured for all $N \geq 8$, and the value of $\mu[\Phi^{\max}]$ appears to stabilize past this point ($0.01 < p < 0.05$: *).

4.3.2. Simulated and empirical time-series

For the comparison between empirical and simulated time-series in the retrosplenial network, $\mu[\Phi^{\max}]$ is plotted against the number of time points as opposed to the number of subjects. The empirical results still correspond to different numbers of concatenations as in Figure 4.3.1, but the simulated data was obtained for a single subject. The overall behavior of the model is similar to the empirical time-series, with a drop of $\mu[\Phi^{\max}]$ that steadies as the number of time points reaches the equivalent of 8-10 subjects. However, the groups of simulated time-series do not present any significant differences.

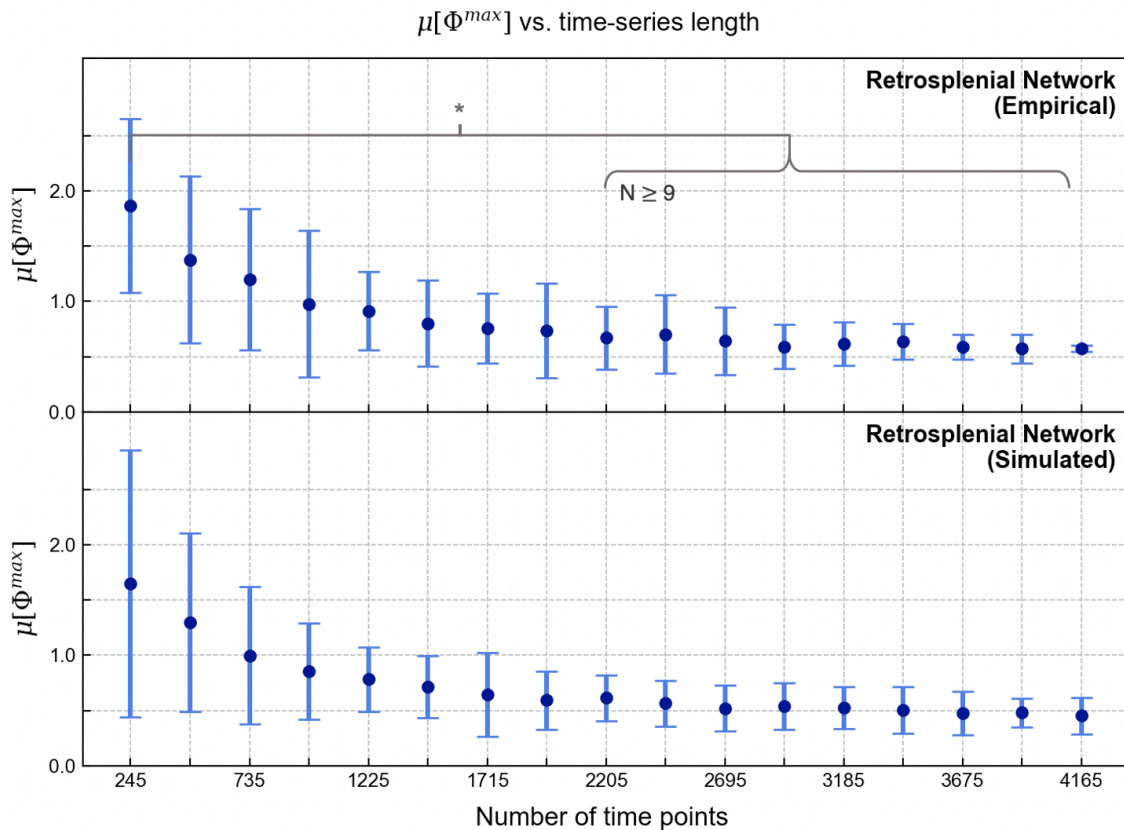


Figure 4.3.2. Variation of $\mu[\Phi^{\max}]$ in empirical and simulated time-series.

$\mu[\Phi^{\max}]$ is plotted against the number of time points, which corresponds to the number of subjects concatenated using empirical data (top plot) and the number of time points simulated for a single subject in the model (bottom plot). For the empirical data, a significant difference from $N = 1$ was measured for all $N \geq 9$. While the simulated behavior of the retrosplenial network is similar to that of the empirical time-series, none of its decreases were statistically significant.

4.4. Discussion

When it comes to obtaining time-series, the principal limitation of fMRI is its low temporal resolution [1]. Recall that consecutive images are separated by the repetition time (TR), which ranges between 2-3 seconds. This means that acquisition of 245 time points (one subject in one condition) required a scan ranging between 8-12 minutes per condition. We saw that for this amount of time points, the procedure yielded spurious results for $\mu[\Phi^{\max}]$, which prompted us to concatenate the population's time-series for each RSN and condition. By analyzing how the length of a time-series affects $\mu[\Phi^{\max}]$, our goal in this chapter was to provide general guidelines for future implementations of IIT in fMRI.

First, the most evident effect of concatenations (and increasing time points in the simulation) was that they reduced $\mu[\Phi^{\max}]$. Considering that each subject's time-series were different, this observation is in line with the fundamentals of IIT. In Chapter 1, we explained that mechanisms generate integrated information through a selective set of causes and effects. When different time-series are combined, the constraints pertaining to each subject's mechanisms are relaxed, meaning their cause-effect repertoires become less selective. For instance, if in the time-series of one subject, state i only transitions to state j , but in another subject, state i transitions to states j and k , the combined signal will have a less selective effect repertoire for the mechanisms related to state i . This effect became more pronounced as the length of the time-series increased, which is what we observed in Figures 4.3.1 and 4.3.2. In the simulated data, the same behaviour was observed because larger number of time points allowed for a broader range of spin flips and state transitions. As a primary guideline, we emphasize that the number of time points used in this procedure must be a fixed control; the comparison of $\mu[\Phi^{\max}]$ must be applied to time-series of the same length.

While increased length resulted in a general decrease of integrated information, the standard deviation corresponding to most samples of the empirical data was quite large. While this may suggest inconsistency regardless of what number of subjects is concatenated, it is important to note that these samples were generated by concatenating

random groups of subjects. Accordingly, there was significant variation in the time-series included for each measurement. In Chapter 2, we obtained a sample of 17 $\mu[\Phi^{\max}]$ values by concatenating 16 subjects and excluding a different subject for each measurement, which resulted in much less variation in the time-series. As the results showed, the standard deviation of $\mu[\Phi^{\max}]$ was much more reasonable, which allowed for observations of statistically significant differences between different states of awareness in certain RSNs. Thus, we suggest this practice for any future studies that involve the concatenation of time-series over a population (e.g., in a study with 12 subjects, this would involve obtaining a sample of 12 $\mu[\Phi^{\max}]$ values by concatenating the time-series of 11 subjects for each measurement).

Although the random grouping of subjects accounts for the large variance of $\mu[\Phi^{\max}]$ in the concatenated time-series, we also observed significant variance in the simulated results. Theoretically, the simulation was intended to represent the time-series of a single subject. While the GIM has been successfully applied to predict a network's correlational functional connectivity, it may not be as consistent when it comes to simulating integrated information. The large variation of $\mu[\Phi^{\max}]$ may be due to the stochastic nature of the model [4,5], which allows for a variety of time-dependent behaviors that vary in their integrated information.

Ultimately, using fMRI to obtain longer time-series from subjects will be the best way to understand how many time points are necessary to accurately compute $\mu[\Phi^{\max}]$ for an individual. While we can conclude with certainty that 245 TRs or less are not enough, it is difficult to provide exact guidelines with the data we used. Based on the statistically significant differences we found, this may require the equivalent of 8 concatenated subjects, or around 2000 TRs. However, this would make for a rather long scanning session (over 65 minutes), which is impractical for most clinical purposes [7]. While the results for time-series with less than 8 concatenated subjects did not indicate any statistically significant differences, it is possible that the equivalent of 5-7 concatenated subjects, or roughly 1200-1700 TRs, would be substantial to obtain consistent results. This would amount to a scanning time of at least 40 minutes, which is realistic for a resting-state acquisition procedure [8].

4.5. References

- [1] Glover GH. Overview of Functional Magnetic Resonance Imaging. *Neurosurgery Clinics of North America* 2011;22:133–9. <https://doi.org/10.1016/j.nec.2010.11.001>.
- [2] O'Connor EE, Zeffiro TA. Why is Clinical fMRI in a Resting State? *Frontiers in Neurology* 2019;10. <https://doi.org/10.3389/fneur.2019.00420>.
- [3] Abeyasinghe PM, Aiello M, Nichols ES, Cavaliere C, Fiorenza S, Masotta O, et al. Consciousness and the Dimensionality of DOC Patients via the Generalized Ising Model. *Journal of Clinical Medicine* 2020;9:1342. <https://doi.org/10.3390/jcm9051342>.
- [4] Fraiman D, Balenzuela P, Foss J, Chialvo DR. Ising-like dynamics in large-scale functional brain networks. *Physical Review E* 2009;79:061922. <https://doi.org/10.1103/PhysRevE.79.061922>.
- [5] Marinazzo D, Pellicoro M, Wu G, Angelini L, Cortés JM, Stramaglia S. Information Transfer and Criticality in the Ising Model on the Human Connectome. *PLoS ONE* 2014;9:e93616. <https://doi.org/10.1371/journal.pone.0093616>.
- [6] Zanoci C, Dehghani N, Tegmark M. Ensemble inhibition and excitation in the human cortex: An Ising-model analysis with uncertainties. *Physical Review E* 2019;99:032408. <https://doi.org/10.1103/PhysRevE.99.032408>.
- [7] Birn RM, Molloy EK, Patriat R, Parker T, Meier TB, Kirk GR, et al. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. *NeuroImage* 2013;83:550–8. <https://doi.org/10.1016/j.neuroimage.2013.05.099>.
- [8] Murphy K, Bodurka J, Bandettini PA. How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *NeuroImage* 2007;34:565–74. <https://doi.org/10.1016/j.neuroimage.2006.09.032>.

Chapter 5 : Summary, Conclusions, and Future Work

5.1. Summary

This thesis began with a comprehensive review on consciousness and how integrated information theory (IIT) attempts to explain it [1]. Up to this point, most developments made with this framework have been purely theoretical, as it is difficult to implement to empirical data. To move past this precedent, we presented a methodology on how Φ , the principal metric of IIT, can be obtained from fMRI measurements. Our approach relied on converting a collection of cortical resting-state networks (RSNs) to discrete systems consisting of binary elements, to which we then applied the computational scheme of IIT.

Altogether, we obtained promising results when this approach was applied to fMRI data of subjects who underwent propofol-induced sedation. In Chapter 2, we found that changes to integrated information from anesthesia closely reflected the conscious evolution of subjects in the frontoparietal and dorsal attention networks, which are frequently associated with executive control, voluntary attention, and higher-order processing of stimuli [2]. In Chapter 3, we extended this analysis and showed that networks with high integrated information demonstrate more intricate causal structures, as their time-dependent behavior was highly susceptible to permutations. This analysis underscored the intrinsic causal structures of each RSN, which gave rise to the integrated information we computed.

An important step in our methodology was the concatenation of time-series from the entire subject population, which were the signals used to analyze the integrated information of each RSN. However, this was one limitation of our study. As we discussed, the need for this stemmed from an inherent disadvantage of fMRI, which is its low temporal resolution. While this approach did not preclude a meaningful analysis of our data, Chapter 4 discussed how the length of a time-series affected our calculations of integrated information. We took this opportunity to present several guidelines for future implementations of our procedure, including an approximation of how long a scan should be to reliably compute integrated information for individual subjects.

Besides the use of concatenations and the low temporal scale of fMRI, there are other limitations worth mentioning. First, the use of time-series for calculations of Φ is constrained by the Markovian assumptions underlying the mathematical formulation of PyPhi [3]. In our case, the fMRI signals did not completely satisfy the conditional independence property, meaning that some of the original signal properties were inevitably lost when inputted to PyPhi. Another limitation of this software is the high computation cost of obtaining Φ and how it scales with increasing network size. Although an improved spatial resolution may provide better insight into each RSN's mechanisms, a scalability of $O(n53^n)$ means that a sixth region would increase computation time by over sixty-fold.

A critical step in the procedure was simplifying the time-series to create systems whose elements take on binary states. Although this is required by PyPhi for computational purposes, describing activity in this form allows for no consideration of the BOLD signal's fluctuation amplitudes, which is also an important factor in each network's dynamics. Moreover, we assigned one of the two binary states to a signal regardless of how much above or below it was relative to the baseline. To maintain a discrete system of elements while improving the signal's discrete resolution, it is perhaps more appropriate to analyze the system using three states, which could be +1 for activity significantly greater than the baseline, 0 for activity near the baseline, and -1 for activity significantly below the baseline. Although PyPhi is currently restricted to work with binary state systems, developments are currently being made to extend the number of possible states for each element within a system [4]. It is also possible to maintain a system of two states, but only consider time points where each region's activity is significantly above or below baseline. Even though binarizing the time-series significantly simplified the behaviour of each network, it is important to emphasize that the binarized signals still reflect variations that occur in the original data. While we were unable to compute integrated information from the BOLD signal in its natural form, our procedure still presents a promising starting point to achieve this goal.

5.2. Conclusions and Future Work

As discussed, we believe that an important step forward will be to apply our method to time-series of substantial length from individual subjects, which will allow for more meaningful comparisons of how integrated information varies from one individual to another. To further investigate Φ as a neural correlate of consciousness, this analysis should be applied to fMRI data obtained for other cases of impaired consciousness, such as the neurological disorders we discussed in Chapter 1. Although we focused on fMRI, our procedure could, in principle, be applied to other temporal measurements of brain activity where time-series are obtained for a set of brain regions, such as electrical activity measured with electroencephalography (EEG) [5]. In particular, the procedure we used for analyzing the Markov property and conditional independence could also be used to test the applicability of IIT to these other techniques.

We emphasize that although our study focused on integrated information, this framework continues to be debated, with some challenging the validity of the axioms used for its foundation. More generally, the importance of causality to consciousness is a subject of great debate, and some in the neuroscientific community believe that causal structure theories cannot account for conscious phenomena [6]. In the context of ongoing arguments about IIT, we believe that empirical applications of this framework are the best way to test the validity of its metrics as indicators of consciousness, which was the goal of our implementation. Moreover, IIT is not the only framework proposed to explain consciousness. Other notable perspectives include global workspace theory [7], Granger causality [8], and even quantum mechanical postulates about the brain [9]. Instead of ascribing to a single theory, the neuroscientific community should investigate this topic through multiple viewpoints that complement, rather than contradict, one another. Ultimately, the multifaceted nature of consciousness calls for a multifaceted approach to advance our understanding of this astonishingly intricate concept.

5.3. References

- [1] Oizumi M, Albantakis L, Tononi G. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology* 2014;10:e1003588. <https://doi.org/10.1371/journal.pcbi.1003588>.
- [2] Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL. Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. *Journal of Neurophysiology* 2008;100:3328–42. <https://doi.org/10.1152/jn.90355.2008>.
- [3] Mayner WGP, Marshall W, Albantakis L, Findlay G, Marchman R, Tononi G. PyPhi: A toolbox for integrated information theory. *PLOS Computational Biology* 2018;14:e1006343. <https://doi.org/10.1371/journal.pcbi.1006343>.
- [4] Gomez JD, Mayner WGP, Beheler-Amass M, Tononi G, Albantakis L. Computing Integrated Information in Discrete Dynamical Systems with Multi-Valued Elements. *Entropy* 2021; 23:e23010006. <https://doi.org/10.3390/e23010006>.
- [5] Kim H, Hudetz AG, Lee J, Mashour GA, Lee U. Estimating the Integrated Information Measure Phi from High-Density Electroencephalography during States of Consciousness in Humans. *Frontiers in Human Neuroscience* 2018;12. <https://doi.org/10.3389/fnhum.2018.00042>.
- [6] Doerig A, Schurger A, Hess K, Herzog MH. The unfolding argument: Why IIT and other causal structure theories cannot explain consciousness. *Consciousness and Cognition* 2019;72:49–59. <https://doi.org/10.1016/j.concog.2019.04.002>.
- [7] Baars BJ. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience, 2005, p. 45–53. [https://doi.org/10.1016/S0079-6123\(05\)50004-9](https://doi.org/10.1016/S0079-6123(05)50004-9).
- [8] Seth AK, Barrett AB, Barnett L. Granger Causality Analysis in Neuroscience and Neuroimaging. *Journal of Neuroscience* 2015;35:3293–7. <https://doi.org/10.1523/JNEUROSCI.4399-14.2015>.
- [9] Hameroff S, Penrose R. Consciousness in the universe. *Physics of Life Reviews* 2014;11:39–78. <https://doi.org/10.1016/j.plrev.2013.08.002>.

Chapter 6 : Supplementary Material

6.1. Propofol Administration and fMRI Acquisition

6.1.1. Participants and Ethics

We recruited 17 healthy volunteers (4 women; mean age 24 years, SD = 5) after posting printed advertisements throughout the university and sharing the study through word of mouth. All were native English speakers, right-handed, and had no history of neurological disorders. The attending MR technician and anesthesiologist instructed volunteers to complete safety screening questionnaires for MRI and propofol, followed by written informed consent forms to confirm their understanding of any potential risks involved. We remunerated volunteers for their time and willingness to participate. Ethical approval was obtained from the Health Sciences Research Ethics Board and Psychology Research Ethics Board of Western University (REB #104755) [1,2].

6.1.2. Administration of Propofol

In preparation for sedation, a 20 G i.v. cannula was inserted into a vein in the dorsum of the non-dominant hand and a propofol infusion system was connected to it. Intravenous propofol was administered with a Baxter AS 50 syringe pump (Singapore). To deliver propofol in an incremental, stepwise fashion, an effect-site/plasma steering algorithm was used in combination with a computer-controlled infusion pump. The infusion pump was adjusted to achieve the desired level of sedation, which was guided by target propofol concentrations predicted by the TIVATrainer (the European Society for Intravenous Anaesthesia, eurosiva.eu) pharmacokinetic simulation program. This model provided target-controlled infusion by adjusting propofol infusion rates, with the goal of reaching and maintaining the target blood concentrations specified by the Marsh 3 compartment algorithm for each participant (also incorporated in the TIVATrainer software) [3].

Subjects underwent four conditions throughout sedation and acquisition:

- 1) **Awake:** Propofol was not yet administered. Participants were fully awake, alert, and communicative.
- 2) **Mild sedation:** At the start of this phase, we began propofol infusion with a target effect-site concentration of 0.6 $\mu\text{g/ml}$. Oxygen was titrated to maintain SpO_2 above

96%. After reaching the target effect-site concentration, we assessed the participants' level of sedation and maintained the effect-size concentration if observations were consistent with mild sedation. Initially, participants became calmer and slow in their responses to verbal communication. Once they stopped engaging in spontaneous conversation, became sluggish in speech, and only responded to loud commands, they were classified as level 3 using the Ramsay sedation scale [4] and considered mildly sedated.

- 3) Deep sedation:** Prior to reaching the deep sedation phase, the target effect-site concentration was increased in increments of 0.3 $\mu\text{g}/\text{ml}$ and responsiveness was assessed with each increase. Once participants reached level 5 on the Ramsay scale of sedation, whereby they stopped responding to verbal commands and were unable to engage in conversation, the level of propofol was maintained. Participants remained capable of spontaneous cardiovascular function and ventilation.
- 4) Recovery:** Propofol administration was terminated after acquisition in deep sedation. Approximately 11 minutes afterwards, participants reached level 2 on the Ramsey scale, which was marked by clear and quick responses to verbal commands.

The mean estimated effect-site propofol concentration was 2.48 (1.82–3.14) $\mu\text{g}/\text{ml}$, and the mean estimated plasma propofol concentration was 2.68 (1.92–3.44) $\mu\text{g}/\text{ml}$. The mean total mass of propofol administered was 486.58 (373.30–599.86) mg. The variability of these concentrations and doses is typical for studies of the pharmacokinetics and pharmacodynamics of propofol [5,6].

Prior to initiating fMRI acquisition, three independent assessors (two anesthesiologists and one anesthesia nurse) evaluated participants with the Ramsay scale. Participants were also asked to perform a basic verbal recall memory test and a computerized (4 min) auditory target detection task, which further assessed each participant's wakefulness/sedation level independently of the anesthesia team. Scanning commenced only after agreement on the wakefulness/sedation level among the three anesthesia assessors.

6.1.3. fMRI Acquisition and Neuroimaging Data Preprocessing

Echo-planar sequencing was used to acquire functional images with the following properties: 33 slices, voxel size: $3 \times 3 \times 3 \text{ mm}^3$, inter-slice gap of 25%, TR = 2000 ms, TE = 30 ms, matrix size = 64×64 , FA = 75° . Resting-state scans had 256 vol. We also obtained an anatomical scan using a T1-weighted 3D MPRAGE (Magnetization Prepared - Rapid Gradient Echo) sequence with the following properties: 32 channel coil, voxel size: $1 \times 1 \times 1 \text{ mm}^3$, TE = 4.25 ms, matrix size = $240 \times 256 \times 192$, FA = 9° .

T1 images were preprocessed using the following toolboxes: SPM (<http://www.fil.ion.ucl.ac.uk/spm>), FSL (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>), SimpleITK (<http://www.simpleitk.org/>) and Dipy (<http://nipy.org/dipy/>). Preprocessing for T1-weighted imaging consisted of the following: manual removal of the neck, removal of non-brain tissue using the FMRIB Software Library (FSL), correction of non-uniformity in low frequency intensity based on the N4 bias field correction algorithm (obtained from SimpleITK), image denoising with the nonlocal means algorithm from Dipy, and spatial normalization to standard stereotactic Montreal Neurological Institute (MNI) space using the SPM12 normalization algorithm. The three initial volumes were discarded to avoid T1 saturation effects in the fMRI data. Head motion and slice timing correction were performed using the MCFLIRT algorithm from FSL. We then ran artifact outlier detection, followed by artifact correction using RapidArt (<https://www.nitrc.org/projects/rapidart/>). fMRI data were subsequently co-registered onto the T1 image and spatially normalized to the MNI space with the SPM12 normalization algorithm. Finally, spatial smoothing was applied to the fMRI data using a Gaussian kernel (8 mm full width at half maximum as implemented in SPM12).

6.1.4. Extraction of Representative Regions from RSNs

The regions constituting each RSN were obtained from a cortical parcellation scheme based on resting-state correlations [7]. Cortical areas were grouped into representative regions of interest (ROI), which were selected from five clusters by running a k-means algorithm over the spatial centroids of each RSN. Five regions were included in each extracted network to maintain reasonable balance between spatial resolution and computational complexity. Note that because calculations of Φ have a scalability of

$O(n53^n)$, with n being the number of nodes (cortical regions), larger networks would have resulted in computation times. The RSNs extracted were the Default Mode, Frontoparietal, Retrosplenial, Dorsal, Ventral, Cingulo-opercular, Cingulo-parietal, Auditory, Visual, Sensorimotor Hand, and Sensorimotor Mouth networks. The ROIs extracted for each network, along with descriptions of their functions, are presented in the supplementary material.

Once the five representative regions of each network were obtained, their means throughout the time-series were computed. The time-series were then cleaned by removing spurious variance by means of nuisance signal regression, which was based on the average time series of external regions of noninterest (white matter and cerebrospinal fluid) [8]. A rigid body transformation, which was obtained from head-motion correction with FSL, yielded six motion parameters for translation and rotation in three dimensions, which were also included in the nuisance regressors. The time-series were then detrended and filtered using a bandpass Butterworth filter with cut-off frequencies set at 0.01 Hz and 0.1 Hz. Finally, we standardized the time-series to have zero mean.

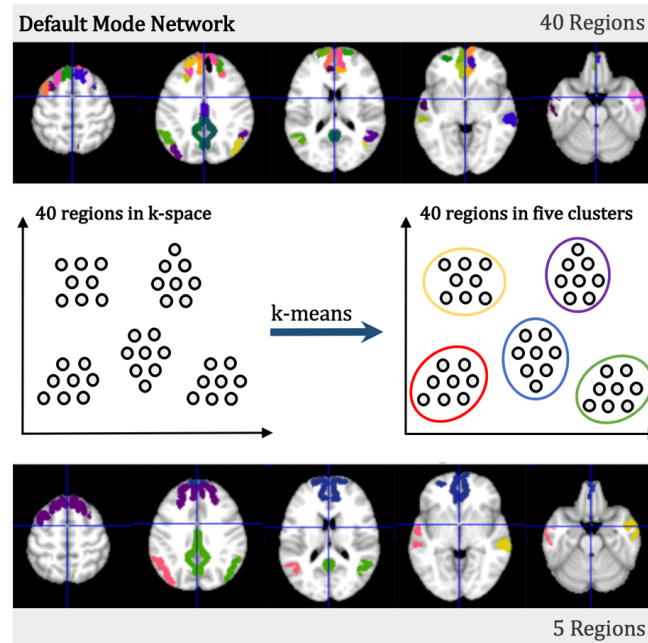


Figure 6.1.1. Obtaining five representative regions for an RSN. Starting with 40 regions in the DMN, a k-means algorithm was used to obtain five clusters, each associated with a centroid. These were used to obtain a total of five time-series.

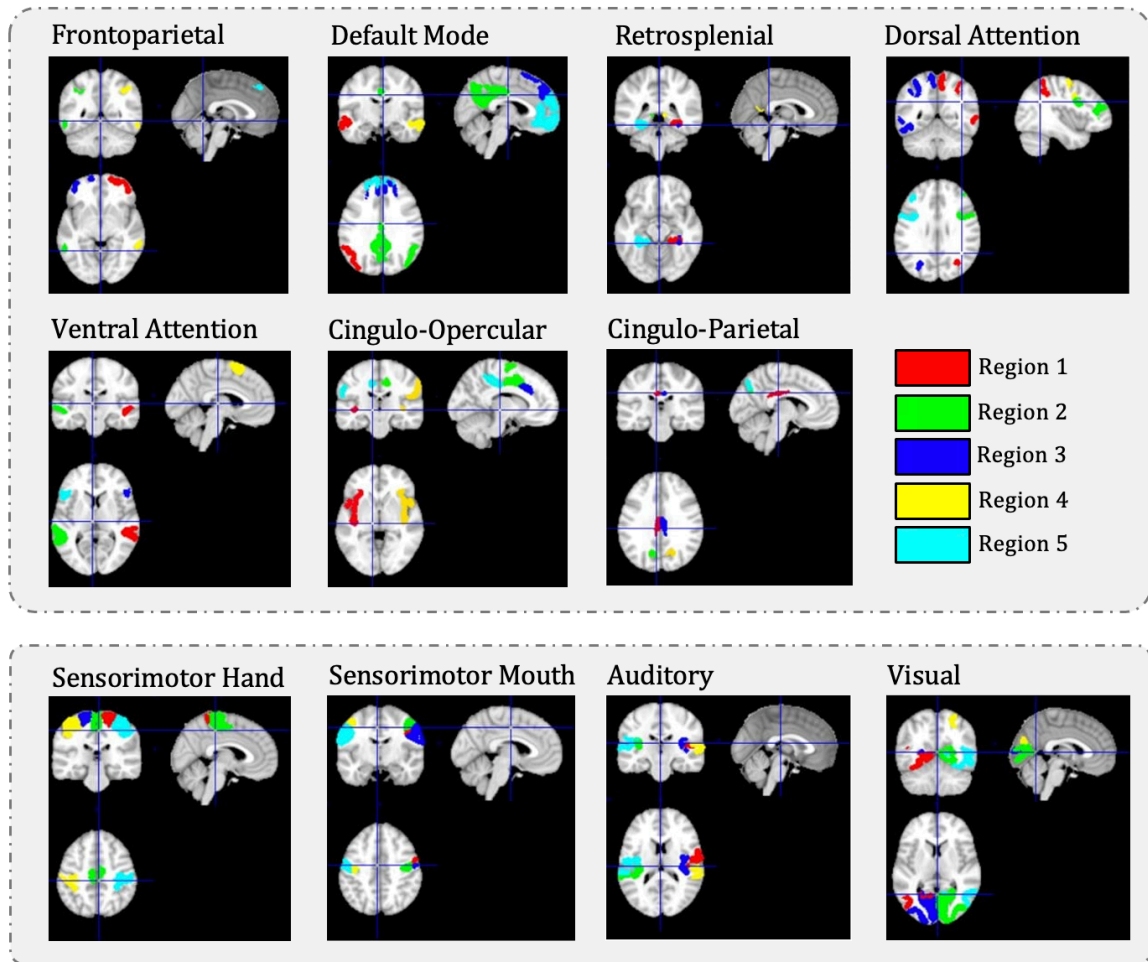


Figure 6.1.2. Representative cortical regions of the 11 RSNs from the parcellation scheme. The five regions obtained for each RSN are presented in the figure above and are separated into the two categories of RSNs we described.

6.2. Propofol Administration and fMRI Acquisition

6.2.1. Markov Property Test

The results for the Markov property test (section 2.2.3) are presented in the table below, where we give the chi-square statistics obtained from each network's contingency table.

	χ^2	ν	χ^2/ν	p	
Frontoparietal	35.66	28	1.27	0.15	<i>Awake</i>
	22.25	28	0.79	0.77	<i>Mild</i>
	29.24	29	1.01	0.45	<i>Deep</i>
	10.52	30	0.35	1.00	<i>Recovery</i>
Default Mode	23.16	28	0.83	0.72	
	16.46	27	0.61	0.94	
	29.59	28	1.06	0.38	
	25.46	30	0.85	0.70	
Retrosplenial	21.94	29	0.76	0.82	
	10.85	22	0.49	0.98	
	29.33	29	1.01	0.45	
	41.82	27	1.55	0.03	
Dorsal Attn.	16.13	28	0.58	0.96	
	17.68	30	0.59	0.96	
	15.43	25	0.62	0.93	
	8.52	28	0.30	1.00	
Ventral Attn.	34.41	22	1.56	0.04	
	34.48	23	1.50	0.06	
	22.99	24	0.96	0.52	
	16.57	29	0.57	0.97	
Cing. Opercular	8.18	31	0.26	1.00	
	13.33	30	0.44	1.00	
	21.72	31	0.70	0.89	
	11.35	29	0.39	1.00	
Cing. Parietal	45.11	27	1.67	0.02	
	27.28	29	0.94	0.56	
	19.34	29	0.67	0.91	
	15.09	30	0.50	0.99	

	χ^2	ν	χ^2/ν	p
Auditory	22.43	30	0.75	0.84
	33.61	31	1.08	0.34
	42.94	31	1.39	0.08
	31.18	29	1.08	0.36
SM Mouth	26.26	30	0.88	0.66
	32.52	29	1.12	0.30
	38.78	30	1.29	0.13
	22.70	29	0.78	0.79
Auditory	22.76	31	0.73	0.86
	41.19	29	1.42	0.07
	17.43	30	0.58	0.97
	20.10	31	0.65	0.93
Visual	17.34	31	0.56	0.98
	24.75	30	0.82	0.74
	36.24	31	1.17	0.24
	21.78	30	0.73	0.86

Table 6.2.1. Results for the Markov property contingency table test. Each block corresponds to an RSN, with four rows corresponding to the four states of awareness (awake, mild, deep, recovery). The first column gives the χ^2 values obtained from equation 1, the second column gives the degrees of freedom ν (number of sequences – 1), the third column gives a reduced χ^2 value, or χ^2/ν , and the fourth column gives p values. Time-series yielding statistically significant differences ($p < 0.05$) were considered to violate the Markov property and their p -values are highlighted with red text.

6.2.2. Modulation of $\mu[\Phi^{\max}]$ by Propofol

For each comparison of $\mu[\Phi^{\max}]$ in the same network (i.e., awake vs. mild sedation, mild vs. deep sedation, etc.), we present all statistics obtained in the set of tables below. These were obtained using Welch’s t-test on a sample of 17 time-series for each RSN and condition, which was obtained by concatenating 16 subjects at a time.

Frontoparietal

		<i>M</i>	<i>D</i>	<i>R</i>		
<i>A</i>		1.17	4.3	-4.12	<i>t</i>	
		31.94	30.43	31.99	<i>v</i>	
		0.25	0.00016	0.00025	<i>p</i>	
	<i>M</i>		3.2	-5.19		
				30.93	31.98	
				0.00318	0.00001	
		<i>D</i>		-7.9		
				30.64		
				0		

Default Mode

		<i>M</i>	<i>D</i>	<i>R</i>		
<i>A</i>		-0.35	2.14	-1.75		
		27.96	30.38	31.9		
		0.73	0.04007	0.0894		
	<i>M</i>		3.04	-1.7		
				31.17	27.06	
				0.0047	0.1	
		<i>D</i>		-4		
				29.62		
				0.00039		

Retrosplenial

		<i>M</i>	<i>D</i>	<i>R</i>		
<i>A</i>		-0.02	1.06	-0.15		
		31.26	30.69	31.56		
		0.98	0.3	0.88		
	<i>M</i>		1.18	-0.15		
				31.9	31.96	
				0.25	0.89	
		<i>D</i>		-1.31		
				31.74		
				0.2		

Dorsal Attention

		<i>M</i>	<i>D</i>	<i>R</i>		
<i>A</i>		1.74	7.21	3.69		
		31.87	29.06	31.89		
		0.0912	0	0.00083		
	<i>M</i>		5.49	2.11		
				29.97	31.54	
				0.00001	0.04282	
		<i>D</i>		-2.73		
				28.16		
				0.01072		

Ventral Attention

		<i>M</i>	<i>D</i>	<i>R</i>		
<i>A</i>		-0.18	1.04	0.85		
		30.45	30.57	27.19		
		0.86	0.31	0.4		
	<i>M</i>		1.08	0.92		
				27.17	23.73	
				0.29	0.37	
		<i>D</i>		-0.31		
				30.46		
				0.76		

Cingulo-Opercular

		<i>M</i>	<i>D</i>	<i>R</i>		
<i>A</i>		0.94	0.24	0.19		
		29.4	19.88	31.27		
		0.35	0.81	0.85		
	<i>M</i>		-1.12	-0.81		
				22.91	31.28	
				0.28	0.42	
		<i>D</i>		0		
				21.21		
				1		

<i>Cingulo-Parietal</i>			
	<i>M</i>	<i>D</i>	<i>R</i>
<i>A</i>	1.29	1.38	0.7
	30.8	31.26	31.97
	0.21	0.18	0.49
	<i>M</i>	0.12	-0.5
		31.93	30.43
		0.9	0.62
		<i>D</i>	-0.6
			30.96
			0.55

<i>Visual</i>			
	<i>M</i>	<i>D</i>	<i>R</i>
<i>A</i>	0.27	-1.48	-0.2
	25.11	31.88	28.96
	0.79	0.15	0.84
	<i>M</i>	-1.28	-0.4
		24.21	30.25
		0.21	0.69
		<i>D</i>	1
			28
			0.33

<i>Sensorimotor Mouth</i>			
	<i>M</i>	<i>D</i>	<i>R</i>
<i>A</i>	1.22	0.14	-0.63
	30.35	29.82	30.87
	0.23	0.89	0.53
	<i>M</i>	-1.26	-1.76
		31.95	27.46
		0.22	0.09002
		<i>D</i>	-0.82
			26.83
			0.42

<i>Sensorimotor Hand</i>			
	<i>M</i>	<i>D</i>	<i>R</i>
<i>A</i>	-0.24	2.75	1.23
	32	30.27	31.96
	0.81	0.00992	0.23
	<i>M</i>	2.96	1.47
		30.29	31.96
		0.00593	0.15
		<i>D</i>	-1.72
			29.82
			0.09601

<i>Auditory</i>			
	<i>M</i>	<i>D</i>	<i>R</i>
<i>A</i>	2.39	-1.62	-2.12
	29.91	28.82	31.81
	0.02361	0.12	0.04201
	<i>M</i>	-3.5	-4.25
		31.82	28.76
		0.0014	0.0002
		<i>D</i>	-0.04
			27.59
			0.96

Table 6.2.2. Results for statistical comparison of $\mu[\Phi^{\max}]$ over the four states of awareness in each of the 11 RSNs. A: awake; M: mild; D: Deep; Recovery: R; Welch's t-test was used to obtain t , ν (degrees of freedom), and p values, which are given in this respective order within each comparison block (highlighted with the bolded borders). Statistically significant p values are highlighted in red ($p < 0.05$). Positive t values indicate a decrease of $\mu[\Phi^{\max}]$, whereas negative values indicate an increase.

6.3. References

- [1] Naci L, Haugg A, MacDonald A, Anello M, Houldin E, Naqshbandi S, et al. Functional diversity of brain networks supports consciousness and verbal intelligence. *Scientific Reports* 2018;8:13259. <https://doi.org/10.1038/s41598-018-31525-z>.
- [2] Kandeepan S, Rudas J, Gomez F, Stojanoski B, Valluri S, Owen AM, et al. Modeling an auditory stimulated brain under altered states of consciousness using the generalized Ising model. *NeuroImage* 2020;223:117367. <https://doi.org/10.1016/j.neuroimage.2020.117367>.
- [3] Marsh B, White M, Morton N, Kenny GNC. Pharmacokinetic Model Driven Infusion of Propofol in Children. *British Journal of Anaesthesia* 1991;67:41–8. <https://doi.org/10.1093/bja/67.1.41>.
- [4] Ramsay MAE, Savege TM, Simpson BRJ, Goodwin R. Controlled Sedation with Alphaxalone-Alphadolone. *BMJ* 1974;2:656–9. <https://doi.org/10.1136/bmj.2.5920.656>.
- [5] Nimmo AF, Absalom AR, Bagshaw O, Biswas A, Cook TM, Costello A, et al. Guidelines for the safe practice of total intravenous anaesthesia (TIVA). *Anaesthesia* 2019;74:211–24. <https://doi.org/10.1111/anae.14428>.
- [6] Sukumar V, Radhakrishnan A, Keshavan V. Effect site concentration of propofol at induction and recovery of anaesthesia - A correlative dose-response study. *Indian Journal of Anaesthesia* 2018;62:263. https://doi.org/10.4103/ija.IJA_670_17.
- [7] Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cerebral Cortex* 2016;26:288–303. <https://doi.org/10.1093/cercor/bhu239>.
- [8] Fox MD, Snyder AZ, Vincent JL, Corbetta M, van Essen DC, Raichle ME. From The Cover: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences* 2005;102:9673–8. <https://doi.org/10.1073/pnas.0504136102>.

Curriculum Vitae

Idan E. Nemirovsky

Education

Master of Science in Physics Sept. 2020 – Apr. 2022
 Physics and Astronomy Department, Western University, London, Ontario, Canada
 Thesis title: *An Implementation of Integrated Information Theory in Resting-State fMRI*
 Supervisor: Dr. Andrea Soddu

Bachelor of Science, Honors in Physics Sept. 2016 – Apr. 2020
 Physics and Astronomy Department, Western University, London, Ontario, Canada
 Thesis title: *The Application of the Boundary Element Method to Concomitant Field Reduction in MRI*
 Supervisors: Dr. William Handler, Dr. Blaine Chronik

Honors and Awards

Natural Sciences and Engineering Research Council of Canada (NSERC)	Apr. 2021 – Apr. 2022
Canada Graduate Scholarship – Master’s Program	
Ontario Graduate Scholarship (Declined)	Apr. 2021
The Raymond Compton Dearle Gold Medal	June 2020
Lynn Fordham Award in Science and Engineering	June 2020
Bennie and Shirley Bradshaw Award in Science	Sept. 2019
Josh Gordon McIntosh Scholarship in Physics (Western University)	Sept. 2019
NSERC Undergraduate Research Summer Award	May 2018 – August 2018
Collaborative Specialization in Musculoskeletal Health Undergraduate Award	May – August 2018
Four Year Continuing Admissions Scholarship (Western University)	Sept. 2016 – Sept. 2020

Publications

Porcaro, C., Nemirovsky, I. E., Riganello, F., Mansour, Z., Cerasa, A., Tonin, P., Stojanoski, B., & Soddu, A. (2022). Diagnostic Developments in Differentiating Unresponsive Wakefulness Syndrome and the Minimally Conscious State. *Frontiers in Neurology*, 12. <https://doi.org/10.3389/fneur.2021.778951> (*shared first authorship*)

Cortese, M. D., Arcuri, F., Nemirovsky, I. E., Lucca, L. F., Tonin, P., Soddu, A., & Riganello, F. (2021). Nociceptive Response Is a Possible Marker of Evolution in the Level of Consciousness in Unresponsive Wakefulness Syndrome Patients. *Frontiers in Neuroscience*, 15. <https://doi.org/10.3389/fnins.2021.771505>

Idan Nemirovsky. (2020). Tehran and Yalta: Failures in Devising Post-War Poland and Germany. *The Mirror: Undergraduate History Journal*, 40, 82–88.

Conference Presentations

Physics Undergraduate Conference Mar. 2020

Oral Presentation

Talk title: The Application of the Boundary Element Method to Concomitant Field Reduction in MRI

Imaging Network Ontario Mar. 2019

Oral Presentation

Talk title: Vibrometer and Accelerometer Measurements of Induced Vibration in MRI

Teaching Experience

Physics 1502B (Enriched Introductory Physics II) Jan. 2022 – Apr. 2022

Department of Physics and Astronomy, Western University, London, Ontario, Canada

Physics 1501A (Enriched Introductory Physics I) Sept. 2021 – Dec. 2021

Department of Physics and Astronomy, Western University, London, Ontario, Canada