Computational network models for molecular, neuronal and brain data in the presence of long range dependence

James Wilsenach

St Cross College

University of Oxford

A dissertation submitted for the degree of

Doctor of Philosophy

Michaelmas 2021

For my grandfather, James

Acknowledgements

I would like to thank my supervisors Gesine Reinert and Charlotte Deane for their unfailing support and guidance during all the challenges and setbacks of the doctoral process. In particular, I have relied on Gesine's kindness and patience, and Charlotte's astute observations and approach to science throughout. Both of them have helped me immensely to become a better statistician, a better communicator and a better researcher. In a similar vein, I would especially like to thank Katie Warnaby for always being available to discuss research and for giving me so many amazing new opportunities to pursue new ideas, such as with the great team at Neuroelectrics.

I will always be grateful for the way Katie welcomed me into her small team of enthusiastic consciousness researchers, with special thanks to Marco Fabuš and Jostein Holmgrem for their insights and assistance. I would also like to thank my collaborators Nigel Emptage and Peter Haslehurst, with special thanks to Peter for his commitment and friendship. Thanks also goes to Rudi Tong for his incredible kindness and passion as an experimenter.

I have been lucky to have made some incredible friends while at Oxford who have helped me through it all. I would like to especially thank Michael and Jamie for being the home I could always rely on. Thanks to Anne for being my COVID confidant. For all the amazing chats that kept me sane and inspired, I would like to thank Greg, Zach, Theo, Martin, Maxence, Donovan and so many others.

I am grateful to my mother, Yvette, for always, always being there with the support and care we needed and fostering an interest in the world. Thanks to my father, Wynand, for pushing us to question everything and push past bad arguments. My brother Greg, can't be thanked enough for showing me what I am capable of. I would especially like to thank my grandparents James and Maryna for their love and support and my grandfather, James, for sharing all his stories and passion for science and engineering with us. Thanks to my partner Cathy who supported me in every way she could. Her patience and love got me to this point.

Abstract

Standard parametric statistical approaches based on comparison to global activity tend to perform poorly when this activity varies over multiple scales. Such multiscale variation, termed long range dependence, is a well-documented features of many biological and neurological data sets. We provide evidence from the literature as well as from data that demonstrates long range dependence across three contexts in: protein, brain and neuronal data. We propose novel non-parametric statistical approaches that account for these dependencies using network models.

In networks of annotated proteins (nodes) connected by their physical interactions (edges), communities of functionally related proteins can provide possible novel drug targets. However, this guilt by association approach is vulnerable to potential bias in protein annotations. By restricted random-walk-based ranking of communities, our method, CommFinedWalker, preserves protein co-expression within well-ranked communities (suggesting functional conservation). The method provides novel drug target candidates that add to the diversity of interesting proteins and potential drug targets for research. As proof of concept, we explore a case study of a community that contains proteins involved in a rare genetic disorder. These results also serve as inspiration for novel methods to rank interesting communities in other contexts including brain state networks.

Brain states have previously been studied across many conditions, including sleep and disorders of consciousness using a variety of models. However, brain state has not yet been studied in pharmacologically-induced unconsciousness. We define a new subclass of models, Hidden Markov Graph Models, to study brain states and communities of brain regions in fMRI data from subjects in both wakefulness and anaesthesia-induced unconsciousness. The framework draws on principles of free energy minimisation and entropy maximisation. We show evidence of characteristic resting state network activity in wakefulness as well as states of reduced temporal and functional complexity that are unique to anaesthesia-induced unconsciousness. These states further relate to established EEG markers of unconsciousness and long range dependent slow waves with implications for effective anaesthesia concentration in surgery.

Lastly, we propose a pipeline to guide experimentation and perform automated detection of synaptic calcium events from single neuron light sheet microscopy data. This pipeline iteratively combines experimentation with experimentally informed simulation to improve experimental outcomes. The model fitting framework we provide draws on the free energy principle, and detection and smoothing methods that account for long range dependence and non-stationary activity present in the microscopy background. Through simulated results we empirically show that our methods can provide consistent estimates of actual experimental conditions that could be used to inform future experimentation.

Author Contributions

All work in this thesis has been produced by myself, except where otherwise indicated here, or acknowledged via citation. Chapter 4 is expected to be published as a co-authored journal article with myself as first author (for full details of each authors contributions see Chapter 4, Acknowledgements). In Chapter 4 and 5, fMRI data pre-processing was performed by Katie Warnaby while EEG SWAS data were produced by Marco Fabus (University of Oxford, Nuffield Department of Clinical Neuroscience). All neuronal microscopy images were produced and deconvolved (where relevant) by Peter Haslehurst (University of Oxford, Department of Pharmacology). Figure 3.2 is adapted from [236], Figure 6.2 is adapted from [364] while Figure 6.4 and 6.5 were adapted from [167].

Contents

Glossary of Terms and Notation

Mathematical Notation

$\mathbb{N}_k = \{n : n \in \mathbb{N}, 1 \leq n \leq k\}$

Terms and Acronyms

ALC	Average Local Clustering coefficient
AP	Action Potential
bAP	backpropagating Action Potential
BOLD	Blood-Oxygen-Level-Dependent
Ca^{2+}	Calcium Ion
CEN	Central Executive Network
CFW	CommFinedWalker
CSF	Cerebral Spinal Fluid
DAG	Directed Acyclic Graph
DMN	Default Mode Network
EEG	Electroencephalography
EPSP	Excitatory Post-synaptic Potential
ESC	Effect-Site Concentration
FH	Functional Homogeneity

FO	Fractional Occupancy
FOV	Field of View
fMRI	functional Magnetic Resonance Imaging
GO	Gene Ontology
GC	Global Clustering coefficient
HMGM	Hidden Markov Graph Model
HMM	Hidden Markov Model
LRD	Long Range Dependence
LSM	Light Sheet Microscope
LTD	Long Term Depression
LTP	Long Term Potentiation
MVN	Multivariate Normal
NT	Neurotransmitter
PCA	Principal Component Analysis
PIN	Protein Interaction Network
PSD	Power Spectral Density
ROI	Region of Interest
RSN	Resting State Network
SN	Salience Network
TO	Temporal Occupancy
VGCC	Voltage-Gated Ca^{2+} Channel

1 Introduction and Background

Biological and Neurological Data in the Presence of Long Range Dependence

Molecular biology, neuronal and neural systems display nested organisation and dynamic interactivity [371]. Networks are one of the most powerful and well established frameworks to quantify these types of complex phenomena [8, 40]. In this thesis I will put forward evidence to support the hypothesis that some of the complexity inherent to these systems is due to interaction occurring at multiple scales such as the microscale and mesoscale, or across dimensions such as space and time. This complexity often manifests as a tendency for correlations to persist, decaying slowly over multiple scales (typically sub-exponentially) [65, 427], the term for which is Long Range Dependence (LRD), also termed long range persistence or memory. We explore the consequences of this property in biological systems, the potential issues it presents in extracting meaningful information from networks data, and provide novel approaches to mitigate these issues across three biological scales, molecular, cellular (neuronal) and whole brain data.

Though LRD in a stochastic process does not imply non-stationarity, the relationship is complex and long range dependent signals tend to violate many of the strong stationarity and independence assumptions that are explicitly or implicitly made by standard statistical methods [329]. Markers of this dependence in statistical data include distributions with long or fat tails indicating a relatively high probability of extreme values which can cause false positives [250]. All these factors can lead to bias in estimates based on standard statistical approaches when LRD is present [49].

We focus on one area where mitigating potential bias caused by LRD is an active area of research, biological networks. Deriving methods for mitigating LRD in networks is made particularly difficult by the multi-scale interdependence of the biological systems [310, 441]. We propose non-parametric statistical methods to address these dependencies; methods that respect the unique topological features of these networks. We examine both classical network

models as well as multiplex networks that can account for the multiple scales that give rise to LRD. Key to much of our analysis are the related concepts of Markov chains, random walks, diffusion, entropy maximisation and free energy minimisation [92, 153, 156]. We also utilise and expand upon community detection methods that can uncover the nested structure of these networks [56].

These ideas link information theory and statistical physics to networks biology and neuroscience in ways that draw from the strengths of both frequentist and Bayesian statistical frameworks [22]. We use a combined approach rooted in these ideas to investigate:

1. Novel drug target pathways in protein interaction networks.
2. Subnetworks that are key to brain dynamics in wakefulness as well as deep anaesthetic-induced unconsciousness.
3. Hot spots of local, synaptic calcium ion (Ca^{2+}) spikes in neurons.

In all cases we will support our findings with information from the literature. In some cases annotations based on the literature can themselves act as an additional data layer, termed metadata. Metadata provides context to data by giving a functional description to the members of a data set. Like the biological data itself, metadata can exhibit LRD in its distribution, either due to an asymmetric focus on some entities over others, or due to structural limitations in the way that data is annotated [335]. We introduce the term metaplexity to describe complexity arising from the distribution of functional metadata. In many ways metaplexity can function like an additional layer on top of a multilayer or single layer network [230].

This chapter provides a background that establishes LRD as a phenomenon in biological systems and data (Section 1.1), as well as providing an overview of networks biology and neuroscience (Section 1.2). We introduce relevant concepts in statistics and computational neuroscience and outline our approach for dealing with LRD in networks. Finally, we discuss the structure of this thesis with reference to the chapters to follow.

1.1 Long Range Dependence

The investigation of LRD as a quantifiable phenomenon began with studies of reservoirs along the Nile River [184], since then it has been observed in data from across the natural and social

sciences [110, 352]. LRD can be caused by a wide range of phenomena including self-organising properties of real world systems, so-called Self Organising Criticality (SOC). SOC in particular leads to the kind of slow-decaying, fat-tailed distributions that are characteristic of typical LRD [35,350]. LRD, particularly LRD associated with SOC, is often thought of in terms of very specific power scaling laws. Given a distribution, $p(X = x)$ over all $x \in \mathbb{R}^+$, this power scaling is defined by

$$p(X = x) \propto x^{-a},$$

and the scaling constant $a > 0$.

Demonstrating power law-like scaling is one way to suggest LRD (and one we rely on in Chapters 3 and 6), however LRD does not require strict power law-like scaling. Many system that are dominated by low frequencies or multi-scale nested structure may be considered to exhibit LRD while only exhibiting this characteristic scaling relationship for sufficiently large x [66, 373].

LRD is ubiquitous in biological and neurobiological systems; so much so that it characterises the background of most types of neuroimaging data [16, 424]. Random noise that follows the power law relationship in the distribution of spectral power is known as coloured noise. Pink noise (noise where exponent $1 < a < 2$) is present in both the spatial and temporal domains in neurobiological data [219]. However, it is not only noise that carries this spectral power property as many important processes appear to be either facilitated by [5], or rely on [285], LRD to perform their function.

In static networks, LRD is often seen in the scaling properties of certain network statistics. Static networks of brain connectivity commonly exhibit LRD and scaling [398]. In proteomics, the study of proteins and their relationships, protein interactivity statistics may exhibit LRD [33, 437]. On the level of metadata, protein annotations provide an important indication of the functional role of proteins. Analysis of the distribution of annotations per protein has shown that the state of information on proteins is dominated by a sparsity of information on most proteins and detailed annotations on only a few. These annotations per protein follow a power law-like relationship [335].

The problem of annotational asymmetry on proteins is difficult to assess as it is unclear what

fraction reflects the difference in the actual relevance of proteins to researchers (which is unlikely to be equal across proteins) and what fraction is due to self-perpetuating sampling bias in the selection of experimental targets. Sampling bias of this kind is known to effect the reproducibility of experiments in biology and neuroscience [82], and unbiased sampling is particularly difficult on networks as many methods do not preserve the properties of the original network [438].

1.2 Networks in Biological Systems

In systems biology, the study of whole biological systems on the level of of their interactions, networks or graphs are structures composed of a set of nodes connected by edges [425]. These nodes could be genes, brain regions, proteins or compounds while the edges could represent chemical, environmental or behavioural interactions. Edges do not need to represent physical interactions and could represent abstract forms of interactions between entities such as correlation in the activity of genes (termed co-expression) or shared information between brain regions (termed functional connectivity).

Network biology has revealed common themes in the structure of biological networks including LRD and the tendency towards the formation of functional modules. Functional modules are composed of entities that work together to perform some common task or structure. In networks, functional modules are subnetworks (parts of a larger network) whose members organise together to perform important tasks, or carry out complex roles or behaviours [309, 323, 407]. Modules can combine hierarchically to perform ever more complex functions [37, 437]. Typically, members of the same module share similar properties, such as a tendency towards intramodular connectivity. We look at the structure of networks at multiple scales to determine which proteins, brain regions or neuronal substructures come together to perform functions in the body and brain. Networks analysis is integral to our reasoning and in Chapter 2, we will develop the necessary networks analysis tools used throughout this thesis. Here, we provide very brief introductions to the fields of networks in molecular biology and neuroscience.

1.2.1 Networks in Molecular Biology

Figure 1.1 shows a simple example of a Protein Interaction Network (PIN). PINs are networks in which the nodes are proteins and the edges are physical or chemical interactions between them [316]. This PIN is centred on RPB2 which is a protein forming part of the DNA-directed

RNA polymerase II protein complex in yeast [188]. This subnetwork also includes interacting proteins RPB3, 7 and 9, which also form part of the same functional module, the protein complex.

Different types of interactions between the same entities can provide complementary information on their function [164]. An example is protein interaction and co-expression networks. In this example the presence of a protein interaction edge indicates the possibility for chemical binding between proteins, while co-expression edges indicate spatial or temporal correlation between them. The additional information given by co-expression provides evidence of whether the proteins are co-located and hence may have an opportunity to interact. Understanding how and why proteins interact to perform biological functions is a central and on-going research problem, one that is made more complex by the impact of LRD, metaplexity and bias on protein networks [236].

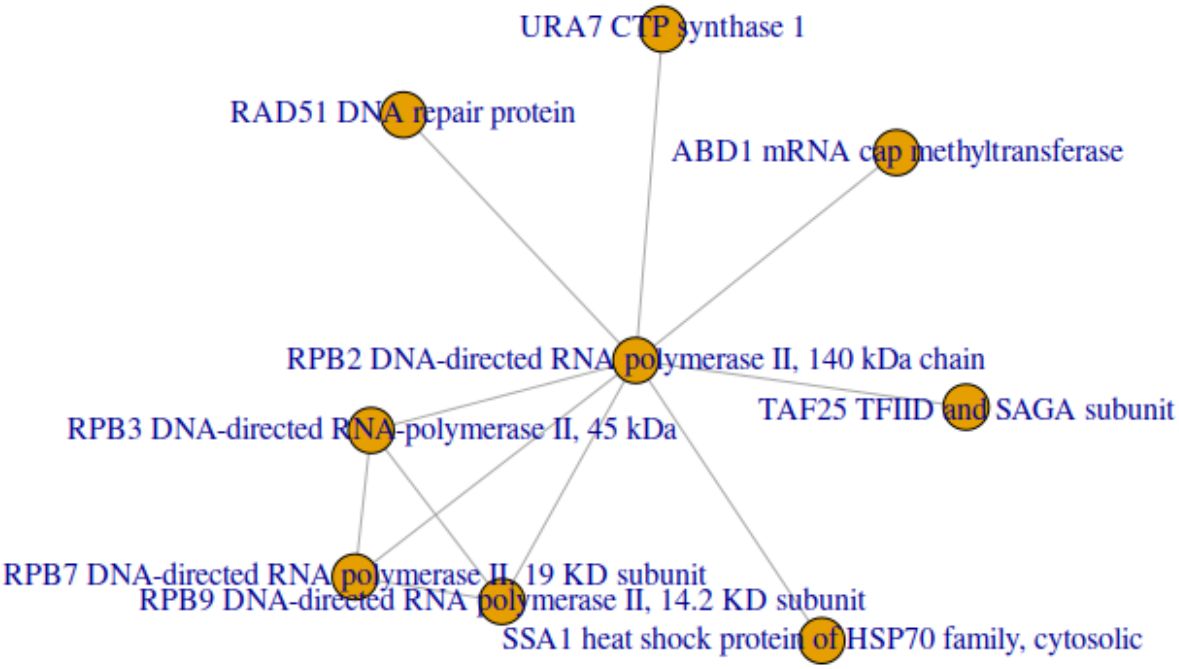


Figure 1.1: This figure shows a small subnetwork of a yeast Protein Interaction Network (PIN) network accessed from [41], and based on data from [70]. Node labels indicate the names of proteins. The network is formed (induced) by considering the immediate interactions of the protein RPB2 DNA-directed RNA polymerase II, which is formed from a 140 kilodalton (kDA) amino acid chain. Induced subnetworks and PINs are explained in more detail in Chapters 2 and 3 respectively.

1.2.2 Network Neuroscience

Network neuroscience is a relatively young area of computational neuroscience which focuses on the interaction between different parts of the brain at multiple scales [40]. Network neuroscience

has benefited immensely from from the availability of data from functional neuroimaging studies. This data may come from one or more of several modalities including Electroencephalography (EEG), Magnetoencephalography (MEG) and functional Magnetic Resonance Imaging (fMRI) [366]. In recent years, dynamic whole brain network analysis has come to rely more and more on fMRI which has the spatial resolution and deep brain coverage needed to resolve dynamic global and regional shifts in activity and connectivity. Though the temporal resolution is lower than MEG, MEG struggles to resolve deep brain sources of network activity [231]. Here, network connectivity is defined by shared information (correlation or mutual information) or phase coupling between brain regions, indicating some level of coordination between them [102].

Despite its relatively recent development, network neuroscience has produced remarkable insights into the function and structure of the brain. The term network in network neuroscience is overloaded as the term can also mean a part of a network or subnetwork (as in the case of the so-called Resting State Networks), for this reason we use the terms graph or graph model to refer to connectivity at the whole brain level and network to refer to important parts (subnetworks) of this whole brain graph.

Anatomical Structure and Connectivity

Neural anatomists divide the brain into cortical (dorsal) and subcortical (ventral) regions as well as left and right hemispheres. The expansive cortex seen in humans is a relatively novel evolutionary development, and many uniquely human higher level brain functions are very dependent on cortical activity, such as language, reasoning and autobiographical thought [112]. Different brain regions within the cortex and subcortex are responsible for different functions and behaviours, with subcortical regions like the thalamus and basal ganglia acting as important bridges for sensory information to reach the cortex [45].

Functional Brain Networks and Brain States

Through mapping the interactions of anatomically defined brain regions, neuroscientists have uncovered core modules composed of multiple regions that work together to carry out specific functions [76,363]. Study of dynamic global changes in whole brain networks in response to task stimulation or internal modulation requires high spatial resolution. For this reason, studies of whole brain networks in recent years have largely used fMRI [400], however, multi-modal studies

that record two or more modalities simultaneously are becoming more common [177, 185, 262].

Figure 1.2 shows a brain map (3D activity map) of the Default Mode Network (DMN), one of the so-called Resting State Networks (RSNs). The image was compiled by performing a metaanalysis of 3D brain activity images from over 10,000 fMRI studies [436], and shows a general consensus on the regions agreed to be part of the DMN. Appendix A.1 expands on the DMN and other RSNs. These are important to the understanding of patterns of activity in network neuroscience for fMRI. As a related concept, brain states are a reconfiguration of brain activity evoked in response to a stimulus or to facilitate more complex responses [68]. Functional networks and brain states are expanded upon on in Chapter 4 [141].

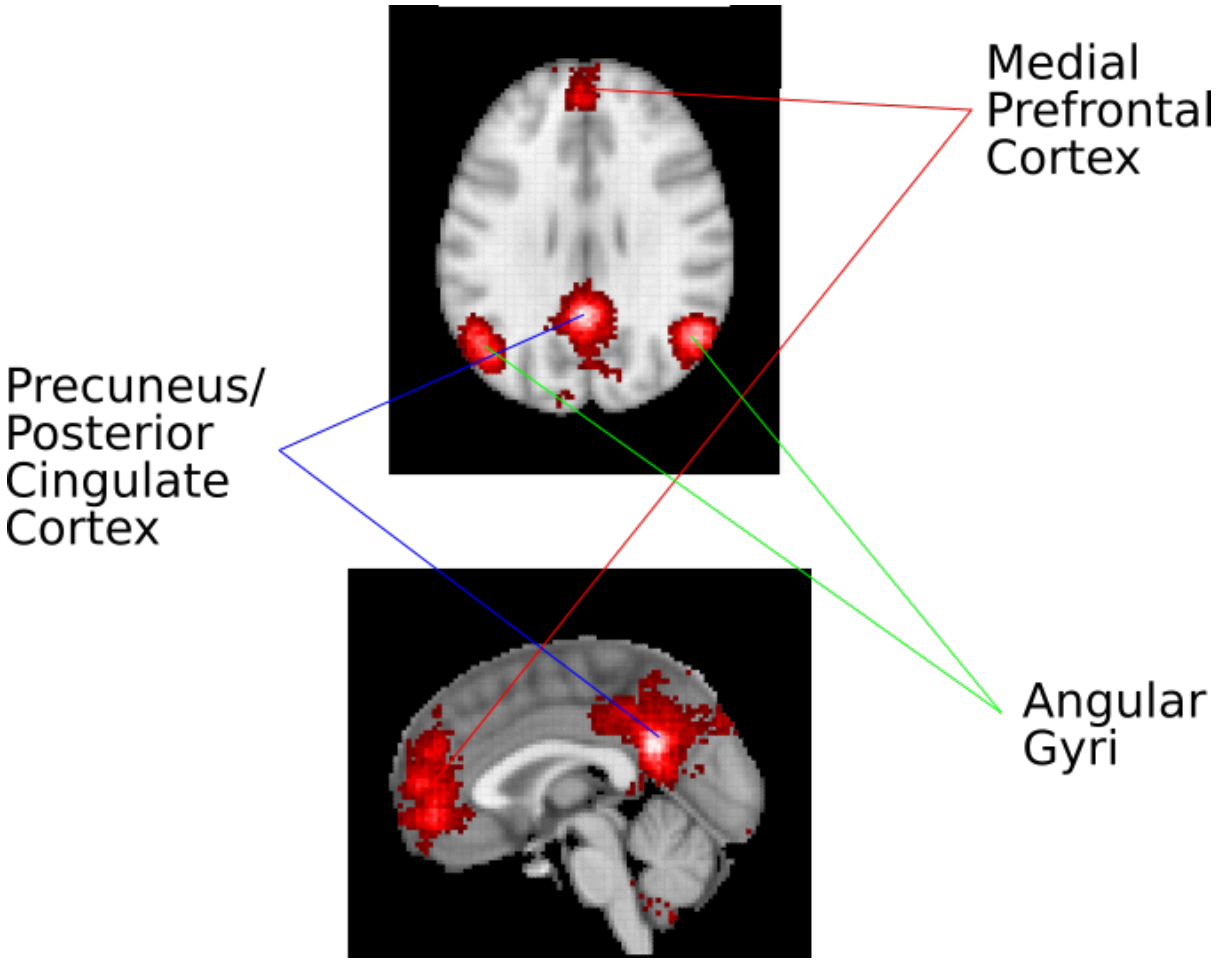


Figure 1.2: This figure shows the core regions involved in the Default Mode Network (DMN) from an axial (top) and a sagital (bottom) view. The DMN is an important Resting State Network involved in maintenance of the self, self-concept and autobiographical information [20]. The core nodes are the neighbouring precuneus and posterior cingulate cortices, the medial prefrontal cortex and the angular gyri. This brain map (3D image of brain activity) was produced using a composite of image data from over 10,000 functional Magnetic Resonance Imaging (fMRI) studies compiled using the *Neurosynth* package [436].

1.3 Free Energy, Diffusion and Inhomogeneous Information

Free energy is a concept borrowed from statistical physics, where the physicist is generally interested in the distribution of the state, Z, that a system is occupying given some partial and approximate information, X, related to Z, Z is often referred to as a hidden state as it is not directly observed in the data [43, 130]. However, the marginal distribution of $p(X = x)$ is often difficult to compute, making it difficult to find the conditional distribution $p(Z|X = x)$ directly.

Energy, G, in this context refers to the mismatch in expected information between the conditional distribution $p(Z|X = x)$ and a candidate distribution $q(Z)$,

$$G(q(Z)||p(Z|X = x)) = E_q[\log(p(Z|X = x))].$$

where $E_q[\cdot]$ is the expected value under q.

The free energy by contrast is the amount of energy available after taking into account the energy unavailable due to it being *lost to entropy*. In statistical terms the entropy term is given by the Shannon information entropy,

$$H(q) = E_q[\log(q(Z))].$$

In statistical physics the free energy F can be seen as the energy available to do work in the system that is not lost to entropy. The free energy in statistics is closely related and represents the available slack in the distribution that is not lost to the information entropy H. It is given by

$$F(q(Z)||p(Z|X = x)) = G(q(Z)||p(Z|X = x)) - H(q) = E_q[\log(p(Z|X = x))] - H(q).$$

Exact minimisation of the free energy is equivalent to Bayesian inference [42]. However, in order to limit the search space, q is usually chosen to belong to some tractable family of distribution parametrised by $\theta \in \Theta$. The goal in approximate free energy minimisation (usually termed Variational Bayesian Inference) is then to find

$$\arg\min_{\theta \in \Theta} F(q(Z|\theta)||p(Z|X = x)).$$

An important alternative formulation of F is as the Kullback-Liebler divergence (also called relative entropy) between $q(Z|\theta)$ and $p(Z|X = x)$, $D_{KL}(q(Z)||p(Z|X = x))$, which measures the expected information required to encode $q(Z)$ given $p(Z|X = x)$,

$$F(q(Z)||p(Z|X = x)) = D_{KL}(q(Z)||p(Z|X = x)) = E_q\left[\log\left(\frac{q(Z)}{p(Z|X = x)}\right)\right].$$

Estimation of distributions with hidden state variables is crucial to the concept of brain state used in Chapters 4 and 5 where the true brain state of a subject is likely hidden. Free energy has been proposed as a unifying principle in neuroscience and biology, which we describe in more detail in Chapter 4, Appendix C.1) and 6. According to this theoretical framework, all living systems seek to minimise the available free energy and neurological systems seek to minimise the free energy between their own internal representation of the state of their environment and the survival-relevant features of the changing external environment itself [140]. In Chapter 4, we extend this framework to account for intersubject variability in state and environment. In Chapter 6 we show that the free energy framework provides a consistent model selection criterion when considering a range of models to approximate local synaptic Ca^{2+} spike arrival in neurons.

Free energy connects to the problem of LRD and inhomogeneity in metadata, as in general, systems with highly asymmetric information about the state or properties of that system will tend to have high free energy relative to those with accurate and unbiased information [354]. Information diffusion as it occurs in message passing systems, e.g. within the brain [71,91], tends to reduce the asymmetry in the initial distribution of input messages through diffusion [157]. In fact, as diffusion increases, the entropy tends to increase, decreasing the free energy [94]. As we shall see, this is one of the possible reasons why limited diffusion-based smoothing and random walks (random sampling by diffusion-like network exploration) can improve our ability to choose the correct model by simulating the local diffusion of information in the system.

1.4 Accounting for Long Range Dependence: Local Exploration

In this dissertation we apply the concepts of random walks, free energy and information diffusion on networks to address three central problems in systems biology and neuroscience related to LRD.

1.4.1 Inspection Bias in Protein Networks

Bias in both functional annotations and the way that PINs are constructed may contribute to LRD in protein annotation and network structure, leading to regions of sparse connectivity and annotation and regions of dense connections and high annotation. The portion of this asymmetry that is potentially due to human attentional and experimental bias is known as inspection bias [236]. We perform modularity-based community detection to identify potential protein functional modules while taking the annotation background into account, thereby reducing potential bias. Our approach uses restricted random walks to perform local sampling of network space, allowing the comparison of modules to a restricted background distribution of similarly annotated proteins.

1.4.2 Characterising Conscious and Unconscious Brain State Dynamics

Dynamic models of changing brain state can be trained by combining data from functional neuroimaging recordings of multiple subjects, occupying a number of different brain states. The level of Long Range dependent background activity is known to vary across subjects and conscious conditions [262]. We perform free energy-based training of a Hidden Markov Model (HMM) of brain state dynamics, and produce a method to convert this HMM to a dynamic multiplex network model of brain state dynamics, termed a Hidden Markov Graph Model (HMGM). By assuming a uniform prior over the probability of a subject occupying any one state, we arrive at a novel model selection criterion for HMGMs based on the free energy principle that takes into account variability between subjects. We also use local random walk-based sampling of the brain state activity background to determine functionally conserved modules in wakefulness and deep anaesthesia. This method takes advantage of the multiplex structure of the HMGM to detect communities and takes into account spatiotemporal relationships between brain regions.

1.4.3 Automated Detection of Synaptic Ca^{2+} Events

Neurons are the basic computational units of the brain whose activity determines its eventual behavioural output through summation or integration of incoming electrochemical signals. Integration depends in a non-linear way on the topology of the neuron [61]. Despite this dependence, the exact relationship the location and timing of input signals, termed synaptic events, shares with neuronal integration is not well understood. Light sheet microscopy-based neuronal Ca^{2+}

imaging provides a method to measure the dynamics of synaptic events as transient, local Ca^{2+} spikes [167]. We investigate the properties of real Ca^{2+} activity from pilot data, identifying the extent of LRD in the microscopy background. We then use our observations to motivate a realistic model of neuronal Ca^{2+} which includes realistic cellular topology extracted from 3D neuronal images and Long Range Dependent background activity. Finally, we provide a novel method which approximates information diffusion on neurons in order to perform model selection for a variety of event detection methods. We empirically show that this selection procedure provides a consistent estimator of the true model parameters under realistic sample size constraints.

1.5 Thesis Structure

This thesis is divided into seven chapters. The first chapter introduces the reader to the topics and structure. Chapter 2 is a primer on the relevant networks analysis tools and terms. Chapter 3 provides a novel method for PIN analysis in the presence of inhomogeneous data quality and quantity based on restricted random walks. Chapter 4 provides the methods of HMM model selection and brain state graph model analysis used in both Chapters 4 and 5, and applies these methods to resting state wakefulness data as a proof of concept. This chapter builds on methods developed in Chapter 3 to analyse inhomogeneity and LRD in the spatiotemporal context. Chapter 5 details the application of the methods developed in Chapter 4 to deep propofol anaesthesia and presents new methods to analyse community structure in HMGMs. Chapter 6 is a largely stand-alone chapter that describes the analysis, whole cell simulation, and detection of synaptic Ca^{2+} events in neurons that shares similar methodological themes to Chapter 4, especially with regards to free energy minimisation.

Where all other chapters except Chapter 2 have Appendices. Chapter 4 and 5 are planned publications where Chapter 4 has already been submitted. The appendices for these chapters are therefore titled Supplementary Information. An additional Appendix C has been added to provide model training details not included in the publication. Chapter 4 also includes an foreword and abstract to provide context to the publication and its application to Chapter 5.

2 | Network Analysis

Networks (or graphs) are a powerful tool in data science that can represent the complex, multi-scale and nested relationships present in biological systems [17]. Networks also make for intuitive models of these relationships that allow for the local structure of complex data sets to be easily conceptualised and visualised [272]. Borrowing heavily from multiple disciplines including biology, social science, mathematics and physics, the field has built up a highly versatile set of statistical and analysis tools for answering questions about the properties of complex data. Phenomena such as long range dependence, well established in time series analysis [329], have known network analogues, where the term *scale-free* is used to describe the characteristic power-law relationships seen in many data types [35]. Hierarchical and modular structure is also common to many networks, where the term *community* is used to describe a subnetwork with a higher than expected level of internal connectivity [275]. Here, we give a brief overview of the network concepts important to our study of biological networks.

In molecular biology, networks have played an important role in the study of molecular relationships, such as the relationships between proteins and genes that facilitate the body's response to drugs, disease and other stimuli [89, 174, 411]. Protein Interaction Networks (PINs), which are networks of proteins (nodes) and their physical interactions (edges), have been key to these studies [319]. PINs have a modular and hierarchical organisation making community detection important for understanding these networks [226,235]. Random walks (a type of Markov process or Chain) have emerged as a method to explore the space of possible modules in PINs [236]. Annotations on the role and location of proteins are also important to making sense of protein and module function [232], adding a layer of metaplexity to PIN analysis [230]. A wealth of annotation data exists for PINs in the form of functional, locational and disease-related labels on nodes (proteins) [2, 26, 336]. This meta layer of information is key to understanding function

and bias in PINs (see Chapter 3).

Brain graphs represent relationships (edges) between brain regions or neuronal populations (nodes) [40]. Brain networks have shown many similar properties to proteomic, genomic and metabolomic networks including hierarchical structure and function, as well as the presence of important regions or hubs [277]. Network neuroscience thus shares many of methods and tools with networks in molecular biology [317]. Dynamic networks in neuroscience can also have modular organisation in time [409]. Given the challenge of combining both spatial (brain regions) and temporal dimensions in brain graphs we introduce the reader to the concepts of multiplex networks to be applied in Chapter 4 and 5. As an improvement from earlier community approaches we also briefly acquaint the reader with the concept of consensus community detection (to be employed in Chapter 5). Consensus community detection methods overcomes some of the resolution limitations of our earlier approaches, taking into account the multi-scale nature of our fMRI data sets [216].

Our work on neuronal microscopy relies on the interpretation of neurons as trees (a special type of graph structure with no cycles). Each neuron is subdivided into segments (nodes) that physically connect (edges) to characterise the branching structure of the neuron [272]. We briefly define these and other types of graphs. Information propagation on neurons is complex but can be approximated by a diffusion process [157]. We examine heat diffusion kernels (heat kernels) [211], which we extend to the spatiotemporal domain in Chapter 6 in order to study intra-neuronal calcium dynamics in time and space. This concept relies heavily on the graph Laplacian [272], a matrix that functions as a differential operator on graph structures that has strong relationships with random walks [266].

This chapter will start with a basic introduction to Markov chains, followed by the graph theoretic framework for understanding neurobiological networks, both directed and undirected, as well as adjacency and Laplacian matrices. We include a discussion of network statistics and centrality (hub) measures as important properties of a graph and its components. From this we will cover community detection, specifically Louvain modularity community detection and its directed extensions. We briefly touch on a multi-resolution extension to community detection as well as a brief primer for consensus clustering. We also introduce the concept of network meta-information and multiplex networks, as an extension of classical networks for understand-

ing metaplexity and temporal structure of biological networks. We then give a brief introduction to the heat kernel for diffusion on graphs. Lastly, we briefly and critically discuss some general observations about real world networks and their properties. We use network interchangeably with the term graph except in Chapters 4 and 5 where graph is used to avoid confusion with specific whole brain networks (such as the resting state networks).

2.1 Networks as Data Structures

At the most basic level, a network or graph $G = (V, E, W)$ is a data structure consisting of a set of nodes, V and a set of edges, E. In addition edges and nodes can be weighted or unweighted by W, annotated or unannotated. Equivalently, G can be defined simply by a set of nodes V and an adjacency matrix A (i.e. $G = (V, A)$). We define the different types of networks we analyse in the chapters to follow in terms of their node and edge structure with reference to applications to biological data.

2.1.1 Network Structure

In PIN data, the vertex set V is composed of labels (often integers) or protein names. In neural data, these are brain regions, or neuronal compartments. Edges in a PIN G may be weighted or unweighted, we will deal almost exclusively with unweighted PINs except in the case of co-expression (which can also be seen as a form of edge annotation), directionality or self-loops and represent physical interactions [319]. In our neural datasets, edges can be either directed, undirected, weighted or unweighted and represented shared information between brain regions [40]. An undirected edge is a set of nodes $e_{ij} = \{j, i\} = \{i, j\}$ for $i, j \in V$. Directed edges are sets of distinct, adjacent node pairs that are denoted $e_{i,j} = (i, j)$ for a directed edge from i to j. For our purposes, two directed edges $e_{i,j}, e_{i,j} \in E$ are equivalent to including a single undirected edge $e_{ij} \in V$. Nodes that occur in the edge together are termed neighbours, while the neighbourhood of a node i is the set $N(i) = \{j \in V : e_{i,j} \in E \text{ or } e_{j,i} \in E\}$. This can be extended in the natural way to sets of nodes $U \subset V$,

$$N(U) = \bigcup_{u \in U} N(u)$$

An edge $e_{i,j}$ can have edge weights $W_{i,j} > 0$ or be unweighted, which is equivalent to an edges having a weight $W_{i,j} = 1$. A network composed of only undirected edges, $e_{ij} \in E$, with no self loops (i.e. $i \neq j$), no multi-edges and $W_{i,j} = W_{j,i}$ for all $i, j \in V$, is termed a *simple graph*.

The edges in a network can be isomorphically represented by the matrix A where

$$A_{i,j} = \begin{cases} W_{i,j} & \text{if } e_{i,j} \in E \\ 0 & \text{otherwise} \end{cases}.$$

This provides an alternative way to define a network simply by providing a vertex set and adjacency matrix, $G \equiv (V, E_A, A) \equiv (V, A)$, where $e_{i,j} \in E_A$ if and only if $A_{i,j} \neq 0$.

An important feature of nodes is their degree, a measure of their in-flows and out-flows. The in-degree, d_i^{in}, and out-degree, d_i^{out}, of a node i is defined as

$$d_i^{in} = \sum_{j \in V} A_{j,i} \qquad \text{and} \qquad d_i^{out} = \sum_{j \in V} A_{i,j}. \tag{2.1}$$

In an undirected graph (consisting of only undirected edges), $d_i = d_i^{in} = d_i^{out} = |N(i)|$.

2.1.2 Markov Chains as Networks on Finite State Spaces

In his original paper and letters, Markov developed the Markov chain as a weapon in a theological debate and also used it as a tool to explore the space of Russian poetry [38]. A Markov chain $M = (S, P, p)$ is defined as a stochastic process of transitions through a finite space of states, $\mathcal{S}$ over times $t \in \mathbb{N}$. The core of the Markov chain is the stochastic matrix P, a square matrix of order $|\mathcal{S}|$ in which each row is a probability mass function over $\mathcal{S}$. The probability vector p_0 over $\mathcal{S}$ defines the initial probability of state occupancy of the system at time $t = 0$ (i.e. $P(S_0 = s) = p_s$). We assume here and in all other chapters that M is a discrete time Markov chain that is homogeneous meaning that transitions occur at regular, discrete intervals, $t \in \mathbb{N}$ and also that M is time invariant (i.e. the probability of state transition does not change with t). Each realisation from M at time t, S_t, is a random variable taking values in $\mathcal{S}$ with conditional

distribution

$$P(S_t = s|S_{t-1} = s') = P_{s,s'},$$

which implies that for $t \geq 1$, S_t is conditionally independent of all $\{S_0, S_1, ..., S_{t-2}\}$ conditioned on S_{t-1}. We say that the system *occupies* the state s_t with conditional probability $P(S_t = s|S_{t-1} = s') = P_{s,s'}$.

Markov chains are ubiquitously and fruitfully applied in both the natural and social sciences, with applications including gene sequence prediction, word and spelling prediction and Hidden Markov Models (HMMs) in machine learning [408, 419]. An HMM is a special kind of Markov model in which the state models some hidden (unobserved) property of the system. The state is observed indirectly through a random variable $O(s_t)$ that depends directly on s_t, the state at time t.

Properties of Markov Chains

Given a Markov chain $M = (\mathcal{S}, P, p)$ and states $s, s' \in \mathcal{S}$ [225]:

1. *Markov Network:* The Markov transition matrix P induces a directed, weighted graph $G(M) = (\mathcal{S}, P)$ with nodes in $\mathcal{S}$ and edge weights in P.
2. *Irreducibility:* The probability $P(S_t = s'|S_{t-1} = s) = P^t_{s,s'}$ $\forall t \in [1, \infty)$ where $P^t_{s,s'}$ is the (s, s') entry of P^t. The Markov chain M (equivalently P) is said to be *irreducible* if $\exists t \in [1, \infty)$ such that

$$P(S_t = s'|S_{t-1} = s) = P^t_{s,s'} > 0.$$

3. *Stationary distribution:* A *stationary distribution* of M is a probability vector $\pi = \pi P$.

   Lemma: If P is irreducible then there exists unique stationary distribution $\pi = \pi P$.

   Proof: The proof follows from the observation that $\mathbf{1}$ is an eigenvector of P ($P = \mathbf{1}P$) and the Perron-Frobenius theorem, for full details see [225].
4. *Aperiodicity/Periodicity:* Let $\mathcal{T}(s) = \{t \leq 1 : P^t_{s,s'} > 0\}$. The state s is said to be

aperiodic if

$$\gcd(\mathcal{T}(s)) = 1,$$

otherwise it is periodic with period $\gcd(\mathcal{T}(s))$. The Markov chain M is said to be aperiodic (respectively periodic) if all of its states are aperiodic.

5. *Return Probability*: The probability

$$f_s = P(\exists t \in [1, \infty) : S_t = s)$$

is called the *return probability* of s.

6. *Transient/Recurrent:* The state s is said to be *recurrent* if

$$f_s = 1$$

otherwise it is *transient*. The Markov chain M is said to be recurrent if all of its states are recurrent.

7. *Ergodicity:* If M is irreducible, recurrent and aperiodic then M is called an *ergodic* Markov chain.

   - It can be shown that since M is a finite Markov chain, if M is irreducible and aperiodic then M is ergodic.

8. *Ergodic Theorem:* The following is a special case of the so-called *Ergodic Theorem*. For a proof, see [225].

   Theorem: Suppose M is ergodic, then for any choice of p a probability vector we have that

$$P(S_t = s) \xrightarrow[\infty]{t} \pi_s$$

   Corrolary: In the above setting.

$$P^t_{s,s'} \xrightarrow[\infty]{t} \pi_{s'} \tag{2.2}$$

These last result in Eq. (2.2) suggest that if M is ergodic the stationary distribution π depends only on P and π in turn determines the long run behaviour of M.

The concepts of the Markov chain and, in particular, of the stationary distribution of M has strong links to random walks on graphs such as PINs (Chapter 3) and functional brain activity graphs (Chapter 4), diffusion in neuronal sub-compartments (Chapter 5), and so-called Hidden Markov Models (Chapter 4).

2.1.3 Subnetworks

Any subset $U \subset V$ induces a sub-network, H, where $V(H) = U$ and $E(H)$ contains all the edges between members of U that are present in E;

$$E(H) = \{e_{i,j} : i, j \in U, e_{i,j} \in E\}.$$

Subnetworks can also be produced by other means from a set of nodes with edges sampled by some other procedure (see random walks in Section 2.1.6 and community detection in Section 2.2).

2.1.4 Paths, Walks, Diameters and Cycles

A set $P = (v_0, v_1, \ldots, v_{n-1})$ is called a *path* of size n from v_0 to v_{n-1} if for each $i \in [0, n)$, $e_{v_i, v_{i+1}} \in E$. In the special case where each edge in P is unique P is called a *trail*. A connected component of a graph $C \subset V$, is ta set of nodes $i, j \in C$ in which every i, j can be connected by a path from i to j. A graph with only one connected component is a *connected* graph (also termed *strongly connected* for a directed graph). If the graph formed by transforming any directed edges in the directed graph G to undirected edges ($e_{i,j} \rightarrow e_{ij}$) is connected, then G is *weakly connected* [59].

For a strongly connected graph G, the *length* of an edge is defined as the inverse of its weight, $l_{i,j} = W_{i,j}^{-1}$ for $i, j \in V$ such that $e_{i,j} \in E$. The minimum length of all paths starting at a and ending at b for arbitrary $a, b \in V$, $\ell_{i,j}$, is then

$$\ell_{i,j} = \min \left\{ \sum_{i=1}^{k-1} l_{i-1,i} : \ P = \{v_0, ..., v_{k-1}\} \text{ is a path} \right\}. \tag{2.3}$$

The *diameter* of G is

$$\text{diam}(G) = \max_{i,j \in V} \ell_{i,j}.$$

2.1.5 Trees and Directed Acyclic Graphs

We briefly define closely related terms related to trees and Directed Acyclic Graphs (DAGs). Consider a graph $G = (V, A)$:

- *Cycle/Acyclic:* A path C in G is a *cycle* of order k (k-cycle) if C is a *trail*, $T = (v_0, ..., v_k)$, of size $k + 1$ with every node unique except for $v_0 = v_k$. A graph G without cycles is *acyclic*.
- *Tree:* G is called a *tree* if G is undirected, connected and acyclic.
- *Directed Acyclic Graph:* G is a DAG if G is directed and acyclic. G need not be (weakly or strongly) connected.

Trees are important in neuronal models where the physical topology of the neuron can be approximated by a tree connecting the subcellular components (nodes) into connected segments [259]. On the other hand, DAGs are important to the understanding of ontologies in Chapter 3, where each node is a term, connected by semantic relationships [2].

It is often useful to define an order o on the node set V of a tree or DAG. The node at the bottom of this ordering is the *root node*, $r \in V$ with order $o(r) < o(v)\ \forall v \in V$. For example in neurons, r is often near the centre of the cell body with segments branching out from r [432]. In an ontology, the root node r is the most general term in the ontology to which all other terms can be traced back by backtracking the directed edges [2]. In a DAG or tree, a i with an edge $e_{i,j} \in E$ from i to j is known as a *parent* of j, while j is termed the *child* of i. In trees the parent of each node is unique (this is not necessarily true for DAGs) [425].

2.1.6 Random Walks

Given a graph G with vertex set V of size $|V| = N$, and adjacency matrix A, any process that sequentially generates a path starting from some node $s_0 \in V$ is termed a *walk*. Walks can be finite or infinite [59]. A random walk is a Markov chain $R = (V, P, \delta_{s_0})$. Here, δ_{s_0} is the probability vector that is 1 at s_0 and 0 elsewhere. A finite realisation from R, $s = (s_0, ..., s_{k-1})$,

of size k is built up sequentially by starting at s_0, adding nodes at random according to the transition dynamics determined by the transition matrix P.

In order to ensure that the sampling process stops for some finite $k \in \mathbb{N}$, we additionally require a stopping rule. We will only consider a stopping rule of the following kind:
The sampling of R will stop when either:

1. $|\{s_i\}_{i\in[0,k)}| = n \leq |V|$ for some $n \in \mathbb{N}$
2. the size of s, k, exceeds some large $N^* \gg n$.

This last rule is an exit to limit the computational time spent in sampling states from R when R is trapped in an isolated region of state space and unable to sample n unique states.

The simple degree-based random walk is the one that depends only on s_0, A and the stopping rule. In this case, the transition $P(A)$ matrix is

$$P(A)_{i,j} = \begin{cases} \frac{A_{i,j}}{\sum_{j\in V} d_i^{out}} & \text{if } d_i^{out} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Many other types of random walks exist including maximum entropy random walks and others [431]. These include random walks with restart encountered in Chapter 3.

2.1.7 Graph Laplacians and Mixing Times

For a simple, unweighted graph G, the combinatorial graph Laplacian matrix (or simply Laplacian) $L(G)$ is the discrete Laplace operator on graphs originally developed in [206]. Laplace operators are a family of operators linked with the processes of diffusion and equilibrium dynamics across many different domains and spaces [161]. The combinatorial Laplacian $L(G)$ is defined as

$$L(G) = D - A \tag{2.5}$$

where D is the diagonal matrix whose i^{th} diagonal entry is d_i. $L(G)$ is positive semi-definite since $L(G) = U^T U$ for

$$U_{e_{i,j},k} = \begin{cases} 1 & \text{if } k = i \\ -1 & \text{if } k = j \\ 0 & \text{otherwise} \end{cases}$$

for $e_{i,j} \in E$ and $, i, j, k \in V$. The matrix U is called the incidence matrix of L.

Consider a function $f : V \to \mathbb{R}$. The function values can be written as a vector with i^{th} entry f_i. Suppose that f represented the distribution of a diffusive quantity, such as heat, across the nodes of V, then for $e_{ij} \in V$, the instantaneous flow along an edge $e_{ij} \in E$ is proportional to $f_i - f_j$. The Laplacian acting on $L(G)f$ is a new vector function taking nodes $i \in V$ to $\mathbb{R}$, which is the sum of all diffusion experienced at $i \in V$

$$[L(G)f]_i = \sum_{j \in N(i)} (f_i - f_j) = f_i d_i - \sum_{j \in N(i)} f_j. \tag{2.6}$$

If $L(G)v = \mathbf{0}$ for a vector v then the diffusion experienced across G is zero at all $i \in V$. This occurs only when v is a zero eigenvector of $L(G)$ which can be shown to correspond to the vectors v_C where C is a connected components of G and

$$[v_C]_i = \begin{cases} 1 & \text{if } i \in C \\ 0 & \text{otherwise.} \end{cases}$$

This implies that there is a one-to-one relationship between the multiplicity of the zero eigenvalue of $L(G)$ and the number of connected components of G [59].

The Laplacian $L(G)$ relates to the *mixing time* of the simple random walk $R = (V, P(A), p_0)$ for $i \in V$ (see Eq. (2.4)). The mixing time is the time taken for the probability vector of state occupancy

$$p(t)_j = P(S_t = j | S_0 = i, R) = [\delta_i P(A)^t]_j,$$

to approach π_j, for $j \in V$. If we let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_N = 0$ be the magnitude of the N eigenvalues of $L(G)$ for G (all positive since positive semi-definite) with at least one odd cycle [59], then λ_2 directly bounds the mixing time of R as

$$||p(t)_j - \pi_j||_2 \leq |\lambda_2|^t \sqrt{\frac{d_j}{d_i}},$$

that is the mixing time depends on the degree of the starting and end nodes as well as the eigenvalues of $L(G)$ [225].

In Section 2.4.3 we show deeper relationships between diffusion processes on G and the Laplacian.

2.2 Community Detection and The Louvain Algorithm

Community detection is used to uncover nested modular structure in a network $G = (V, E, W)$ in applications across the biomedical sciences [237, 391]. A *community* is a set of nodes with relatively strong (or many) edges connecting to nodes of the same set and comparatively weak (or few) edges to or from nodes outside the set [135]. Community detection methods differ widely in their mathematical interpretations of what a community is (e.g. what does comparatively weak mean in the context of edges between nodes), as well as the mathematical approaches they employ to find them [299].

Divisive methods such as iteratively removing high importance edges to obtain distinct islands of activity were some of the first methods to be applied to PINs (see Section 2.4) [113]. Other methods include spectral decomposition of the adjacency matrix A of an undirected graph G to obtain a partition of the network into communities [109]. Yet another approach is to model the probability of a node belonging to one or more *blocks* within the network [7].

2.2.1 Modularity under the Configuration Null Model

One notion of high intra-connectivity at the edge level which has been widely adopted was formalised as the community modularity in [275]. *Modularity* was originally formulated for simple graphs as the tendency for specific subsets of nodes to have more internal edges than would be expected by chance. Key to the definition is that it requires a model of expected edges between nodes. Here, we address the modularity as it is defined for unweighted networks, with minor details in the generalisation weighted networks detailed in Supplementary Information,

Section D.0.2.

The modularity Q is an example of a quality function that scores the performance of a community partition of V, $\mathbf{C}$, given a probabilistic null model $p_{i,j}$ of the existence of an edge between $i, j \in V$. The community modularity score is

$$q(C) = \sum_{i,j \in C} A_{i,j} - p_{i,j},$$

for a community $C \in \mathbf{C}$ [272]. The modularity score for $Q(\mathbf{C})$ is then the sum over all $C \in \mathbf{C}$

$$Q(\mathbf{C}) = \frac{1}{2m} \sum_{C \in \mathbf{C}} q(C), \tag{2.7}$$

where $m = \frac{1}{2} \sum_{i,j \in V} A_{i,j}$

One of the most common null models is the configuration null model which is based on a simple relationship between degree and edge probability [135]. The intuition behind this model is that nodes that already have many neighbours tend to accrue more; leading to a high level of inhomogeneity in the degree distribution as observed in real world networks (see Section 2.5). Other possible null models exist that emphasize different shared attributes between community members, such as intracommunity density and intermember distances (see Section 2.4 for details on these properties) [86, 345]. We use the configuration null model here as a relatively simple model with well established properties shared by many real world networks, including some protein and brain networks [36, 220, 273] (see Section 2.5).

The configuration null model assumes that

$$p_{i,j} = \begin{cases} \frac{d_i d_j}{2m} & i \neq j \\ 0 & \text{otherwise.} \end{cases} \tag{2.8}$$

Directed Modularity

We focus on the directed, weighted modularity in Chapters 4 and 5 developed in [222]. The directed configuration null model, of which Eq. (2.8), is a special case is

$$p_{i,j} = \begin{cases} \frac{d_i^{out} d_j^{in}}{2m} & i \neq j \\ 0 & \text{otherwise} \end{cases} \tag{2.9}$$

where d_i^{out} and d_j^{in} are the directed degrees (see Eq. (2.1)) and m is defined as in the original modularity formulation in Eq. (2.7).

In Chapters 4 and 5 as an aid to calculate the modularity score of $\mathbf{C}$ for additional classes of networks we introduce the generalised symmetrised modularity matrix B

$$B_{i,j} = \frac{A_{i,j} + A_{j,i} - \gamma(p_{i,j} + p_{j,i})}{2}. \tag{2.10}$$

This matrix is symmetric regardless of the edges of G (undirected or otherwise) and represents the connectivity strength between i and j in V, subtracting the expected connectivity under the null model in Eq. (2.9) regardless of the community structure of $\mathbf{C}$. Note that B reduces to $B_{i,j} = A_{i,j} - \gamma p_{i,j}$ for simple graphs [299]. The resolution parameter $\gamma > 0$ is designed to make the size of detected communities an explicit and variable parameter with large $\gamma > 1$ favouring smaller communities than in Eq. (2.7) (which effectively assumes $\gamma = 1$) and $\gamma < 1$ corresponding to preference for large communities [136, 308].

Given a vector $s(C)$ such that $s(C)_i = 1$ if $i \in C$ and zero otherwise, the directed community modularity (generalising Eq. (2.7)) takes of a community $C \in \mathbf{C}$ as

$$q(C) = s(C)^T B s(C),$$

and so the modularity of the partition is

$$Q(\mathbf{C}) = \frac{1}{2m} \sum_{C \in \mathbf{C}} q(C) = \frac{1}{2m} \sum_{C \in \mathbf{C}} \sum_{i,j \in C} \frac{A_{i,j} + A_{j,i} - \gamma(p_{i,j} + p_{j,i})}{2}, \tag{2.11}$$

which reduces to the classical modularity in Eq. (2.7) in the special case where G is simple.

2.2.2 Optimising Modularity with the Louvain Algorithm

Given the modularity quality function Q in Eq. (2.11) it is tempting to simply attempt to find the partition $\mathbf{C}$ that maximises Q by exhaustively testing the possible partitions of V, however, the space of partitions of V into an arbitrary number of communities is prohibitively large [315], and not computationally scalable to even moderately sized networks [62]. This necessitated the development of a number of approximate algorithms. The Louvain algorithm is one such approximate algorithm which is known to scale better with network size than most of its competitors [29]. The Louvain algorithm employs a stochastic bottom-up approach in which each node is initially placed into its own community and nodes are assigned agglomeratively to neighbouring communities based on whether or not the re-assignment would increase the modularity sum of the network communities. Once an apparent maximum in total modularity is reached, a higher level network is constructed in which communities form the new nodes with edges weighted by the total weight of edges between nodes of the two different communities in the previous network. The agglomerative clustering step is then repeated for the new network and the whole process is iterated until convergence (i.e. when no modularity increases are possible through re-assignment). The algorithm also has strong analogues with random walk based sampling on networks which is favoured in our approach due to the methods of surrogate community data generation we employ in Chapters 3, 4 and 5 [115].

2.2.3 Limitations of The Louvain Algorithm

The Louvain algorithm is broadly applied across a wide range of fields [369]. However, problems relating to the limitations of the algorithm emerged relatively quickly. One of the earliest, which can be extended to a criticism of direct modularity maximisation approaches in general is the so-called resolution limit [136]. That is, for fixed resolution parameter γ, detected communities will depend not only on γ and the embedded community structure but on the size and density of G.

Another problem, specific to the Louvain algorithm has been put forward in [388]. This issue can lead to arbitrarily poorly connected communities under certain conditions. Namely, when a node assigned to one community by the Louvain procedure is acting as a bridge between two edge disjoint subnetworks of the community, a reassignment of the node to a second community can cause the original to become disconnected. This lead to the proposition of the Leiden algorithm

which addresses this shortcoming.

In Chapters 3 and 4, we apply the Louvain algorithm as it was originally implemented because of its wide usage in both neuroscience and molecular biology [236, 299, 369]. In later Chapter 5, we attempt to mitigate some of the above shortcomings by aggregating across multiple runs of the Louvain algorithm. This is known to help to average out poor random assignments. Resolution limitations can be similarly mitigated by aggregating community partitionings across multiple resolutions [226].

2.2.4 Multiresolution and Consensus Partitioning

Varying the resolution parameter between partitionings makes it possible to explore the multiple scales of community complexity. Various methods exist to select an appropriate resolution at which to perform partitioning. One possible method is simply to select the partition with a certain mean or median community size. Another is to minimise the so-called Variation of Information, VI which measures conditional entropy of the community occurrence [215, 258]. Given two partition $\mathbf{C} = \{C_k\}_{k=1}^{K}$ and $\mathbf{C}' = \{C'_{k'}\}_{k'=1}^{K}$ the probability $P(k)$ of community occurrence is the probability that a node chosen uniformly at random lies in C_k,

$$P(k) = \frac{n_k}{N},$$

where $n_k = |C_k|$. On the other hand the probability of community co-occurrence across $\mathbf{C}$ and $\mathbf{C}'$ is

$$P(k, k') = \frac{n_{k,k'}}{N},$$

where $n_{k,k'} = |C_k \cap C'_{k'}|$ for $C_k \in \mathbf{C}$ and $C'_{k'} \in \mathbf{C}'$, with $|V| = N$.

The entropy of $\mathbf{C}$ is

$$H(\mathbf{C}) = -\sum_{k=1}^{K} P(k) \log(P(k)),$$

while the entropy of $\mathbf{C}$ conditioned on $\mathbf{C}'$ is

$$H(\mathbf{C}|\mathbf{C}') = -\sum_{k=1}^{K}\sum_{k'=1}^{K'} P(k,k')\log\left(\frac{P(k,k')}{P(k)}\right),$$

The VI between partitionings $\mathbf{C}_1$ and $\mathbf{C}_2$ is

$$\mathrm{VI}(\mathbf{C},\mathbf{C}') = \frac{H(\mathbf{C}|\mathbf{C}') + H(\mathbf{C}|\mathbf{C}')}{\log(N)},$$

The goal is to find local minima of VI in the partition space of V by varying some parameter of the community detection procedure such as the resolution parameter γ [215]. The VI constitutes a distance metric on the set of partitions [258].

Similar to the desired goal of the VI, the Rand Index, is a measure of the agreement between two partitions that is adjusted for chance [180]. The Rand Index is limited in that its value may depend on the granularity of the partition and so is not best suited to compare across resolutions. The Adjusted Rand Index, ARI, is a version of the Rand Index that is adjusted for chance agreement.

$$\mathrm{ARI}(\mathbf{C}_1,\mathbf{C}_2) = \frac{\sum_{i\in\mathbb{N}_k}\sum_{j\in\mathbb{N}_{k'}}\binom{n_{i,j}}{2} - \left[\sum_{i\in\mathbb{N}_k}\binom{a_i}{2}\sum_{j\in\mathbb{N}_{k'}}\binom{b_j}{2}\right]\Big/\binom{N}{2}}{\frac{1}{2}\left[\sum_{i\in\mathbb{N}_k}\binom{a_i}{2} + \sum_{j\in\mathbb{N}_{k'}}\binom{b_j}{2}\right] - \left[\sum_{i\in\mathbb{N}_k}\binom{a_i}{2}\sum_{j\in\mathbb{N}_{k'}}\binom{b_j}{2}\right]\Big/\binom{N}{2}},$$

where $a_i = \sum_{j\in\mathbb{N}_k} n_{i,j}$, $b_j = \sum_{i\in\mathbb{N}_k} n_{i,j}$ and $\mathbb{N}_k = [1,k]$.

Both ARI and VI are used to attempt to find am optimal partition of V, for example across multiple resolutions. Alternatively, one can attempt to aggregate signal across multiple partitions or resolutions, say L separate partitions $\mathcal{C} = \{\mathbf{C}_1,...,\mathbf{C}_L\}$, for $N \in \mathbb{N}$, the consensus matrix D is given by

$$D_{i,j} = \begin{cases} \frac{1}{L}\sum_{m=1}^{L}\mathbf{I}_{(\exists C_k\in\mathbf{C}_m : i,j\in C_k)} & i \neq j \\ 0 & \text{otherwise} \end{cases},$$

where $\mathbf{I}_{\{A\}}$ is the indicator function that is 1 if A is true and 0 otherwise. The consensus matrix can itself be viewed as the adjacency matrix of a simple weighted network that can be inputted into any community detection procedure such as in Section 2.2.2 [216]. The consensus matrix is

important to the single layer and multiplex community detection procedures in Chapter 5.

2.3 Multiplex, Annotated and Dynamic Networks

When networks are permitted to change or have multiple types of relationships between nodes, it is often natural to consider such a network as having layers, where each layer represents a different time point or mode of activity. A multilayer network in which each layer consists of a copy of the same set of nodes but possibly different edge structure is termed a multiplex network [272].

Annotations on networks are additional sources of data outside of the network topology, about the nodes or edges in the network. Annotations are sometimes incorporated into a network model as a layer [230], but this need not be the case [236]. Protein annotations on PINs and functional activity levels in brain networks are both examples of network annotations [30, 276].

Dynamic networks are networks which model a dynamic process [213]. A dynamic network may be a potentially infinite time series of networks

$$\mathcal{G} = (G_1, G_2, ...).$$

We focus on a particular class of generative networks covered in Chapters 4 and 5. These are probabilistic, static generative models for generating such a network time series $\mathcal{G}$. While not dynamic networks in the sense above, we class such networks as dynamic networks as they encode a dynamic process. In Chapter 6 we consider spatiotemporal dynamics occurring in the Ca^{2+} signal of CA1 hippocampal neurons. This could be considered both a dynamic and annotated network in which we use multiplex networks to model activity at each time point in the unfolding Ca^{2+} dynamics.

2.3.1 Multiplex Networks

A multiplex network $G = (V^t, E, W)$ is a network with vertex set given by the Cartesian product

$$V^t = \underbrace{V \times V ... \times V}_{t \text{ times}}.$$

In a multiplex network, edges can be between or within layers. Intralayer edges are those within the same copy of V, these can be denoted $e_{i,j}^{k,k}$ for i and j in the k^{th} copy of V. Edges can also be between layers, or interlayer. An interlayer edge $e_{i,j}^{k,l}$ is an edge between i in the k^{th} copy of V and j in the l^{th} copy of V. In a dynamic multiplex network interlayer edges can represent the temporal relationships between layers [233, 270]. The studies of communities in dynamic networks is an active area of research [313]. We analyse the spatiotemporal community structure of such network models in Chapter 5.

2.3.2 Annotated Networks

An annotated network $G = (V, E, W, \mathcal{B}, \mathcal{A})$, is a network with additional edge properties $\mathcal{B}_{i,j}$ for $e_{i,j} \in E$ and node properties $\mathcal{A}_i$ for $i, j \in V$. Generally, these properties are secondary to edge weights (if they exist) and provide additional information on the properties of the network. Node properties may be real-valued such as a regional mean activity value (in the case of brain networks where nodes represent regions), a term or name, or a set of terms annotating a node in the case of proteins in PINs [30, 276]. Edge annotations may similarly be numerical or abstract. Annotations provide information over or above the structure of the network but are treated differently to layers of a network and indeed can be combined with network layers to produce a rich data structure. Layered annotated networks are of particular importance in Chapters 4 and 5 where neural activity can often vary by region while functional connectivity (edges) varies with time as well.

2.4 Network Statistics

There are a number of commonly used network statistics that summarise and contrast the topological properties of a network $G = (V, E, W)$ or its components [272]. These may be local properties of the individual nodes or edges, or global properties of the network as a whole.

2.4.1 Local Network Statistics

There are many statistics that represent the local topological properties of nodes and their local neighbourhoods for a graph $G = (V, E, W)$ with adjacency matrix A. Here we list a number of such node statistics. Of special interest to us are node centrality measures. Node centrality measures are local network statistics that indicate the importance of a node to the

structure of a network. This importance can be local or global topological importance and each measure emphasises different topological features as important. The choice of centrality measure therefore depends on the goals of the analysis. In Chapter 4, we employ node centrality measures to determine which nodes are central to the dynamics of a network. A node which scores highly for a particular centrality measure is sometimes termed a network *hub*.

Local Clustering Coefficient

One common centrality metric on simple, weighted graphs is the clustering coefficient, which measures the tendency for neighbours of a node $i \in V$, $j, k \in N(i)$ to have strong pairwise connections between each other as well as with i,

$$\mathrm{CC}(i) = \frac{\sum\limits_{j,k \in V} A_{i,j}, A_{j,k} A_{k,i}}{d_i(d_i - 1)}. \tag{2.12}$$

The value of $\mathrm{CC}(i) \in [0, 1]$, can be interpreted as a measure of the efficiency of local communication in the node neighbourhood, since higher clustering coefficient $\mathrm{C}(i)$ lessens the degree to which any one node may act as a bottleneck to communication between the three nodes. It is nearer to one when edge weights are high between the three and low when they are weak or absent [281]. The clustering coefficient is used in brain networks as a measure of local efficiency of functional connectivity [238].

Degree Centrality

The simplest and most direct of the node centrality measures is the degree centrality [272]. This measure is simply the average of the in and out degree of the node (or the degree for unweighted networks).

PageRank Centrality

A popular node centrality measure is the PageRank centrality (without dampening), named after one of its developers, Sergei Brin (of Google fame) [64]. The PageRank centrality weights a node as more central if it has higher probability of being selected at random from a well mixed random walk $R = (V, P, \delta_{s_0})$ on G. This centrality is best understood using the conversion of G into a Markov chain, M, as discussed in Section 2.1.1. The algorithm to calculate the PageRank centrality consider $R = (V, P(A), \delta_{v_0})$ at a node $v_0 \in V$, $\mathrm{PR}(v_0)$ and $P(A)$ as defined in Section

(2.1.6)). The $\mathrm{PR}(v_0)$ can be calculated by Singular Value Decomposition of $P(A)$ to obtain the left eigenvectors and scaling by the sum to produce unit eigenvectors (see Section 2.1.2, Property 3). The PageRank centrality is then,

$$\mathrm{PR}(v_0) = \pi_{v_0}.$$

There are also iterative algorithms to approximate $\mathrm{PR}(v_0)$, these methods simulate the mixing of information across G by an iterative procedure. This can be computationally more efficient and distributed than SVD which scales polynomially with the size of G [439].

2.4.2 Global Network Statistics

Global network statistics quantify general properties of the network.

Network Density

The *density* of the network, $\mathrm{D}(G)$, is defined as the proportion of edges in a network given the edges in a maximally connected network of the same size,

$$\mathrm{D}(G) = \frac{2|E|}{|V|(|V|-1)}.$$

Network Clustering Coefficients

The network clustering coefficients of G are two indicators of the tendency towards strong edges between neighbours of nodes. These are the local (nodal) clustering coefficient in Eq. (2.12). The Global Clustering (GC) coefficient measures the total strength of transitive connections in the network [272]. In contrast, the Average Local Clustering (ALC) coefficient measures the average of the nodal clustering coefficient $\mathrm{C}(i)$ over all nodes $i \in V$ as in Equation (2.12) [423]. These network clustering measures are given by

$$\mathrm{GC}(G) = \frac{\sum\limits_{i,j,k \in V} A_{i,j}, A_{j,k} A_{k,i}}{\sum\limits_{i \in V} d_i(d_i - 1)}; \tag{2.13}$$

$$\mathrm{ALC}(G) = \frac{1}{N} \sum_{i \in V} \mathrm{C}_i. \tag{2.14}$$

These measures differ in their interpretation and behaviour. ALC represents the empirical probability that the neighbours of any one node selected uniformly at random are also strongly connected. The GC represents the probability that any node triplet selected uniformly at random is strongly connected. This leads to a tendency towards inhomogeneously clustered graphs having high ALC but low GC. This has been demonstrated for a number of network models but we make this explicit in Section 2.5 with real world examples [123].

Average Shortest Path

Another important statistic in measuring the efficiency of communication within a network as a whole is the Average Shortest Path length (ASL). This measure quantifies average efficiency of connectivity between all nodes in the network.

$$\mathrm{ASL}(G) = \frac{1}{N(N-1)} \sum_{i,j \in V, i \neq j} \ell_{i,j}.$$

where $\ell_{i,j}$ is the shortest path length between $i, j \in V$ as defined in Eq. (2.3).

2.4.3 Heat Kernels

The heat kernel on a simple weighted network $G = (V, E, W)$ of size N, $\Psi(G, t)$ for time $t \geq 0$, connects random walks with the process of diffusion on graphs. Similar to the Laplacian (Eq. (2.5)), $\Psi(G, t)$ measures the diffusion over time of the quantities of the vector function $f : V \to \mathbb{R}$ (see Section 2.1.7). Just as the Laplacian is the Laplacian Operator on graphs, the heat kernel is the discrete, graphical analogue to the heat equation on the real line which models the diffusion of a heat source on a uniformly conducting wire [51, 126]. The heat kernel is

$$\Psi(G, t) = \exp(-tL(G)).$$

where $L(G)$ is the combinatorial graph Laplacian of G, see Equation (2.5). The heat kernel relates directly to Markov chains and in particular the familiar random walk on G, $R = (V, P, \delta_{s_0})$ in the following way

$$\Psi(G, t) = \sum_{k=0}^{\infty} \frac{t^k e^{-t}}{k!} P(A)^k,$$

where $P(A)$ is the random walk transition matrix defined in Eq. (2.4). This implies that the heat kernel is the expected mixing of the chain for a random walk, where each step occurs probabilistically at random with rate t. As the rate t increases, so does the number of steps and the amount of mixing. In the limit

$$\Psi(G,t)f \xrightarrow[\infty]{t} \bar{f}\mathbf{1},$$

where

$$\bar{f} = \frac{1}{N}\sum_{i\in V} f_i.$$

That is, the heat kernel applied to f approaches uniformity over V in the limit as diffusion continues for a long time.

In the special case where f is a probability vector, then $\Psi(G,t)f$ can be thought of as describing a family of distributions over V, parametrised by t. We will apply this interpretation when applying the heat kernel to understand error propagation on dynamic neuronal trees in Chapter 6.

2.5 Proposed Properties of Biological Networks

Real world networks arising from data in biology and neuroscience have many distinct properties, however, broad observations can be made about tendencies in behaviour of these networks. These characterisations will be used as a reference in later chapters where common approaches are applied across data from very different modalities and fields. We also provide common critiques of these proposed general properties and their applicability in specific contexts.

2.5.1 Small World Networks

In their original paper Watts and Strogatz demonstrated shared behaviour across a range of networks, including the neuronal network of the nematode *C. elegans* [423]. Since that time many other networks including brain, gene, protein and metabolic networks have been suggested to share the same property in both weighted and unweighted contexts [39, 196, 368, 415]. This property of *small world* network structure is characterised by low average shortest path and diameter, and high average local clustering coefficient. These networks are small in the sense

that communication between disparate areas of the network is relatively easy. In the biological context, this has been suggested to be due to metabolic and evolutionary demands on these networks [443].

Though it continues to be observed and reported [401], critiques of the generalisability of small world properties mostly stem from the limitations of the data to accurately reflect small world structure [288]. While we do not make direct reference to the small world property in our neural networks analysis in Chapters 4 and 5, we do assume some level of local clustering to be occurring at both temporal and spatial scales in order to apply our community detection procedures.

2.5.2 Scale-free and Fat-tailed Distributions

So-called scale-free behaviour has long been popularised in the biological networks literature starting with the works of Barabási, Albert and others [9, 34, 35, 152], including publications in both the neuroscience (across conscious conditions) and bioinformatics (proteomics and genomics) literature [33, 221]. Network statistics, most notably the degree of many real world networks have been shown to have higher density in the empirical distribution than would be expected under many common distributions (such as the normal or Poisson distributions) at the extreme ends [33]. Scale-free activity is characterised by power law-like behaviour of the statistic. Strictly scale-free behaviour for a network statistic k, with distribution $P(k)$ is

$$P(k) \propto k^{-\alpha}$$

for some constant α, typically $\alpha \in [1, 4]$.

In recent years, strictly scale-free behaviour has been difficult to find in many large real world data sets [402]. This trend highlights a problem with stringent assumptions on the distribution of statistics like the degree. However, model fitting suggests that alternative models remain similarly or more implausible than the scale free model [66]. This could suggest a laxer interpretation of the scale-free property is needed. We refer to scale-free behaviour in Chapter 3, however, we interpret scale-free broadly to mean that these statistics follow a generally fat-tailed distribution (preferring this term where suitable) with a relatively high number of extreme values (more so than would be expected under a normal or Poisson distribution).

3 | CommFinedWalker

Controlling for Inspection Bias in Protein Networks

Biases in the detection of protein properties and protein-protein interactions arise naturally as an artefact of what is historically possible and interesting to experimenters. These biases are reflected in the structure of Protein Interaction Networks (PINs), through factors including the proteins that experimenters choose to screen for interactions. This selectivity leads to so-called 'inspection bias' in that further research often only occurs in proteins about which a fair amount is already known [149]. These same study selection biases affect which proteins are annotated with functional data and to what level of detail [335], which creates a general pattern across the network characterised by islands of highly annotated, well studied proteins interspersed within the majority of less annotated proteins about which very little is yet known. The large scale interaction and biological information biases which lead to informational inhomogeneity in protein networks hinders the discovery of interesting proteins and protein communities outside of those that have already been comprehensively studied [236].

Community detection has played an important role in finding meaningful and manageable protein subsets of large PINs [6,85,377]. Similarity of protein function has been used to refine these subsets to those proteins most likely to be interesting for future experimentation [198]. The functional annotations themselves are often sourced from the Gene Ontology (GO) database [26, 127, 377]. However, GO annotation of proteins is highly inhomogeneous across PINs, with clusters of highly annotated proteins interspersed with poorly annotated regions [236]. The exact scale of the information asymmetry regarding proteins is unknown as the 'completeness' of a proteins functional profile in GO cannot be easily assessed.

Evidence suggests that proteins within communities in highly annotated regions of the network are disproportionately selected for by classical approaches to community filtering [236]. However, highly annotated does not always mean most interesting to experimenters who are often moti-

vated by a community's potential for further novel research for instance as a novel drug target. Thus interest in a protein set should be relative to the amount that community member proteins have already been researched and annotated. CommWalker (CW) is a recent method designed to take relative differences in local annotation into account during community selection. CW does this by comparing communities to the local network background of protein annotations in nearby PIN regions [236]. This is in contrast to classical Functional Homogeneity (FH) filtering which selects communities based on comparison to a global functional similarity level, favouring globally well-annotated communities without adjusting for long range dependence in annotation quality or connectivity.

In order to sample the local annotation background of a given community found using a community detection method, CW explores the network by random walk, starting from within a community and moving from randomly chosen neighbour to neighbour to form a sample from that community's local background. CW then computes the FH of the sample. CW compares the FH of the proteins it found with those of the actual community, filtering out communities that are less functionally similar their local background. This method works well on a regular lattice network, however, a problem with this approach can be identified when PIN topology is considered. PINs are well known to be very easy to traverse, with all proteins separated from each other by only a few steps in the network on average (common in real world networks, see Section 2.5.1). Naïve PIN exploration via random walks could therefore lead to comparison with non-local regions of the network outside the original context [415]. This potential problem with the CW algorithm has been highlighted in previous work and may be exacerbated by the higher network connectivity of well studied proteins [236, 414].

Here, we present the CommFinedWalker (CFW) method, which restricts network community exploration to be more local. The method was tested on two previously studied PINs of the human proteome: BIOGRID-AP (including 15405 proteins with 49301 interactions), from a database including various species specific PINs [83], and HINT-P (including 10927 proteins with 165343 interactions), a human only database [100]. Here, results were validated against a co-expression based community quality measure [399]. CFW succeeds in filtering for proportionally more, potentially interesting communities (having a high level of co-expression) with poorly annotated neighbourhoods. Despite mostly agreeing on which communities to include or exclude, the communities selected only by CFW also tend to be more sparse than their CW

counterparts suggesting that CFW selects for novel features not seen in the original CW.

As important contextual information, this chapter discusses the background behind PIN community analysis as well as CW, the previous method for dealing with inspection bias (Section 3.1). This is followed by an overview of the data used in network construction and testing of the new CFW method (Section 3.2). Section 3.3 starts with an overview of the communities detected, followed by detailing the differences between CW and CFW in their exploration of the local background and in the communities they detect. Comments on overall performance of the new CFW method including on its tendency to select sparser, more pathway-like communities are presented in Section 3.4.

3.1 Background

In this section PIN source data and construction is explained. Source data biases are used to justify the previous CW framework. Further caveats arising from the topological properties of networks (i.e. the small world property) are used to motivate the need for a new restricted exploration approach based on CW.

3.1.1 Structure and Bias in PIN Source Data

Some of the earliest PINs were constructed using data from the Yeast 2-Hybrid (Y2H) assay [188,340]. Y2H is one of two commonly used screening methods for detecting possible interactions between large numbers of proteins, the other method being Tandem Affinity Purification with Mass Spectrometry or TAP-MS [301]. Although many PIN databases include interactions from smaller-scale, targeted studies [100], these two high-throughput methods still make up the bulk of PIN construction data [149]. This makes understanding the errors these two high-throughput screens are prone to important in understanding inspection bias.

While true interaction discovery rate has been estimated to be as high as 70% or higher [343], Y2H and TAP-MS both suffer from high false positive rates in detecting interactions. The false positive rates have been estimated for Y2H to be between 35% and 70% of total interactions, and near 35% for TAP-MS [14, 166].

Experimentally, both methods utilize a library of possible interactors augmented with tags which are simultaneously expressed in the host (usually yeast) providing an opportunity for interaction.

Interaction detection differs between the two methods as in Y2H binding takes place in the yeast nucleus, simultaneously activating a transcription factor signal. In the case of TAP-MS potential interactions occur in the cell and later undergo purification and binding partner identification via mass spectometry [50]. These are foreign environments and modes of expression for human proteins that allow for interactions that would never have occurred under normal patterns of temporo-spatial expression in vivo. In addition, proteins must be modified by fusion of a reporter or tag component (that allows interactors to be detected). The unnatural localisation conditions in these experiments and protein modifications necessary for detection can prevent biologically relevant interactions from being detected (false negatives) and allow for aberrant interactions (false positives) [166].

In terms of interaction complexity, Y2H studies are more limited and can detect only direct physical (P-type) interactions. On the other hand TAP-MS utilizes a pull-down approach in which entire complexes of proteins can be captured by a single tagged bait protein [147]. This implies that while not all proteins interact with the bait directly (P-type), the relationship may still be biological and can be recorded as evidence for an associative (A-type) interaction between co-complexed proteins. Inclusion of A-type interactions opens up new possibilities for false positives since unnatural co-complexing may make additional aberrant interactions possible [163].

Differences also exist in binding affinity between the two methods. In Y2H studies, even weak binding can trigger transient activity from the reporter module, contributing to the false positive rate [342]. This leads to further discrepancies between networks constructed using different types of experiments [55].

The errors in both methods are compounded when combining the multiple assays used in PIN construction. Repeated tests involving the same proteins within or between assays increases the chance of including false positive interactions particularly for high-interest proteins that are repeatedly screened [403]. Screening bias both causes and perpetuates structural inspection bias by giving repeatedly screened proteins the appearance of having many interaction partners.

Mistakes in database curation can also lead to forms of error, for example through protein name or species confusion [99]. Some databases also make use of complementary data types such as co-expression and protein function data to validate interaction quality, which may help to

alleviate some of the method specific bias in high-throughput interaction screens [378].

3.1.2 Network Analysis

PINs have several key research applications from protein function identification to visualisation of regulatory pathways [347]. Considered as graphs, PINs are composed of nodes (the proteins) and links (the interactions). PINs tend to be sparse, having few interactions per protein with subsets of highly interacting proteins [437]. PINs also tend to be easy to traverse, meaning that average distances between proteins in the same connected component tend to be small (as low as 2 in many networks [57]), even for PINs containing many thousands of proteins [415]. This traversibility phenomenon is true of many real world networks and is known as the small world property (see Section 2.5.1) [365].

The Graph Representation of PINs

PINs are constructed from a combination of evidence from multiple assays (See 3.1.1) and are often thresholded for high confidence interactions with large amounts of evidence. In this way PINs can either be represented by a weighted graph or simple graph, G. In the latter case G represents only the most high confidence protein interactions. The graph (or network) G is composed of a set of node indices (numerical protein labels in a PIN), $V(G)$, and edges (interactions), $E(G)$. A network's edge structure can be isomorphically represented by the symmetric matrix A (see Section 2.1.1). Here, we use historical PIN data that has already been processed to produce high confidence, simple networks [236].

Any subset $U \subset V(G)$ induces a sub-network, H, where $V(H) = U$ and $E(H)$ contains all the edges between members of U that are present in $E(G)$;

$$E(H) = \{e_{ij} : i, j \in U, e_{ij} \in E(G)\}.$$

Communities can be understood as induced sub-networks of G formed of proteins from $V(G)$.

3.1.3 Selecting Out Proteins with Metadata

There are many types of metadata associated with proteins, from annotation terms relating to possible biological functions to the protein's spatio-temporal expression patterns. Not all proteins are equally well represented and some may not be associated with any metadata at all.

We define a general selection function $\pi_\bullet[\star]$ for all types of metadata relating to PINs. This map takes two arguments, a set of proteins $\star \subset V(G)$ and a set of metadata on proteins $\bullet$ (such as a set of annotation terms). The function returns the subset of proteins in $\star$ associated with $\bullet$ metadata ($\pi : 2^{V(G)} \times 2^{\text{metadata}} \to 2^{V(G)}$). The purpose of π is to select proteins associated with $\bullet$ metadata, for enrichment, semantic or co-expression analysis.

3.1.4 GO: The Information Content of Protein Function

In the context of protein function, ontologies are vocabularies formed of a strictly controlled and curated set of terms $\mathscr{A}$ which act as a source of functional protein metadata. These terms, known as functional annotations, are used to describe the molecular, mechanistic and/or subcellular role ascribed to individual proteins. These termed form a DAG with a term as root node (see Section 2.1.5). GO is one of the most widely used ontologies for describing protein function [26]. The ontology covers three distinct domains which are organised by the root terms: cellular component (CC), molecular function (MF) and biological process (BP). BP is the domain most closely related to the mechanistic and thus functional role of proteins. We focus on BP for comparison of CFW and CW results presented in [236], BP is also favoured because of its widespread use in the literature when identifying a possible functional module (a set of proteins with a specific functional role) [237, 391].

Hierarchical Structure of Semantic Relationships in GO

In each of the three GO domains, child terms are linked to more general terms, their parents, by simple relationships such as constituent (part of) or subtype (is a) relations. This definition implies a means to construct an *ancestor* set $\mathscr{A}_(B)$ for a given set of terms, B. This set contains all parental and increasingly general ancestral terms of the given set and is the result of considering the union of all paths from the set to the root term. This process is known as performing an ancestor call on the set, $B \subset \mathscr{A}_(B) \subset \mathscr{A}$.

Proteins can be annotated with terms from any specificity level on the GO hierarchy and so comparison of term ancestry provides a meaningful way to compare the function of proteins [286,293]. Two main approaches exist for comparing functional annotation within communities or between proteins: semantic similarity measures and gene set enrichment analysis (see Appendix B.2). Semantic similarity measures form the basis of community to background annotation compar-

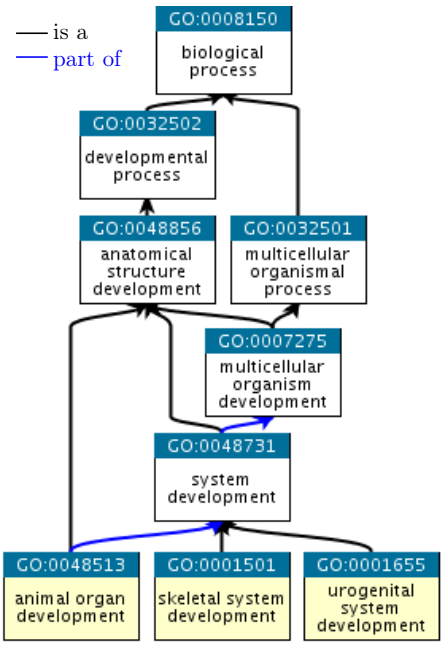


Figure 3.1: Subgraph of GO's BP domain generated using the three bottom most terms (yellow) as the input set, B, of an ancestor call $\mathscr{A}(B)$. Terms are shown as descending from the root term (biological process) with their GO ID above. Terms are linked by directed edges with one of the two relations *is a* (black line) or *part of* (blue line) with arrows showing the direction of the relationship. For instance animal organ development is both a *part of* system development and *is a* type of anatomical structure development. This graph was generated using QuickGO [53].

ison in the CW framework. We next address the semantic similarity method used by CW in more detail.

3.1.5 Scoring Functional Similarity: Pandey Measure

Semantic similarity measures are information theoretic indices that reflect the information content of the terms shared between proteins [293]. Averaging the semantic similarity of interacting proteins within a community can provide an indicator of the amount of shared biological information. Average semantic similarity is therefore a proxy for the known similarity in function (functional homogeneity) of proteins within the community, where high average similarity within a community suggests a possible functional module [90].

Naïve definitions of "informative" protein term relationships often fall short by only considering the abundance of terms shared between proteins or their relative scarcity in the protein universe (as defined by the PIN). Crucially, these approaches ignore the relative specificity of terms as reflected in the ontology hierarchy and therefore tend to be biased due to correlations between

related terms [12].

The Pandey measure is a semantic measure between two proteins that takes topology into account by considering the proportion of annotated proteins that share the same ancestor set [286]. The Pandey semantic similarity between proteins i and j is calculated by considering the number of proteins annotated by the intersection of the ancestor terms annotated to i ($\mathscr{A}_i$) and j ($\mathscr{A}_j$), $\pi_{\mathscr{A}_i \cap \mathscr{A}_j}[V(G)]$, normalised by the set of proteins $\pi_{\mathscr{A}}[V(G)]$ in the PIN,

$$s_{ij} = -\log_2 \left(\frac{|\pi_{\mathscr{A}_i \cap \mathscr{A}_j}[V(G)]|}{|\pi_{\mathscr{A}}[V(G)]|} \right).$$

The Pandey measure can be used to construct an aggregate measure of the FH of any PIN subnetwork or community C_k, through the average of protein semantic similarity. This mean is taken over the set of all annotated pairs of proteins in a subnetwork using the natural extension of π to the set of annotated pairs within a community C_k,

$$\pi_{\mathscr{A}}[V(C_k)]^2 = \{\{i,j\} : \{i,j\} \in \pi_{\mathscr{A}}[V(C_k)], i \neq j\}. \tag{3.1}$$

The resulting community aggregate FH measure is

$$F_{C_k} = \sum_{\{i,j\} \in \pi_{\mathscr{A}}[V(C_k)]^2} \frac{s_{ij}}{|\pi_{\mathscr{A}}[V(C_k)]^2|}. \tag{3.2}$$

The community FH score allows higher scoring communities with similar annotations between (annotated) proteins, to be distinguished from lower scoring communities with more common and/or less related annotations.

3.1.6 Using Co-expression to Validate Functional Modules

In order for co-expression data to be a plausible source of validation data for annotation-based methods, the scoring methods should be comparable and the co-expression data should provide additional information that informs the presence of functional modules in PINs. Notably, while there should also be some commensurability between co-expression and annotation data so as not to unfairly penalise the method. The measures should also not be identical as to avoid over-estimation of performance. We therefore briefly demonstrate that co-expression is indeed

a suitable source of validation data for FH-based functional module detection methods with a few notable caveats.

Co-expression measures and in particular, co-expression networks have been shown to contain both overlapping and complementary information to PINs. For instance, it was found that in a multiplex network containing both interaction and co-expression layers, communities in this multiplex network tended to be more functionally enriched than an aggregation of communities from both single layer networks [107]. In addition, the multiplex network communities were shown to be more robust to missing data than their single layer counterparts, strengthening the hypothesis that the data types are complementary. Studies show a high degree of correlation between co-expression of proteins in a co-expression network and the presence (absence) of an interaction between them in a PIN [158]. However, such co-expression-based networks are topologically distinct and are hypothesized to reflect regulatory more so than physical interactions between proteins [434]. Lastly, research into the applicability of gene co-expression to predict protein function generally show only a weak signal when not paired with other forms of data such as co-evolutionary information [399]. Given this evidence, we contend that co-expression scores do therefore provide a meaningful set of complementary data for comparison with term-based annotations on PINs.

Co-expression analysis relies on correlations in abundance between protein precursor messenger RNAs (mRNAs). These mRNAs are the necessary intermediary between genetic code and protein expression. In order to be expressed in a cell or tissue proteins must first be translated from mRNA. Co-expression analysis often uses expression data sourced under different spatiotemporal profiles or biological conditions in order to cover the whole interaction space [143,148,413]. High levels of co-expression within communities has been used as a predictor of community functional conservation [399].

RNA-seq has made large-scale co-expression analysis possible by providing a high-throughput method based on next generation sequencing of mRNA [420]. RNA-seq provides counts of protein precursor mRNAs that indicate the protein's expression level, but these counts must be normalised to account for experimental count variation. The relationships between mRNA counts for different proteins suggest direct regulatory and indirect functional relationships, making co-expression analysis a useful surrogate for functional annotation analysis on communities [191].

Scoring Co-expression within Protein Communities

Here, the data we used in our analysis is based on RNA-seq experiments so we assume the expression metadata $\mathscr{E}$ to contain normalised mRNA counts of protein i, across multiple time points, denoted by the $x_i \in \mathscr{E}$, where the number of experiments and time points (entries in x_i) may vary by protein.

Average co-expression in a community C_k, E_{C_k}, has previously been used as a measure of functional homogeneity within the community, as follows [236]. The mean of the absolute inter-protein co-expression correlation $|\rho_{ij}|$, between shared entries of the vectors x_i and x_j (or if values are missing, the vectors of their mutually occurring entries), is used to score a possible functional relationship. However, since the co-expression data includes indirect correlations between non-interacting proteins, it is necessary to refine the co-expressional enrichment score to be defined only for proteins with an interaction present. This makes it analogous to the FH score and in addition this increases ease of comparison by making both measures consistent with an underlying PIN topology. The magnitude of the correlation is used because both strong negative and positive correlation can plausibly inform protein interactions (for instance through direct inhibition or up-regulation). The community co-expression score, E_{C_k}, is then given by the average of pairwise co-expression scores, $|\rho_{ij}|$, of community members i and j for which expression data exists (similar to Eq. (3.2)),

$$E_{C_k} = \sum_{e_{ij} \in \pi_{\mathscr{E}}[V(C_k)]^2} \frac{|\rho_{ij}|}{|\pi_{\mathscr{E}}[V(C_k)]^2|},$$

where $\pi_{\mathscr{E}}[V(C_k)]^2$ (see Eq. (3.1)) represents the pairs of proteins in C_k for which expression data from $\mathscr{E}$ is available for both proteins. Although E_{C_k} is not used directly to determine whether a community is a functional module, it provides a secondary measure of the overall quality of potential modules selected by a FH-based method.

3.1.7 CommWalker: Conditioning on the Local Annotation State

CW provides a measure of how likely a community is to be a functional module. It does so while also attempting to take local inspection bias into account [236]. Like many other methods that define community quality by functional modularity, CW makes use of FH measures. Inspection bias is a problem for FH-based functional module detection because annotation level, a result

of research history, directly informs FH filtering of communities, which should reflect research agreement rather than research abundance. Figure 3.2 shows the relationship between local availability of annotations and apparent local FH. In this section the effect of inspection bias on local FH scoring is reviewed. The random walk-based methodology behind these results is then used to justify the CW framework.

Random Walks Can Provide a Measure of Inspection Bias

Researchers in [236] used short random walks to measure the amount of inspection bias in local FH due to annotation by calculating a Local Homogeneity Measure (LHM) for each protein in the network and comparing it to the local level of annotation amongst proteins sampled. The assumption behind this is that short random walks will tend to stay local and thus, on average provide a rough indication of the FH background. In Section 3.1.9 we examine this assumption in more detail and provide alternatives.

The Protein LHM is calculated as

$$LHM_i = \sum_{m=1}^{10^4} \frac{F_{L_{im}}}{10^4}, \tag{3.3}$$

where L_{im} is the m^{th} random walk of length 6 starting at i and 10^4 was chosen to provide a stable estimate of the local homogeneity. Each random walk induces a small (and thus locally restricted), community-like sub-network of size at most 6 around the protein i. This motivates the borrowing of community FH notation from Eq. (3.2). The fraction of locally annotated proteins can be measured similarly by substituting $F_{L_{im}}$ for the fraction of annotated proteins in the L_{im} walk.

Previous results (see Figure 3.2) suggest that a high number of locally annotated proteins is strongly correlated with high protein LHM [236]. This demonstrates the annotation dependence that needs to be addressed and shows that random walks can provide a local measure through which the local annotation inhomegeneity can be measured and taken into account.

The Critique of Classical Approaches to Functional Module Selection

The direct approach to assessing community functional conservation is to compare the community mean FH (see Eq. (3.2)) to some globally applied threshold derived from the network

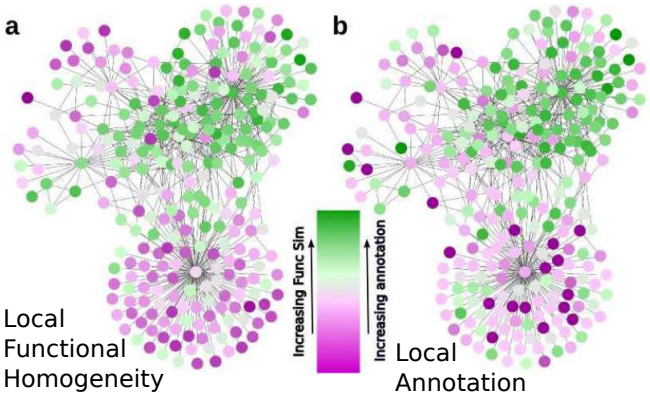


Figure 3.2: This Figure, adapted from the original CW paper [236], shows the clear similarity between (a.) local homogeneity score (measured by LHM) and (b.) local annotation fraction, in the same PIN subnetwork of HINT-P. The local functional homogeneity is measured by random walk as in Eq. (3.3), giving a score for each node based on the average community FH of 10^4 short walks that originate from it. Similarly, local annotation fractions per node were calculated from the average of walks originating at that node. This fraction was a simple average of nodes sampled in each walk rather, ignoring local connectivity. Levels of local annotation or functional homogeneity for each protein (node) in the example subnetwork (a subnetwork induced by the three step neighbourhood of the protein FAT1) are shown on the colour scale from purple (low) to green (high). Regions of low homogeneity have a clear tendency to coincide with regions of low annotation fraction and vice versa.

(such as the median FH of interacting protein pairs). But under this method, communities from well-annotated regions will tend to have higher FH and thus to be over selected. CW was developed to filter for communities that contain proteins that are functionally similar enough to be interesting given the amount of local annotation available.

CommWalker Compares Communities to Locally Defined Pseudocommunities

In order to condition on the state of local annotation surrounding each community C_k, CW uses, as a proxy, sets of proteins sampled through random walks in the vicinity of the community. Each random walk starts from a protein in the community and contains the same number of proteins as the community. Each protein set sampled in this way near a particular community C_k, induces a subgraph C_{km}. This subgraph resembles an equally sized community sampled in the vicinity of C_k, and is termed a pseudocommunity of C_k. We use C_{km} to refer to the m^{th} pseudocommunity sampled in this way. Since each C_{km} was sampled by random walk, results suggests that pseudocommunities may share many local topological properties with the C_k community found by modularity maximisation [106]. This makes them particularly apt for functional comparison to the original community.

Apart from having the same size as the original community, each pseudocommunity generated

also overlaps with the original community by starting within it. Pseudocommunities are therefore more likely to contain proteins more similar to the original community in terms of annotation and FH level, than would be expected by chance for a random sample of proteins of the same size [236]. The distribution of pseudocommunities according to their FH is an approximation of the original communities local annotation background. Any community with higher mean community FH than the vast majority of its pseudocommunities can be considered highly functionally homogeneous (i.e. enriched) given this annotation background.

CW relies on extending community FH as defined by Eq. (3.2) to each pseudocommunity (with the FH of C_{km} denoted $F_{C_{km}}$) and then comparing these values to the original F_{C_k}. In short, the fraction of pseudocommunities C_{km} with FH exceeding the original C_k is recorded as a tail statistic or T-value, T_{C_k},

$$T_{C_k} = \frac{1 + \sum_{m=1}^{M_{C_k}} \mathbf{I}_{[F_{C_k} < F_{C_{km}}]}}{1 + M_{C_k}}, \tag{3.4}$$

where M_{C_k} is the total number of walks for community C_k which depends on the size of C_k (see below). Under Eq. (3.4) locally unexceptional communities produce T-values close to 1, that could then be filtered out when exceeding a task-specific threshold, T^*, and kept when $T_{C_k} \leq T^*$.

The T-value can be interpreted as an indicator of community functional modularity or quality, with lower T-value corresponding to higher quality, more significantly locally exceptional communities. With regards to controlling for inspection bias, CW has been shown to select for communities with more proteins with lower local homogeneity (determined by LHM) than classical mean FH filtering. Thus CW potentially mitigates the effects of inspection bias [236].

Selecting the Number of Walks

In order for sampling to avoid size biases, the total number of walks per node was chosen so that the total number of nodes explored per community, C_k, remained constant. This number is $10^5/|V(C_k)|^2$, where the constant 10^5 was found to be sufficient to produce numerically stable T-values. This means that the total walks per community, M_{C_k}, will be equal to $10^5/|V(C_k)|$,

if all walks for each node member are completed successfully.

A walk can only be a complete pseudocommunity if it is the same length as the original community size $|V(C_k)|$. Walks may be more difficult to complete when starting from regions which are poorly connected to the rest of the network, such as degree one nodes. In the case of CW, a walk was discarded if it exceeded $|V(C_k)|^2$ steps without reaching $|V(C_k)|$ distinct nodes. It is thus also necessary to place a limitation on the maximum number of discarded walks which scales with size, since more walks means more possible failures. In the case of CW this limit is $2 * 10^5/|V(C_k)|$.

3.1.8 Motivation for Alternative Sampling Methods

CW assumes that any community-sized random walk still samples proteins with similar levels of annotation as the community (is from the same local annotation region). As noted in Section 3.1.2, average shortest path lengths within PINs can be low, as low as three to five [415] (a property of many small world networks, see Section 2.5.1). In addition, mixing times on protein networks have been observed to be short when compared to other real world networks (such as many social networks [303]). Work on the BIOGRID human PIN shows that relatively low average shortest path is conserved even for nodes sampled uniformly at random from the network [46]. Together, these properties suggest that pseudocommunities sampled in an unconstrained way, such as by unconstrained random walk could be sampling nodes from across the network. Lastly, functional similarity is suggested to decrease with shortest path length [383]. As a result, nodes sampled by random walk via CW may possess an entirely different annotation distribution to the original community whose local annotation background the pseudocommunity is supposed to represent.

The problem is exacerbated when highly screened proteins are considered. Highly screened proteins have disproportionately high degree (see Section 3.1.1) and are also likely to be well-studied. This implies that these proteins will tend to be part of communities with high levels of annotation. The result is that these highly annotated communities are more likely to be 'escaped' by their random walks. The resulting unrepresentative pseudocommunities make for unfair comparisons between the relatively poorly annotated network proteins present in the pseudocommunities and the highly annotated community proteins. Since highly annotated communities are likely to have high FH, these communities will tend to T-values that do not reflect

relative enrichment and instead are too low (too apparently significant).

The effects of unconstrained walk sampling can be summarised as having an averaging effect on pseudocommunity FH. Aggregations of non-local FH scores can lead to pseudocommunity FH scores which are skewed towards the global FH distribution of the network rather than reflecting the local state of relevant research near each community. The FH averaging of pseudocommunities may cause higher mean FH communities in annotationally richer regions to be under-penalised and low FH communities from poor regions to be over-penalised. This potential community test bias makes constraining the sampling region a priority in the optimisation of CW and the primary goal of the CFW method.

3.1.9 Alternative Psuedocommunity Sampling Methods

In order to accurately compare the FH or co-expression-based scores of communities to that of the local background, representative samples from the local background should be as close in topological and annotational distribution to the original community as possible while maximising the coverage of this background space. Here, we examine the original assumption that generic random walk sampling accomplishes this best when compared to other methods and give a brief review of alternative methods for use in our new method. We prioritise locality as a means of retaining similarity to the original community but acknowledge other methods may exist to sample areas of the network with similar local properties (see discussion in Section 3.4). Properties such as connectedness can be guaranteed by all of the sampling methods which we examine here. Other properties, like similarity of annotation quality and topological properties of the sample, may depend on the method.

Table 3.1 shows a number of possible techniques for sampling contiguous nodes in a network (nodes that form a connected induced subnetwork). All of snowball, breadth and depth first sampling share the weakness of being non-stochastic and therefore unable to explore the full range of possible communities in the network. These methods also produce samples that are necessarily topologically biased towards either long, thin induced subgraphs (depth) or star-like subgraphs (breadth).

Of the stochastic methods, expansion sampling favours the addition of high degree nodes that quickly leads to non-local sampling [242]. Forest fire and (generic) random walk sampling both

have high topological fidelity [438]. These methods have both been shown to preserve important topological properties such as the clustering coefficient (see Section 2.4) and other topological features [223, 438]. Forest fire sampling can allow for interpolation between preserving local and global properties by tuning of parameters, however local properties are arguably more computationally expensive since this requires a lower probability and thus a higher number of iterations [116]. The recursive forest fire sampling method is comparatively more expensive than random walk sampling which is sub-linear in sample size k [27]. More recent, PageRank forest fire sampling attempts to solve the inefficiencies and improve the sampling of less dense communities [385].

Metropolis-Hastings random walks are designed to approximate mean global node features (such as annotation level) from a small subset of nodes and requires a computationally expensive rejection step [218]. This is the opposite of what is desirable for our purposes where the goal is to approximate the local background with high coverage, not necessarily obtain a tight global estimator at high cost. The opposite problem is true for maximum entropy random walks which tend to be too locally concentrated around the most dense regions of the network with minimal coverage [72]. Stochastic restart random walk on the other hand is good at remaining local to the starting node when correctly parametrised and was first proposed for local node importance ranking [386]. Notably, it has also shown promise in predicting disease associations from a multiplex PIN, annotation and co-expression network [396]. It is therefore a promising candidate sampling method for local background sampling with the caveat that the appropriate restart probability can dependent on local network topology.

Another promising method is community Markov Chain Monte Carlo. The method samples directly from the distribution of possible communities by exchanging nodes with neighbouring communities, taking as starting state the original partition. The method tends to sample plausible communities (according to a predefined objective function). It is good at remaining local but generating sufficiently different samples is not computationally scalable to large PINs [241]. Similarly complex is node2vec sampling (based on the feature learning algorithm in [159]), which can theoretically select both local and community-like samples but requires tuning of additional parameters.

What is apparent from all of these methods is that there is a trade off between community-like

local sampling and coverage as well as concerns over parametric and computational complexity. PINs are large networks with no gold standard for validation. Sampling methods thus need to be scalable and generalisable to all regions of the network without recourse to too much fine tuning of parameters. We propose a slight modification to stochastic restart random walk as our sampling method that directly restricts sampling to a given snowball sample (neighbourhood) of the starting node. In so doing we hope to cover much of the local space while still capturing community-like samples as our psuedocommunities.

Sampling Procedures	Basic Description	local	community
Depth First	From starting node samples according to depth first ordering [108]		
Breadth First	From starting node samples according to breadth first ordering [108]	X	
Snowball	n-neighbourhood centred at a node [154]		
Expansion	Greedily adds neighbours that increase the neighbourhood size the most [241]		
Generic FF	From starting node recursively add neighbours with probability p [224]	X	X
FF with PageRank	Uses PageRank to prioritise low ranking local hubs [385]	X	X
Community MCMC	Accept a community sampled from an MC over community space with probability p [241]	X	X
Generic RW	Starts at a node and uniformly samples successive neighbours [228, 291]		X
Maximum Entropy RW	All paths of same length are equiprobable given start and end [72]	X	X
Metropolis-Hastings RW	Weights transition probabilities to give uniform stationary distribution [168, 374]	X	X
Stochastic Restart RW	Each step in the walk has probability p of returning to starting node [386]	X	X
node2vec RW	Interpolation between a depth or breadth first RW depending on p and q [159]	X	X

Table 3.1: This table lists procedures for sampling contiguous nodes in a network with citations providing further details on each method. Each row gives a basic description of the method and a summary assessment of whether the it tends to select community-like samples and/or remains local to the original community (marked with an X if so). These methods include some involving Forest Fire (FFS) and Random Walk (RW) sampling as well as Markov Chain (MC) and Markov Chain Monte Carlo (MCMC) methods. The values $n \in \mathbb{N}$ and $p, q > 0$ represent parameters of the methods that must be chosen by the experimenter.

3.2 Materials and Methods

This section deals with the specific PIN data used, and describes the three approaches to community selection that will be compared and contrasted in Section 3.3.

3.2.1 CommFinedWalker: Constrained Local Sampling

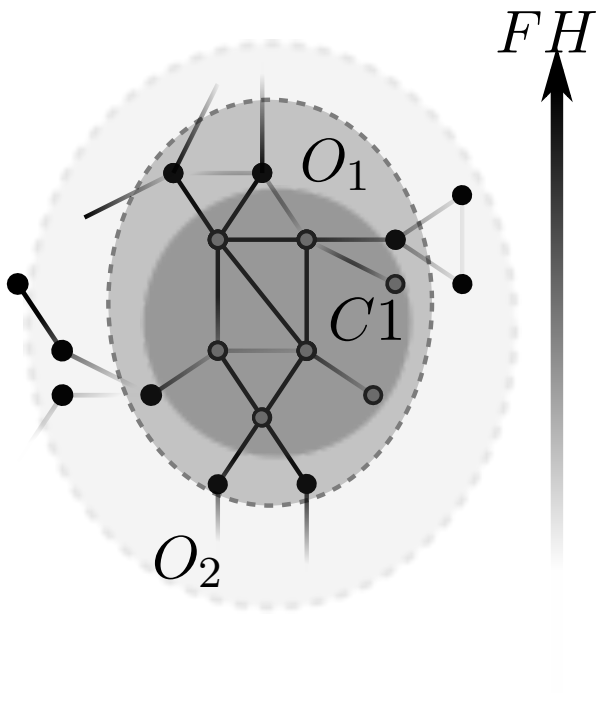


Figure 3.3: This network subregion illustrates a typical example of what a hypothetical relatively homogeneous community $C1$ might look like as larger and larger restart radii are considered. The shading of edges by their FH measure reflects that the internal-internal FH scores are generally higher than their external-external or external-internal counterparts and that FH score dissipates with distance from the original community. This dissipation is expected for highly homogeneous, highly annotated communities that differ from the rest of the network. This is what makes comparison to similar homogeneity conditions near to the community more appropriate for assessing the community's relative research potential as a functional module.

The goal of CFW is identical to that of CW in controlling for inspection bias while providing a community quality statistic that can be used to select for communities that may contain novel and interesting proteins. Just as in CW local quality is assessed by comparison to local pseudocommunities (Section 3.1.7).

In order to account for the small world and aggregation effects discussed in Section 3.1.8, CFW uses a form of restricted sampling known as Random Walk with Restart (RWR) [386]. RWR is

an augmented form of random walk sampling where instead of moving from successive node to successive node, the walk has some probability of returning to the initial protein i and sampling the next protein from amongst those interacting with i instead. The walk may return to i an unconstrained number of times in the process of generating a community sized sample. It is important that the probability of restart should not be so stringent as to prevent a complete sampling of the local network region or cause an infinite loop.

Here the probability of restart depends on the next protein selected to be the successor. If the would-be successor j exists outside of some radius of the community (as given by a snowball sample of proteins around $V(C_k)$) then j is not sampled and the process returns to i. We define the restart probability, p_r for a successor protein j for a walk starting in community C_k as

$$p_r(j, C_k) = \begin{cases} 0 & j \in O_2(V(C_k)) \\ 1 & \text{otherwise} \end{cases},$$

where $O_2(V(C_k))$ is the 2^{nd} order snowball sample (neighbourhood) of community C_k (see Appendix B.1). In this way we can define a local neighbourhood or “restart radius" for each community C_k and the boundary outside of which no proteins are sampled. The mean shortest path length provides a lower bound on how many steps are needed to escape the local area of the community. We assume this restart radius parameter to be two throughout as this is lower than the upper limit, assumed to be the mean shortest path length for the two PINs under consideration (see Table 3.2) as well as many other PINs [118,430]. We call this parametrisation CFW-R2 (abbreviated to R2 in figures and tables).

As discussed in Section 3.1.8, CW T-values may not be reflective of local annotation quality, due to escape from the community's local background region. CFW walks are more likely to stay near the starting node and thus sample more of the original community, which should produce a shift in T-value distribution (see Section 3.3.2). A new threshold T^*_{R2} may therefore be required to replace the one used in the previous assessment of CW, $T^*_{CW} = 0.5$ [236]. Selection of the T-value threshold T^*_{R2} for the CFW method was tuned so that the total number of proteins in a network that were in significant communities, those below threshold, was similar to the number obtained for CW with $T^*_{CW} = 0.5$. We term proteins from significant communities to

be significant proteins (though many may not be functionally enriched).

Important to the assessment of CFW is another measure, community LHM (CLHM). CLHM is calculated using the Local Homogeneity (Eq. (3.3)) of each protein in the community. For a community C_k, $CLHM_{C_k}$ is calculated as the mean over LHM_i for $i \in V(C_k)$,

$$CLHM_{C_k} = \frac{1}{|V(C_k)|} \sum_{i \in V(C_k)} LHM_i, \tag{3.5}$$

where LHM is calculated as in Eq. (3.3). This gives a measure of the local homogeneity context, and a proxy for the local annotation quality of each community.

3.2.2 Global FH Benchmark Testing

The universal threshold selected for the classical FH benchmark community filter is the global median FH between all pairs of interacting proteins in the network, FH_{med},

$$FH_{med} = \underset{e_{ij} \in \pi_{\mathscr{A}}[E(G)]}{\text{med}} s_{ij}, \tag{3.6}$$

where s_{ij} is the pairwise semantic similarity as defined in Eq. (3.1.5). Any community that exceeds this threshold was considered to be significantly homogeneous under this test and a possible functional module.

3.2.3 Sourcing of Interaction Data

For comparison with previous results (CW), we use the same data from the BioGrid and HINT databases [83,100]. The BioGrid data includes both Associative (A) and Phyical (P) interaction data, whilst the HINT data contains a smaller set of higher confidence interactions, only of the P-type. By isolating the largest connected component and eliminating self-loops, two undirected networks were produced which we call BIOGRID-AP and HINT-P (accessed August 2015, to be consistent with CW results in [236]).

The two networks differ not only in their interaction types but in their reliability. HINT-P requires two corroborating sources per interaction, while BIOGRID-AP is richer in both proteins and interactions but has a higher rate of false positives than HINT-P. Table 3.2 summarises the properties of both sparsely clustered networks.

network	size	edges	GC	ALC	% annot.	$\overline{\ell_{ij}}$
BIOGRID-AP	15405	165343	0.055	0.13	84%	3.27
HINT-P	10927	49301	0.034	0.099	90%	4.02

Table 3.2: Network statistics for the two PINs BIOGRID-AP and HINT-P including network size and number of links. GC and ALC are global and average local culstering coefficients respectively. The mean $\overline{\ell_{ij}}$ is the mean minimum path length between all pairs of nodes in the network and % annot. reflects the percentage of proteins in the network annotated by GO terms.

3.2.4 Parametrisation of Community Detection

As meaningful biological structure can exist at all levels [284], we investigated 51 values of the γ resolution parameter using CMM clustering, taking values on a $\log_{10}$-scale in the interval $[10^{-1}, 10^{3}]$. Here, 10^3 represents the maximal resolution which most favours the detection of smaller communities. This is the same parametrisation as used in the assessment of CW [236].

In order for communities to be investigated they must be both computationally tractable and experimentally viable. These considerations constrain the size of communities investigated. We selected the size range of interesting communities to be between 6 and 35 proteins, as in the previous CW study.

3.2.5 Sourcing and Measuring Community Co-expression Data

Protein expression profiles, $\mathscr{E}$, were sourced from whole mRNA data from 8500 post-mortem, tissue-specific samples which were sequenced using RNA-seq and normalised by both read length and sample size using the RPKM method [234, 269]. We use a commonly applied measure of co-expression (see Section 3.3.4). The measure was based on the absolute value of the Pearson correlation between two proteins i and j, $|\rho_{ij}|$ [217].

3.3 Results

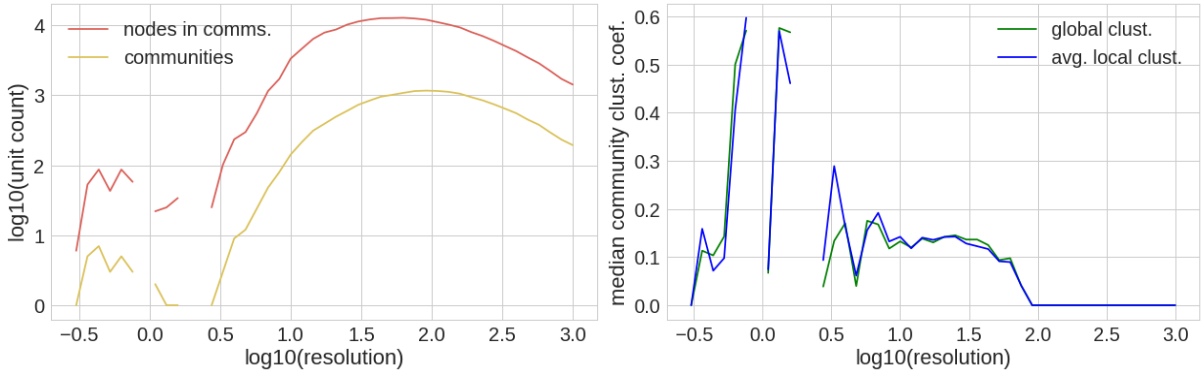
The results consist of a comparative analysis of CW, CFW with restart radius 2 (CFW-R2 or just R2) and the direct mean FH benchmark method, all using the Pandey similarity measure to define FH (Section 3.1.5).

3.3.1 Community Properties

Community detection was carried out using the CMM method on both the HINT-P and BIOGRID-AP networks as detailed in Section 2.2. The size of communities detected by the CMM method

is highly sensitive to the value of the resolution parameter γ. For low values of γ, larger communities are favoured that could very well occur outside of the 6 to 35 community size range of interest. Similarly, for very high values of γ, many more very small communities are chosen that may be smaller than the lower bound of 6, resulting in a peaked distribution of testable communities for both networks. Figure 3.4 shows these results for BIOGRID-AP (with qualitatively similar results for HINT-P, not shown).

Non-uniformity in community sizes extends to community clustering coefficients in BIOGRID-AP where only a minority of high resolution communities exhibit any of the highly clustered structure classically associated with high modularity. This non-uniformity is likely a reflection of the fragmented distribution of clustered regions throughout most PINs which are broken up further by forcing smaller partitions of the network at high resolutions. Similar resolution-community and resolution-clustering relationships exist for HINT-P with the median clustering coefficient vanishing to 0 at $10^{1.48}$. The variation in clustering suggests that resolution selection may be dependent not only on the biological scale of interest but also on the structural requirements of desired communities.

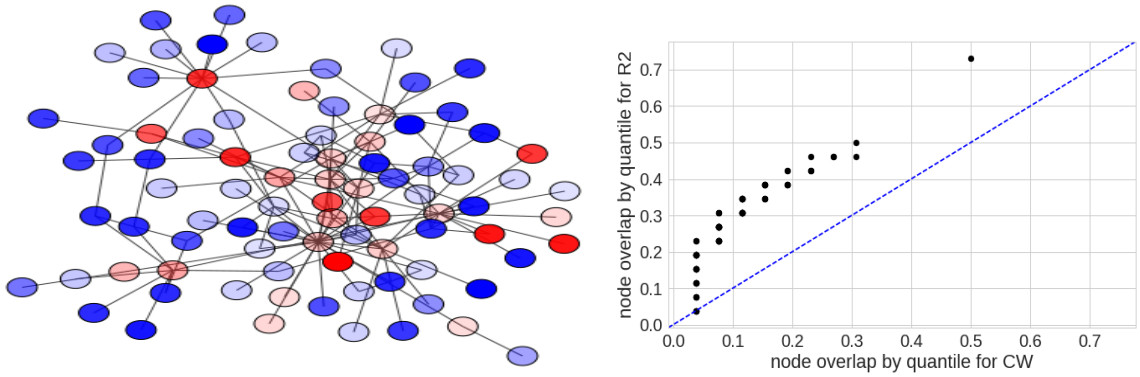


(a) (b)

Figure 3.4: This figure shows the number of communities detected at each resolution within the 6 to 35 size range and the mean number of nodes per community for BIOGRID-AP (Figure 3.4a). The number of nodes in communities have similar trajectories, increasing as expected as the resolution increases (due to detection of more sufficiently small communities), until reaching a saturation point around $10^{1.8}$ (as communities identified by CMM become too small). Note that for certain low resolutions no communities were detected in the given size interval. Figure 3.4b shows median community clustering coefficients per resolution. The gradual decrease and eventual vanishing of the clustering coefficients suggest how resolution effects community structure with less than half of communities just above the saturation point exhibiting any global or local clustering. The large amounts of clustering for low resolution parameter values are likely unreliable due to the small sample size of communities in the 6 to 35 band.

3.3.2 The Relationship Between T^* and Sampling Method

The most important methodological difference between CW and CFW-R2 lies in where nodes can be sampled from in each random walk. In CW walks begin within the community but could eventually leave for unrelated regions of the network. In CFW-R2 this is much less likely as leaving the community is likely to result in restart after only a few steps. The result is a higher proportion of overlap between the pseudo and true communities for CFW-R2. An example from HINT-P illustrating the difference in overlap distribution for a specific community (shown in Figure 3.5a) is given in Figure 3.5b where quantiles of CFW-R2 pseudocommunities are much more skewed towards high overlap with the original community than are CW pseudocommunities.



(a) (b)

Figure 3.5: An example single step neighbourhood of a 26 node community from HINT-P is shown in Figure 3.5a with protein members (nodes) coloured by community membership (red) or non-membership (blue). Nodes are shaded darker if they appear in a high total number of CFW-R2 walks as a fraction of total CFW-R2 plus CW walks. All proteins in this local region are visited more by CFW-R2 (74% of the total CFW-R2 plus CW walks). Figure 3.5b compares CFW-R2 to CW in terms of their pseudocommunity overlap (expressed as a node fraction) with the original community. The comparison is a Quantile-Quantile plot of pseudocommunity overlap that shows that CFW-R2 pseudocommunities are clearly skewed towards much higher overlap. These plots suggest a general trend in which CFW-R2 tends to sample nearby nodes with higher overlap with the original community than CW.

Table 3.3 provides the T-value and FH thresholds used in filtering for significant communities. The larger threshold values associated with CFW-R2 are due to the tendency for high overlap, region-restricted pseudocommunities to be more homogeneous than unrestricted pseudocommunities which aggregate annotations from distantly related proteins. The result is that CFW-R2 T-values tend to be greater than CW T-values for all communities, making a higher threshold necessary to allow similar numbers of proteins to be selected. We address the potential caveats of such a coarse-grain adjustment in the discussion, Section 3.4.

network	T^*_{CW}	T^*_{R2}	med(FH)
BIOGRID-AP	0.5	0.5600	6.10552
HINT-P	0.5	0.5875	6.04852
	n_{CW}	n_{R2}	n_{FH}
BIOGRID-AP	51869	51930	40212
HINT-P	49979	49822	40157

Table 3.3: Threshold values for methods on BIOGRID-AP and HINT-P along with the associated number of signigicant proteins summed over all resolutions (allowing for repeated counts) that were optimised over to provide a close match between the total number of significant proteins for CW and CFW-R2 in both networks. This was done to allow for comparison of the properties of the proteins and communities selected.

3.3.3 Community Selection and Inspection Bias

Without knowing the level of inspection bias in each region of the network, it is impossible to know the level of homogeneity which is appropriate to a community's local background. However, it is reasonable to suppose that a somewhat more uniform distribution of selected communities across homogeneity conditions will be favoured when annotation bias is reduced. Previous comparisons between CW and the median FH method (see Eq. (3.6)) have shown that the distribution of selected proteins is skewed towards higher local homogeneity (measured by LHM) for both CW and median FH, although this bias is weaker for the random walk method (CW) [236].

Similar results to those for individual proteins are seen at the community level when sorting each community from across all resolutions according to their Community Local Homogeneity score (CLHM, see Eq. (3.5)). Table 3.4 shows that all methods tend to select the majority of locally well annotated communities in CLHM quartile one but CFW-R2 communities make up a much larger (but still comparatively small) proportion of quartile four communities, which are the communities most vulnerable to being overlooked due to annotation bias. However, for the majority of communities CFW agrees with CW. It is interesting to note that for BIOGRID-AP the distribution is markedly more even between methods and the levelling effect of CFW-R2 less severe than for HINT-P. This could be due to possibly higher levels of inspection bias in HINT-P, where a larger proportion of data comes from a smaller set of targeted and double validated studies.

Table 3.5 breaks down proteins into LHM quartiles and gives the percentage of proteins in that quartile that were ever part of an appropriately sized community that was selected by the

HINT	R2	CW	FH	R2∩FH	CW∩FH	CW∩R2	R2∩CW∩FH
Q1	53.1	63.6	70.3	52.8	61.3	52.7	52.4
Q2	36.5	34.4	24.7	24.7	24.7	32.5	24.6
Q3	21.1	13.5	4.0	4.0	4.0	13.4	4.0
Q4	14.1	5.2	0.5	0.5	0.5	5.2	0.5
Total %	31.2	29.2	24.9	20.5	22.6	26.0	20.4
Total #	4793	4484	3823	3147	3476	3989	3130
BIO	R2	CW	FH	R2∩FH	CW∩FH	CW∩R2	R2∩CW∩FH
Q1	56.3	59.2	55.9	53.5	54.8	56.2	53.5
Q2	26.5	26	18.4	18.4	18.4	25.4	18.4
Q3	13.1	11.5	5.3	5.3	5.3	11.5	5.3
Q4	9.0	6.4	2.9	2.9	2.9	6.4	2.9
Total %	26.2	20.6	24.9	20.0	20.3	24.9	20.0
Total #	4700	4619	3695	3587	3645	4460	3587

Table 3.4: Percentages of selected (below threshold) communities in the 6 to 35 size band in each quartile when distributed by mean LHM in HINT-P (HINT) and BIOGRID-AP (BIO). All three methods (CW, R2 and FH) exhibit decreases in selected communities as community LHM (CLHM, Eq. (3.5)) decreases. FH shows the largest discrepancies in representation between quartiles, followed by CW. Differences are most apparent in HINT-P where CFW-R2 communities with the poorest annotated neighbourhoods (lowest CLHM) more than double CW communities. Both CFW-R2 and CW largely encompass FH communities, partially because there are much fewer such communities, especially for poorer neighbourhoods. In comparison to CFW-R2, CW has marginally higher overlap with FH selected communities regardless of the level of local homogeneity (LHM).

HINT	R2	CW	FH	R2∩FH	CW∩FH	CW∩R2	R2∩CW∩FH
Q1	80.0	87.2	82.9	76.6	82.6	79.0	76.6
Q2	68.3	70.4	58.8	55.1	58.4	64.0	55.0
Q3	55.6	49.2	32.1	31	32.0	46.1	30.9
Q4	44.6	30.1	13.4	13.1	13.3	29.1	13.1
Total %	62.2	59.2	46.8	44.0	46.6	54.6	43.9
Total #	6793	6473	5116	4807	5090	5964	4797
BIO	R2	CW	FH	R2∩FH	CW∩FH	CW∩R2	R2∩CW∩FH
Q1	76.3	77.3	72.3	71.7	72.0	76.0	71.7
Q2	63.6	63.6	52.2	51.8	52.1	61.8	51.8
Q3	52.6	50.8	37.7	37.5	37.6	49.7	37.5
Q4	35.9	31.9	18.5	18.3	18.5	30.7	18.3
Total %	57.1	55.9	45.2	44.8	45.0	54.6	44.8
Total #	8799	8609	6960	6909	6937	8405	6908

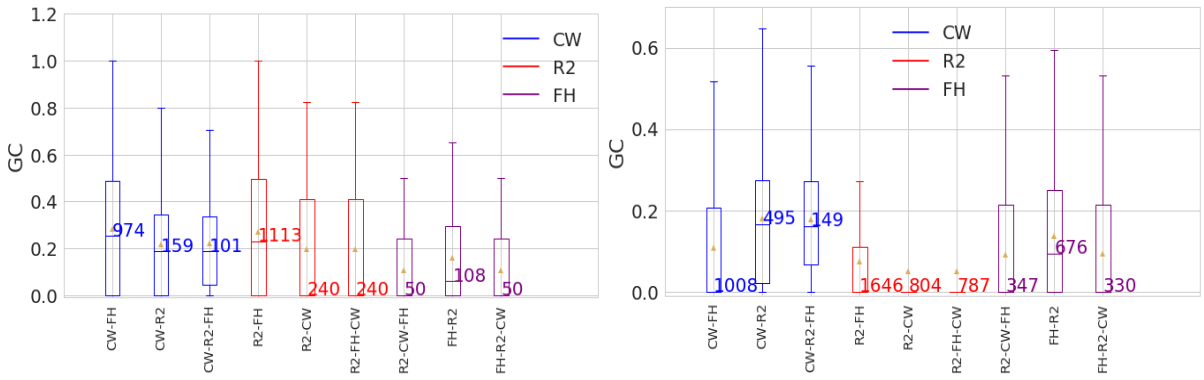
Table 3.5: The distribution of proteins into quartiles by their LHM. Percentages indicate the fraction of proteins ever selected by one of the three methods or their intersection as part of a community at any of the 51 resolutions. The Total %s are the percentages of total PIN proteins selected by that combination of tests at any resolution. The Total # is the number of such proteins selected (without normalization). Again, there appears to be widespread agreement between methods meaning that a significant fraction of proteins (particularly those in lower quartiles) are never selected by any method, indicative of the gap in our knowledge of large regions of the interactome (either annotation or interaction space).

method shown, at any resolution. The high proportion of proteins (Total %) selected by each of the methods at some resolution corroborates the hypothesis that interesting structures do exist

across multiple biological scales within the networks and that of the three methods CFW-R2 selects a slightly broader set of proteins than CW in both networks.

Table 3.5 also suggests that even though varying the resolution can help to detect different high quality communities, some communities have either been poorly partitioned or so poorly annotated that no functional enrichment is detectable under any of the methods investigated here. This could be due to multiple factors from inspection bias at the annotation or link level to shortcomings in the community detection algorithm which prevent it from finding counter-intuitive but otherwise interesting communities. PINs as they are constructed and analysed are only a noisy approximation to the real functional relationships between proteins, which become less accurate the less data is available.

The global clustering coefficient (GC) can be used as a measure of member association between



(a) BIOGRID-AP (b) HINT-P

Figure 3.6: Comparative Box plots of global clustering (GC) coefficients of communities (size 6 to 35) exclusive to each method when compared to the others. Each group is a cluster of three plots comparing communities selected by one method A, and not one or both of the others B, using the set notation $A - B$ for A but not B. The coloured text reflects the number of communities in that comparison class and is in line with the GC median of that class. Communities selected by CFW-R2 and not CW tend to have a lower GC while the relationship between both methods and FH seems to depend on whether network, BIOGRID-AP (Table 3.6a) or HINT-P (Table 3.6b).

proteins in a community (see Eq. (2.13)). Figure 3.6 shows that communities that are unique to each of the three methods do not have the same distribution of GC. CFW-R2 only communities tend to have a lower GC than CW communities independent of the PIN investigated while the relationship between FH and the other two methods seems to depend on the network. This could be partially explained by the tendency of CFW-R2 to select out more low homogeneity communities coupled with weakly positive correlation between local homogeneity and GC value (Spearman correlation of 0.31 for HINT-P and 0.18 for BIOGRID-AP, $p < 10^{-6}$). Communities

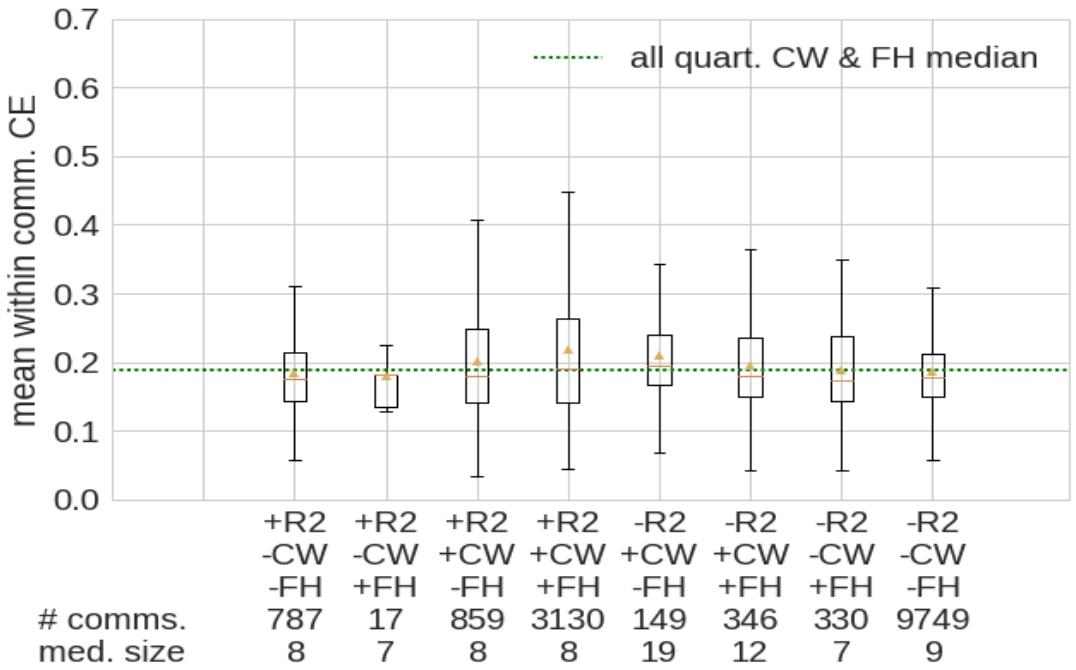
with low GC tend to be more tree-like, lacking transitive edges and therefore a lower number of cycles in general. Proteins in these communities could therefore be some of those most neglected and therefore vulnerable to link and annotation inspection bias.

3.3.4 Co-expression-based Community Validation

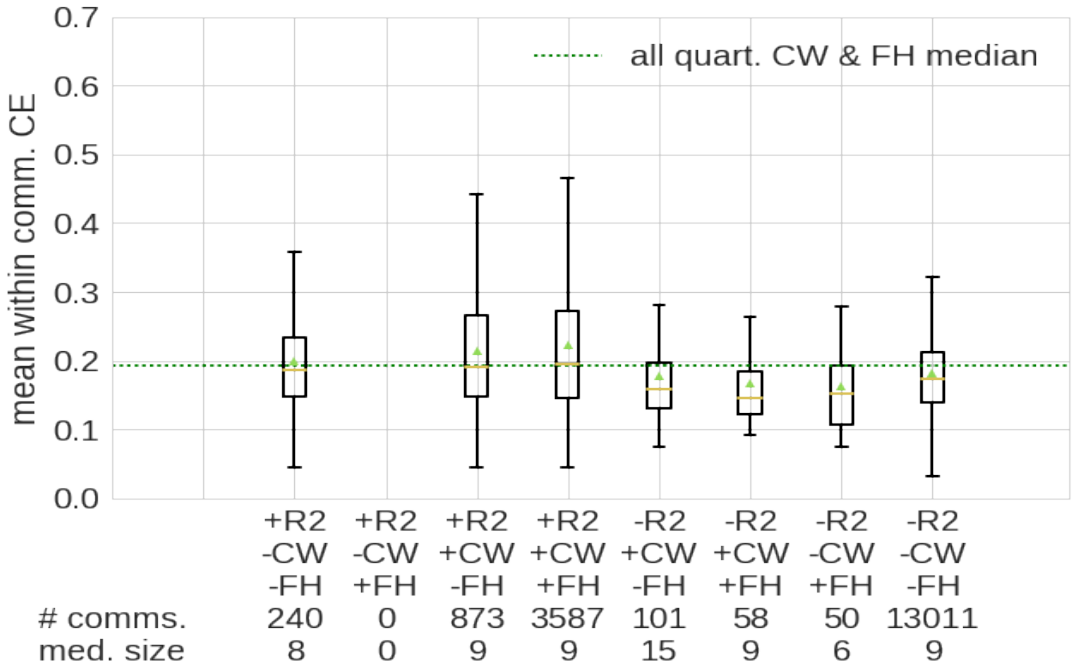
The distribution of community co-expression scores is not well understood, and small sample sizes (e.g. of FH communities in lower homogeneity quartiles) and repeated communities make approximate normality and independent sampling assumptions dubious. Therefore, we utilized a Monte Carlo test comparing the difference in mean community co-expression score between methods, d, to the differences in mean co-expression score of samples of equal sizes to the originals drawn with replacement from the entire score population. The calculation of each re-sampled difference d_i was performed a total of 10^5 times. The null hypothesis of the test is that the difference is no greater than would be expected by random sampling of communities from either method which can be rejected with approximate p-value $\hat{p}$ when $d > d_{(1-p)}$ (the empirical $(1-p)^{th}$ quantile of the 10^5 measurements of d_i). Tests are motivated by differences in the median or spread of their co-expression (see Figure 3.7 and Figure B.2 to B.5 in Appendix B.3).

Figure 3.7a (HINT-P) and 3.7b (BIOGRID-AP) summarise the mean co-expression behaviours of communities selected by any combination (combined by set intersection and difference) of the three methods. Both networks suggest that CW and CFW-R2 communities have very similar overall expression profiles. CFW-R2 communities perform slightly below the baseline (given by CW∩FH) in both but not as badly as CW only communities in BIOGRID-AP. Note also that CW communities tend to be much larger (as indicated by the size median) in both networks. The small number of FH only communities also have a lower median co-expression quality than either random walk method. This indicates that of the methods investigated, co-expression relationships may be better preserved by using the two random walk methods. This difference is particularly pronounced for BIOGRIOD-AP (Figure 3.7b), but tests suggest that FH only communities tend to have lower co-expression then CW and CFW-R2 only communities in both BIOGRID-AP and HINT-P ($\hat{p} < 10^{-4}$). In general, effect sizes at the population level, particularly for HINT-P are not large when compared to differences between communities when they are sorted into quartiles by CLHM.

When partitioned by quartile according to their CLHM (Appendix B.3, Figure B.2 and B.3



(a) HINT-P



(b) BIOGRID-AP

Figure 3.7: Box plots of HINT-P (Figure 3.7a) and BIOGRID-AP (Figure 3.7b) community co-expression for all sets of communities selected (+, intersection) or not (-, set difference) by each combination of methods (means for each combination marked with a triangle). The text below each plot shows the number of communities above the median size of communities in that class. CW∩FH median co-expression (with or without CFW-R2 as in [236]) is used as benchmark value (dotted line). Similar co-expression trends exist in both HINT-P and BIOGRID-AP with the highest community co-expression levels occurring for cases where two or more of the random walk-based methods intersect. Community co-expression values for CFW-R2 only communities appear to be slightly below the benachmark in both BIOGRID-AP and HINT-P but comprable to CW only communities.

for HINT-P, and Figure B.4 and B.5) for BIOGRID-AP, the networks exhibit largely similar patterns to Figure 3.7 that hold across levels of local homogeneity. As before (Table 3.5), all methods generally agree for higher CLHM quartiles (Figure B.2a and B.4a) with co-expression amongst CW∩FH communities and communities with all three showing the highest levels of co-expression quality. The first quartile is where the large CW communities are located, but CFW-R2 also tends to select larger communities in these quartiles.

Note that on BIOGRID-AP there is less disagreement between selected communities across all homogeneity quartiles and thus less clear distinctions between methods, particularly in lower quartiles (Figure B.5). However, some differences are clear across networks as for quartile two and beyond there is a lack of FH only communities (Figure B.2b and B.4b), demonstrating the classical methods shortcomings when distinguishing even moderately annotated communities. CFW-R2 communities on the other hand have limited spread in co-expression quality, even where moderate sample sizes are available, with a comparatively small interquartile range. This is true even in lower quartiles where CFW-R2 has many more communities than CW or FH (see Figure B.3 and B.5). This suggests that CFW-R2 communities tend to have relatively consistent co-expression quality across homogeneity levels. This may be due to decreased levels of inspection bias and comparatively greater sampling of important functional modules.

In general communities that were selected by at least both random walk methods had the highest co-expression. Counter-intuitively, this improvement in co-expression quality with random walk sampling becomes more apparent with decreasing local homogeneity. We compare the first quartile (2024 and 2517 CFW-R2 and CW only communities in HINT-P and BIOGRID-AP respectively) to the lower three quartiles (1965 and 1943 communities for CFW-R2 and CW in HINT-P and BIOGRID-AP respectively) using the one-sided Monte Carlo test detailed above to determine whether high homogeneity communities tend to have lower co-expression scores. The test indicates ($\hat{p} < 10^{-4}$ for both networks) that high homogeneity (quartile one) communities selected by CFW-R2 and CW tend to have lower co-expression than their lower homogeneity counterparts (quartile two to four). This suggests that the random walk methods are able to pull out important functional information in these annotationally poorer regions. This could indicate that utilizing both random walk methods may help to pick out high quality communities in low information regions of the network that are co-expressionally active and interesting.

3.3.5 Case Study: Bardet-Biedl Syndrome

In this case study we walk through the analysis of a module ranked important according to CFW but not CW. This shows the potential usefulness of CFW as well as how the method could be used in practice. In Section 3.4 we discuss how CFW can be part of a combined approach using multiple methods to each highlight different areas of functional module space.

We use topGO (See Appendix B.2), a gene set enrichment analysis tool, to examine the GO term enrichment of the top communities (the gene sets) selected by CFW-R2 but not CW across a range of resolutions in HINT-P. This range was chosen so that the majority of communities had clustering coefficients greater than zero ($[10^{-1}, 10^{1.48}]$ in HINT-P), removing repeated communities across resolutions. The resultant community ($T_{R2} = 0.29 < 0.50 = T_{CW}$) was found to be associated with a rare genetic disorder, namely Bardet-Biedl Syndrome (BBS).

BBS is a rare autosomal recessive genetic disorder of the cilia with highly variable pathology and symptoms ranging from polydactyly to retinal deterioration. BBS is known to involve dysfunction in one or more of at least 14 BBS genes [134]. Of these, five are suspected to interact directly: BBS1, BBS2, BBS4, BBS7, BBS9 as suggested by co-localization and functional analysis. All five appear in the selected community along with BBS8 (TTC8), the first protein through which the cilial malformation hypothesis of disease development was discovered [23]. Also present are homologs of suspected interactors and related components of proteins thought to mediate the function of BBS genes including six keratin family proteins, two exocyst complex components (EXOC2 and EXOC8) [279].

The community was found to contain proportionally less annotated proteins (74%) than the network as a whole (90%) with a community FH of $5.01 < 6.04 = FH_{med}$. Terms found to be highly significant ($p < 10^{-4}$, topGO in Appendix B.2) amongst proteins in the community when compared to the rest of the PIN included those involved in visual perception, photoreceptor cell maintenance and cilium assembly and maintenance. Figure B.6 in Appendix B.4 shows the BBS community subnetwork embedded in its two-step network context. At present the mechanism behind BBS is still poorly understood; but identification of interesting but understudied proteins strongly associated with BBS and other rare diseases may help to bridge the gap.

3.4 Discussion

CFW presents a modest improvement over CW in selection of communities from more diverse annotation backgrounds resulting in a possible decrease in inspection bias. However, the less homogeneous communities selected by CFW are also smaller, more tree-like modules and have lower clustering coefficient, suggesting that they may not possess features typically selected for by community detection methods. This atypical lack of clustered structure could be due to the already established relationship between protein annotation and degree in PINs (interaction or edge inspection bias) [331]. It may also be due to these regions containing regulatory and metabolic pathways that are relatively isolated in terms of their functional interactions with the rest of the network but which nevertheless are highly co-expressed.

Despite the apparent differences in the structure of some communities selected by CFW and CW there is a high degree of overlap between the communities they select when compared to the benchmark method, especially at higher levels of local homogeneity. This is partially explained by the high level of similarity between the two random walk-based sampling methods employed and the small world property of PINs. Methods with greater coverage of protein community-annotation space may be those that are less restricted by locality as a proxy for functional homogeneity and annotation quality. In this respect, an exciting area for possible future exploration is combined network-annotation embeddings of proteins for example using node2vec [159] or heat kernels. Such an embedding could retain both locality and annotational similarity information. This could then be used to sample topologically and annotationally similar areas of the network (e.g. by a kernelised probability centred on the community members) without the restriction of having to necessarily stay local in network space.

Previously, researchers have found that weighted networks provide certain advantages when working with gene annotations [98]. Translating CFW into the weighted network context is relatively straight forward to implement but may have undesirable effects as low probability (low weighted) edges could provide potential ways for walks to escape the local area. This could be addressed by using a combined approach, censoring low confidence interactions while retaining the weighted character of the network edges.

A reasonable limitation of our method is the computational constraints imposed by having to generate large random walk samples which force us to consider only relatively small communities

for analysis. This causes possible splitting of related proteins across communities, possibly inflating T-values. Newer pseudocommunity sampling methods that can be more easily parallelised or optimised may be useful here [385]. Another possible issue is our use of the T-value threshold to adjust for the lower number of pseudocommunities sampled by our method when compared to CommWalker. Although this is likely due to the higher overlap when using CommFinedWalker, it is important to note that since the contribution of each pseudocommunity is binary, it should be relatively robust to whether the apparent difference between the pseudo and original communities is large or small, diminishing problems caused by partial overlap.

One possible way to further validate the effectiveness of the method, that is outside the scope of this chapter, would be to contrast more current annotation and interaction data from the same sources as the historical data used here. Continued monitoring of annotational inspection bias as measured by correlation between local levels of annotation and homogeneity should provide an indication of the continued importance of this research and inform ways in which the method succeeds or s. As more proteins become annotated it may also become possible to assess the predictive power of the quality measure presented here for finding novel biology when compared to more global methods of functional module detection. In this vein, in Chapters 4 and 5, we propose applications of a CW/CFW-like framework to a novel context, discovery of key functional in brain state networks.

Recent research suggests that the gap between poor and well annotated proteins is widening and is likely to remain a feature of protein data as long as medical, personal and financial incentives exist to investigate some proteins and pathways in lieu of others [169]. This means that further research and adoption of bias control methods will be necessary. As text mining and experimental error controls develop further future methods that take more explicit account of differences in source assays and literature, will likely become more important. It is clear that protein research will need to adapt as the opportunity cost of over investigation of only a few key proteins continues to mount.

4 | Brain State Modelling Framework

This chapter includes a paper that has been submitted to the journal *Applied Network Science* as part of the *Communities in Networks 2021* conference special issue. The content of the article was first presented as a contributed talk at the *Communities in Networks* satellite conference of *Networks 2021*.

The article to follow presents an analysis framework for a class of Hidden Markov Models (HMMs) we term *Hidden Markov Graph Models* (HMGMs), but are not referred to as such in the article itself. The framework is applied to the same wakeful fMRI data that is part of the anaesthesia neuroimaging dataset presented in Chapter 5 (but applied to the full 15 subject pilot study dataset).

HMMs have been used to construct multivariate dynamic representations of changing brain state [408]. Appendix C presents an overview of brain region parcellation and model training not presented in the article. This includes a cross-sectional view of the parcellation (i.e. subdividing of the brain into regions) used in the model, Figure C.1. It also includes details on the Variational Bayes approach to model training using free energy C.1 based on the method in [408].

HMGMs are constructed by converting the multivariate brain state models into brain state graphs in which each node is a brain region (annotated with mean activity) and weighted edges represent information shared between regions. The brain state graphs form the layers of the HMGM which is an annotated, multiplex graph model (see Section 2.3.1). Layers are connected by directed, weighted edges that represent the probability of interstate transition. This article develops and validates the modelling framework and applies it in a context in which brain ac-

tivity is relatively well understood, resting wakefulness. The framework is then modified and expanded upon in Chapter 5, in the study of deep anaesthesia which was the intended use-case for this HMGM framework.

Ranking of Communities in Multiplex Spatiotemporal Models of Brain Dynamics

James B. Wilsenach, Catherine E. Warnaby, Gesine D. Reinert

As a relatively new field, network neuroscience has tended to focus on aggregate behaviours of the brain averaged over many successive experiments or over long recordings in order to construct robust brain models. These models are limited in their ability to explain dynamic state changes in the brain which occur spontaneously as a result of normal brain function. Hidden Markov Models (HMMs) trained on neuroimaging time series data have since arisen as a method to produce dynamical models that are easy to train but can be difficult to fully parametrise or analyse. We propose an interpretation of these neural HMMs as multiplex brain state graph models. This interpretation allows for dynamic brain activity to be analysed using the full repertoire of network analysis techniques. Furthermore, we propose a general method for selecting HMM hyperparameters in the absence of external data, based on the principle of maximum entropy and use this to select the number of layers in the multiplex model. We produce a new tool for determining important communities of brain regions using a spatiotemporal random walk-based procedure that takes advantage of the underlying Markov structure of the model. Our analysis of real multi-subject fMRI data provides new results that corroborate the modular processing hypothesis of the brain at rest as well as contributing new evidence of functional overlap between and within dynamic brain state communities. Our analysis pipeline provides a way to characterise dynamic network activity of the brain under novel behaviours or conditions.

Keywords: community ranking, generative models, model selection, multiplex networks, networks neuroscience, spatiotemporal networks

4.1 Introduction

The brain activity of healthy subjects at rest is commonly used as a baseline against which a wide range of both pathological (e.g. dementia) and healthy (e.g. sleep) conditions are compared [104, 265, 302]. Often, activity under one condition is modelled as a single static pattern of activity, ignoring largescale dynamic shifts. However, neuroimaging researchers have begun to recognise that subjects move through a wide array of brain activity configurations even while relaxed or asleep [201, 375, 408]. A brain state is a configuration of brain activity

evoked in response to a stimulus or to facilitate more complex responses [68]. Neuroimaging time series provide a way to observe these reconfigurations as spatial patterns of metabolic or electrophysiological activity, termed *functional activity* [287]. In order to generate these patterns, brain regions must coordinate through transfer of information. This exchange between brain regions defines the state's *functional connectivity*. Functional activity can therefore be interpreted as a realisation from a brain state graph model which describes brain dynamics and the relationships between brain regions in the state [40]. In these graph models, nodes are anatomically or functionally defined brain regions and edge strength is determined by the level of information shared between these regions.

We provide a novel framework to determine the number of brain states from data and analyse brain state dynamics as a multiplex graph with modular (community) structure at both the temporal (state switching dynamics) and spatial (brain region communication) levels. Our new modelling and analysis pipeline (see Figure 4.2) makes it possible to determine complex relationships between networks of brain regions changing with time, doing so in an entirely unsupervised way directly from recorded brain activity. Here, *network*, rather than being largely synonymous with *graph* is more closely linked to *community*, and is used exclusively to refer to modular subgraphs of coordinated brain regions within a state. Networks form the basis of our understanding of the functional connectivity pathways within the brain and are integral to our understanding of the role of changing brain configurations in wakefulness and beyond [311].

We have developed a method to identify the importance of possible brain networks using random walks to ascribe to each module in each state an importance or T-score based on their functional connectivity and co-activation. Notably, the method does not apply random walk information to partition the graph but rather to determine the relative importance of communities within a partition [314]. Our method provides a means to characterise dynamic functional activity under novel conditions or behaviours. As a proof of principle, we apply our pipeline to neuroimaging data from subjects at rest and provide new evidence for both modular and nested functional activity in the awake brain.

4.1.1 Static Brain State Models

In the simplest brain state models (see Figure 4.1A), functional activity arises as noisy realisations of a single static brain state. Considerable progress has been made using this static

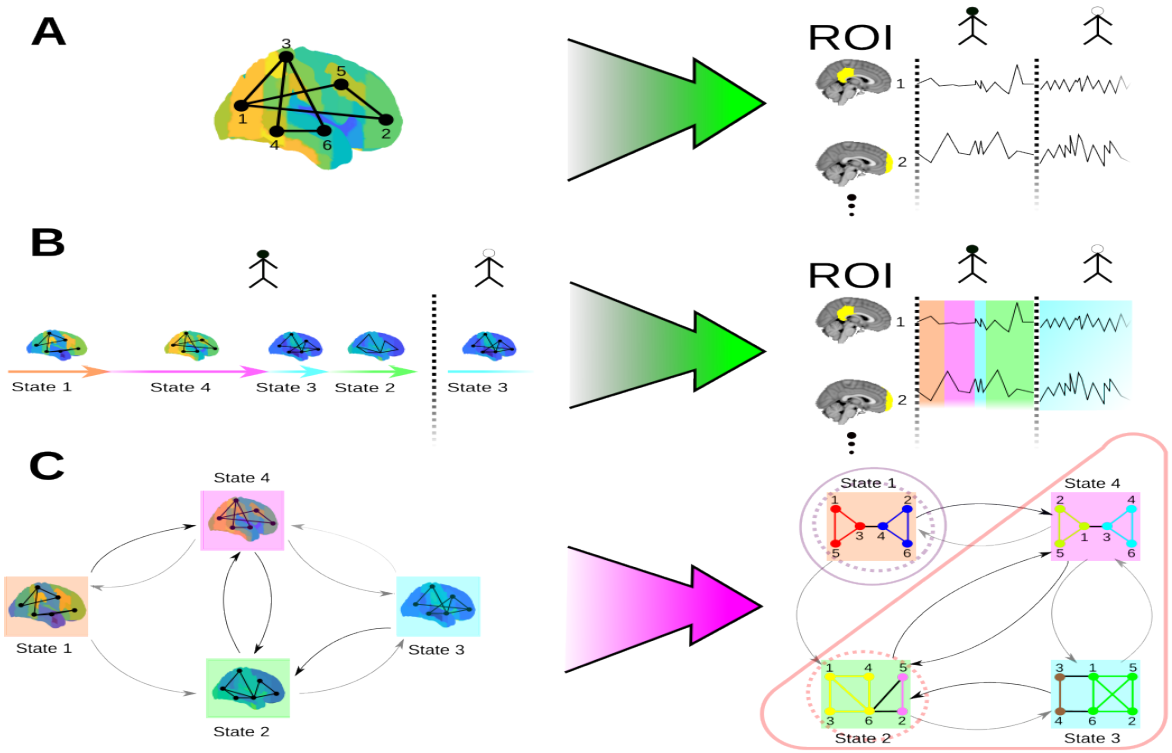


Figure 4.1: This figure shows how brain activity can be modelled as being generated by a system of either static (A) or dynamic (B) states, and, in particular, how such a dynamic brain state model can be interpreted as a multiplex network with modular structure (C). In **(A)** a static pattern (left) of functional activity (colour of functional activity map) and connectivity (edges between regions) is observed (green arrow) as a stationary multi-ROI multi-subject time series (right) in which each dimension is the activity observed for a particular Region of Interest (ROI) in each subject (separated by a dotted line). **(B)** Shows state dynamics for a multi-subject system with multiple states (left). In this system each state is represented by a colour and arrow length indicates its duration in time. This is observed as a multivariate time series composed of weakly stationary segments (right). Segment colour indicates the state that generated it. In **(C)** we use temporal relationships between states to represent the system as a dynamic multiplex graph. This system is decomposed (purple arrow) into its essential temporal (coloured ellipses) and spatial modules or functional *networks* (coloured subgraphs). In this case the system is dominated by a single community with an additional outlying state with weak connectivity, that occupies its own community. In larger networks more complex community structures may arise. Dashed circles around states show state hubs, important states in each community which are central to the dynamics and facilitation of brain activity across subjects.

framework to characterise the vast repertoire of activity patterns observed during wakefulness. Models using both weighted and unweighted graph structures derived from Independent and Principal Component Analysis (ICA and PCA respectively) have revealed key modules within the brain across a wide range of conditions. These include both behavioural and task-based conditions (sensory, motor etc.) and resting state conditions in the absence of direct stimulation [76, 77, 208, 328, 362]. Recent results from both electrophysiological data derived from

Electroencephalography (EEG) and Blood Oxygen Dependent (BOLD) data derived from functional MRI (fMRI), suggest that weighted network models produce more reliably reproducible and robust results than do binarised network models [190, 306, 360].

Studies using static models have helped neuroscientists to build up vast libraries of associations between cognitive functions and specific brain regions [296]. However, the static approach makes it difficult to account for inter-subject variability as well as dynamic changes in state that occur in time as different cognitive and functional demands are placed on the brain [263]. These demands result in activity in one moment that is often functionally incompatible with activity in the the next, driving the need for dynamic approaches to brain state modelling [367].

4.1.2 Dynamic Brain State Models

Moving window-based approaches produce a series of snapshots of the activity pattern of the brain. Although, these methods have proved incredibly useful in understanding changing brain state, they are limited in their ability to reliably detect changes in functional connectivity between regions over time [171]. By contrast, state-space models (Figure 4.2B) and in particular, Hidden Markov Models (HMMs) have arisen as an alternative to the sliding window approach and use a number of simplifying assumptions to improve on these models' tractability and specifiability [375, 409]. The chief underlying assumption of HMMs is that brain dynamics can be parametrised by a finite state, positive recurrent, Markov process where functional activity and connectivity is determined by a simple multivariate normal distribution [408]. Due to the computational role of the brain, recurrent discrete brain states are a likely necessity as the brain needs to be able to reliably store and retrieve information while operating under biological and physical noise constraints [380].

In practice however, HMMs can be difficult to interpret. Here, we interpret the dynamic switching between states as a temporal graph of probable state transitions (Figure 4.1C). The full model can thus be seen as a nested, or multiplex graph in which the layers are brain states (with brain regions as nodes) and the interlayer directed edges are transition probabilities between state layers.

4.1.3 Novel Multiplex Approach

A state characterises a pattern of activity across the whole brain at a given time; however, it is most often characterised in terms of just a few key subgraphs of interacting brain regions (see Figure 4.1C). Much progress has been made to characterise the vast repertoire of activity patterns observed during resting states and task performance. These enquiries have given rise to a number of re-occurring and important networks, associated with a wide range of brain functions and behaviours [54,260,353]. The most prevalent and widely characterised of these are the so-called *resting state networks*, termed the Default-Mode (DMN), Salience (SN) and Central Executive (CEN) Networks as well as those active during sensory and motor tasks including: the sensorimotor, visual and auditory networks [320]. The mechanisms underlying these networks are interdependent with recruitment of one network often necessitating the further recruitment of other networks [201]. Conversely, some networks are known to be largely mutually antagonistic in activity, with DMN and SN activity generally being anticorrelated with sensorimotor-like activity in resting wakefulness [406].

Although state space modelling of brain dynamics is a relatively young field, one key finding has been the multi-scale modularity of brain states. In particular, Louvain modularity-based community detection applied to the temporal graph of state transitions has shown that states are organised modularly into communities under a variety of conscious conditions including resting wakefulness and sleep [369, 409].

In order to construct a set of plausible brain states models we train a number of HMMs with different numbers of states on resting state data. We then utilise our novel cross-validated maximum entropy procedure, based on the maximum entropy principle, to select the HMM that best generalises across subjects [192]. We convert the selected HMM into a dynamic graph model by transforming the state covariance matrices into weighted, directed graphs based on the regional correlations within each state and node attributes given by the state mean activity.

4.1.4 Ranking the Importance of Networks within the Brain

We perform two-level Louvain community detection to discover important communities of brain states (temporal communities) and brain regions within a state (spatial communities). We use community centrality statistics to identify the hub states of key activity in each network. Within each hub state we look at spatial community structure to determine the key actors in the dynam-

ics of the model that may be important to the overall dynamics of wakefulness across subjects.

Random walks provide an effective way to construct representative samples from a graph in a way that preserves local structure [115, 223]. In complex interdependent data sets random walk sampling can be used to remove baseline levels of interdependence and discern the most robust relationships in a one dimensional model, by conditioning out local inhomogeneity in noisy activity [236]. Here, we extend this principle to network sampling across two dimensions, space and time. Our method is based on a non-parametric random walk statistic that combines a temporal walk between layers with a spatial walk between regions. We use random walks to sample plausible patterns of functional network activity from the local functional activity background. We then use the samples as a benchmark against which to score functional coordination in our spatial communities. This statistical score, termed the T-score, is simple to compute given the graph model and putative network and is inspired by a similar method for analysing large, complex protein graphs with metalayer information [236].

Our method allows us to determine which spatial communities are highly co-activated or inactivated relative to the expected dynamics across states in that brain area, providing a generalisable procedure to determine functionally relevant brain state communities. Our within-state community functional associations largely agree with macroscopic analysis of the state functional activity maps, but provide an additional layer of information in the form of networks that provide clarification and depth to our understanding of brain states at the mesoscale.

4.1.5 Metatextual and Network Analysis of Brain State Models

We use the powerful metanalysis tool, NeuroSynth [436], to determine functional associations between each brain state, its most important networks and important functional terms from the literature. Neurosynth provides scores based on either correlations between brain images and the occurrence of a predefined set of terms in the literature or, in conjunction with the NiMARE package [1], a posteriori probabilities of associations between the image and an exhaustive list of literature terms. Using these tools and images derived from our brain states, termed functional activity maps, we provide evidence to corroborate the modular processing hypothesis in resting wakefulness [308]. Key to our findings is that the states associated with resting state networks tend to self-associate while being anticorrelated with sensorimotor associated states.

4.2 Methods

In the following sections, Sections 4.2.1 and 4.2.2, we explain the preprocessing of the data and define the state space (HMM) model and novel model selection criterion. We will see that each brain state s can be thought of as a pattern of activity represented by a weighted graph $G(s) = \{V, a(s), W(s)\}$ in which each node is a brain region $x \in V$, (with $|V| = D$ nodes), each with a level of functional activity $a(s)^x$ attributed to x. Similarly, each edge in $G(s)$ is weighted by $w(s)^{x,y} \in W(s)$ the level of information flow from region x to region $y \in V$ (edge absence is represented by $w(s)^{x,y} = 0$), with $w(s)^{x,y} \neq w(s)^{y,x}$ in general.

As we shall show, Hidden Markov Modelling with our new model selection method, provides a means to construct a dynamic state-space model from multi-subject fMRI time series data in a data driven way. We use inter-regional within-state correlations to determine the state graphs and use the temporal relationships between states to determine the directed interlayer edges (see Figure 4.1C). Lastly, in Sections 4.2.5 to 4.2.8 we set out methods to explore the spatiotemporal modular and functional structure of these multiplex brain state models.

4.2.1 Acquisition and Pre-processing of fMRI Data for HMM Modelling

We used ten minutes of whole brain wakeful fMRI recordings from $N = 15$ wakeful subjects (with eyes closed), acquired as part of a previous pilot study of deep anaesthesia [262]. This data set was chosen to act as proof of concept for later work with the full anaesthesia data set (Chapter 5). The brain volumes produced by the scanner were aligned to the MNI152 standard brain template [132]. This resulted in a high dimensional time series of each subject's fMRI (BOLD) signal for each voxel, with a temporal resolution of 3s and a spatial resolution of 2mm^3 [428].

Recordings were collected separately from each subject. Of the 200 volumes recorded per subject (each time point is one volume), four dummy volumes were removed to exclude any non-steady-state magnetisation effects. This was followed by motion correction with MCFLIRT (Motion Correction FMRIB's Linear Image Registration Tool), spatial smoothing using a Gaussian kernel of 5mm full width half-maximum, global intensity normalisation, and temporal high-pass filtering with a cutoff of 0.02Hz to remove low frequency scanner drift. Automated removal of non-brain tissue was initially performed before statistical analysis using BET (Brain Extraction Tool), with further manual correction in FSLview. Further spatiotemporal artefact removal was

carried out by independent component in FSL melodic [428].

We selected regions of interest in our study based on the Harvard-Oxford (HO) probabilistic cortical and subcortical brain parcellations, which assigns to each voxel a probability for each brain region. We assign each voxel a unique region identity according to the maximum probability across regions in the HO parcellation. Excluding white matter regions the resulting parcellation of 63 Regions of Interest (ROIs) includes 48 cortical and 15 subcortical brain regions [79, 245]. ROI time series were calculated using the ROI spatial mean BOLD signal at each time point. This results in a $D = 63$ dimensional time series with $T = 196$ time points per subject. Each of the D constituent ROI time series were temporal mean subtracted and normalised by the standard deviation.

Model fitting presents two challenges, the first is that the time taken to fit the model scales with parametric complexity, and the second is that a poorly parametrised model may lead to overfitting or underfitting. To address these challenges, dimensionality reduction by principal components of the original D dimensional time series was performed to reduce parametric complexity while also reducing overall noise. This approach is justified by the generally low embedding dimension of most real world data, including neuroimaging data [239, 348]. In order to balance dimensionality reduction and retention of signal, Parallel Analysis is used (see Supplementary Information Section D.0.1) to obtain a $D \times d$ eigenmatrix A of the first $d < D$ eigenvectors [175]. This method assumes roughly linear separability of uncorrelated noise from signal, but has been shown to outperform a number of methods, including maximum likelihood estimation, in simulation [183]. The reduced d dimensional time series $\{X^*_{n,t}\}_{t\in\mathbb{N}_T}$ is then inputted to train a noise reduced HMM model of the data.

4.2.2 Model Specification and Generalisability

We use the HMM-MAR package to train HMMs with multivariate normal observations by Variational Bayes [408], whilst separating the data by subject into distinct trials of length T (the number of time points per subject). For further details on model fitting see [408]. Figure 4.2A shows how observations of the fMRI BOLD signal at each time point are modelled across subjects. Dynamics for each subject are modelled and fitted using a shared set of states $\mathcal{S}$ with finite $S = \{1, 2, ..., K\}$ and Markov transition matrix P.

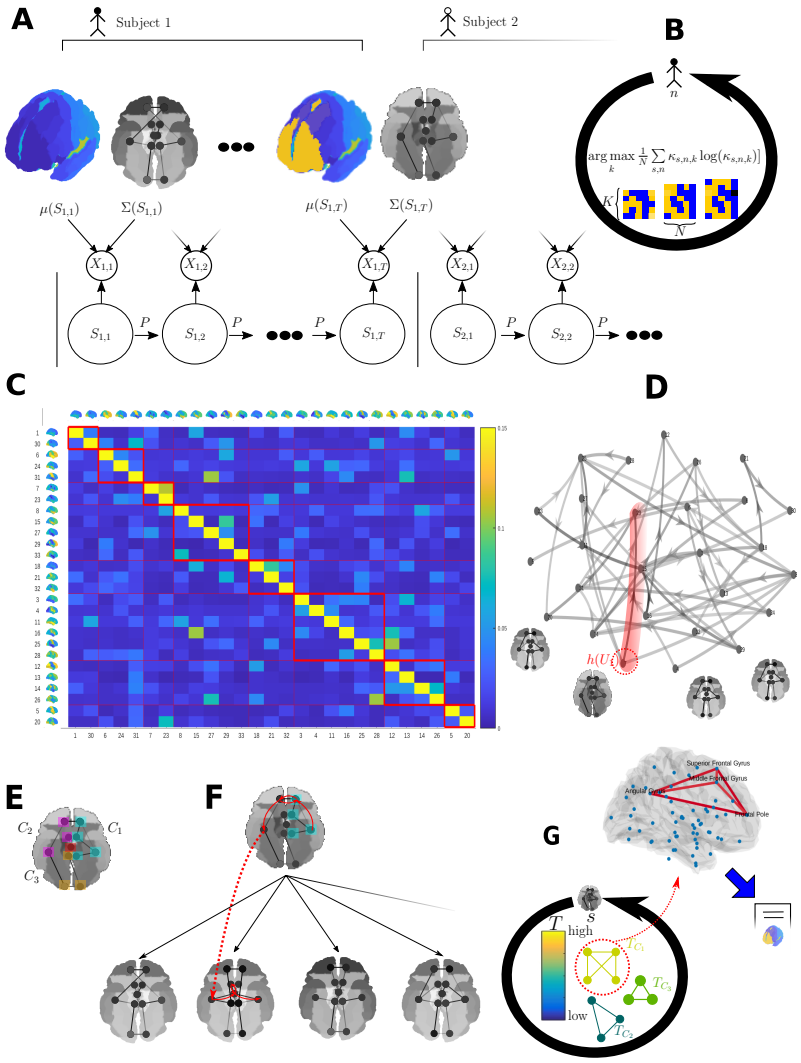


Figure 4.2: Flow diagram of the graph modelling and analysis pipeline. Following preprocessing of the fMRI data we obtain multivariate regional brain activity time series for all N subjects. Variational Bayes inference is then used to train HMMs (using the HMM-MAR package [408]). **(A)** is a sketch of an HMM fitted to the $X_{n,t}$ data for subject n and time point t. Each hidden brain state $S_{n,t}$ $\Sigma(S_{n,t})$ has mean activity $\mu(S_{n,t})$ and covariance $\Sigma(S_{n,t})$ (after backprojection). State change from $S_{n,t}$ is determined by the transition matrix P. **(B)** The number of hidden states, K, is determined using mean subjectwise cross-validated maximum entropy, which is calculated over the fractional occupancies, $\kappa_{s,n,k}$ for each subject-state pair up to K states. **(C)** Adjacency matrix of the interlayer temporal directed transition graph determined by the Markov transition matrix of the HMM, with temporal communities in red along the diagonal. **(D)** Each state itself can be considered a layer with edges relating brain regions by their correlation in activity derived from their modelled covariance $\Sigma(s)$, with node weights (regional mean activity) determined by $\mu(s)$. Of the states, some are highly connected state hubs, $h(U)$, belonging to a temporal community U (red shading). **(E)** Each hub state (layer) $h(U)$ is analysed and internal spatial communities are determined. **(F)** Internal communities are ranked according to their level of coherent brain activity compared to many repeated random walk samples from the multiplex model. **(G)** The results of ranking summarised by the community T-score. High T-score corresponds to a higher than expected level of community coherent activity when compared to the rest of the multiplex graph in this brain area. We propose functions for highly ranked communities by mapping these regions onto a 3D functional activity map and compared them to maps and terms drawn from the neuroscience literature with *NeuroSynth*.

We give a brief overview of HMM dynamics. We note that a key parameter, for these dynamics, the number of brain states, $K = |\mathcal{S}|$, that best generalises these dynamics across subjects is unknown. Consequently, we introduce a novel framework for selecting K based on an information theoretic criterion that maximises generalisability by maximising entropy of the state dynamics across subjects.

In each HMM state trajectory, the initial state of each subject's trial is selected independently at random. Under the Markov assumption of the model the resulting subject-specific state dynamics are assumed independent realisations of the same stochastic process, $S_{n,t}$. For $t > 1$, $S_{n,t}$ is conditionally dependent on the previous time step $S_{n,t-1}$ so that

$$Pr(S_{n,t} = s|S_{n,t-1} = s', \mathcal{M}) = P_{s,s'}, \tag{4.1}$$

for $s' \in \mathcal{S}$. Each brain state $s \in \mathcal{S}$ is associated with an observation model $O(s) \sim MVN(\mu^*(s), \Sigma^*(s))$. The $O(S_{n,t})$ model the dimensionally reduced brain data $X^*_{n,t}$ (row reduced by ROI). In order to obtain the full model, the reduced model is then back-projected into D dimensional brain region space (see Eq. (4.3)).

4.2.3 Novel Model Selection Criterion Based on Fractional Occupancy

The Markov chain defined by P and any given initial state $s_0 \in \mathcal{S}$, has a unique stationary distribution π_s that is independent of s_0 assuming the chain is irreducible and the states are positive recurrent. The probability π_s is the long run probability of the re-occurrence of state s. Selection of the number of these hidden states is carried out by cross-validated entropy maximisation over the related fractional occupancy distribution. The fractional occupancy distribution κ is defined by subject n for each state s and given by

$$\kappa(s, n|\mathcal{M}, X) = \frac{1}{T}\sum_{t=1}^{T} Pr(S_{n,t} = s|X, \mathcal{M}) \tag{4.2}$$

where $P(S_{n,t} = s|X, \mathcal{M})$ is the posterior probability of state s occurring at time t given the model $\mathcal{M}$ and data X. The fractional is the probability of finding subject n in s over the entire trial of length T. The distribution κ for subject n is related to the stationary distribution π_s by

the well-known limit

$$\kappa(s, n|\mathcal{M}, X) \xrightarrow[\infty]{T} \pi_s.$$

That is to say that κ asymptotically approximates the long run average state dynamics of the model as trial length increases. Knowing this, our goal is to select the model whose fractional occupancy maximises the entropy pooled across subjects by maximising the objective function

$$H(k) = -\sum_{n=1}^{N} \kappa(s, n|\mathcal{M}(n, k), X_n) \log[\kappa(s, n|\mathcal{M}(n, k), X_n)]$$

where the model $\mathcal{M}(n, k)$ is the model trained using all trials except the data from subject n assuming k hidden states, and X_n is the trial data from subject n (see Figure 4.2B).

By selecting the initial number of states $K = \arg\max H(k)$, we appeal to the information theoretic principle of maximum entropy which states that the model which maximises the uncertainty over the data tends to be the one that best approximates the true data distribution [192]. More specifically, our goal is to obtain a set of states with similar uncertainty about subject behaviour over the course of the experiment. We shall see in Section 4.3.2 that the goal of state-subject uncertainty maximisation relates closely to that of optimal model selection. We note that to the best of our knowledge this is the first application of such a subject-specific entropic criterion in state-space model selection.

4.2.4 The State Markov Information Graph

First model parameters $\mu^*(s)$ and $\Sigma^*(s)$ for state s from the HMM model $\mathcal{M}$ are backprojected using the transpose eigenmatrix A to obtain a model in D dimensional brain space so that the full D dimensional model has mean $\mu(s)$ and variance $\Sigma(s)$ defined over the ROIs and given by

$$\Sigma(s) = A\Sigma^*(s)A^T \qquad \text{and} \qquad \mu(s) = \mu^*(s)A^T. \tag{4.3}$$

Using the full model, each state s has normally distributed observations with mean $\mu(s)^x$ and covariance $\Sigma(s)^{x,y}$, for brain regions $x, y \in V$. We use these to define a graph $G(s) = (V, a(s), W(s))$ over the set of R brain regions, node weights $a(s)$ and edge weights $W(s)$, which we take to be a proxy for the information flow between regions. More specifically, we estimate the weights $W(s)$ by the correlation matrix $|\rho(s)|$, as derived from the state covariance matrix

$\Sigma(s)$.

Here, $a(s)^x = \mu(s)^x$ are the mean regional functional activity at brain region x in s. The weighted edge (directed information flow) from regions x to y are

$$W(s)^{x,y} = \frac{|\rho(s)^{x,y}|}{\sum_{y=1}^{D} |\rho(s)^{x,y}|}. \tag{4.4}$$

The resulting edge weights matrix $W(s)$, defines a Markov transition matrix, a model of information flow between brain regions in state s.

4.2.5 Louvain and Hierarchical Clustering at the Temporal Scale

We perform Louvain modularity detection on the directed Markov transition graph $G(P) = (\mathcal{S}, P)$, where $\mathcal{S}$ is the set of states (nodes) and the directed edge weights are given by the state Markov transition matrix P [56]. The Louvain algorithm involves the greedy optimisation of an objective function $Q(\mathcal{U})$, termed the modularity score for $\mathcal{U}$ a partition of $\mathcal{S}$ [151]. The algorithm allows for a resolution parameter γ which determines the relative size of communities and goes to one as $\gamma \to \infty$ [215].

The significance of the community partitioning is calculated by comparing $Q(\mathcal{U})$ to an empirical distribution composed of modularity scores from 10,000 partitions constructed by random permutation of the community labels. In addition, in order to examine the state-subject relationships directly, we perform agglomerative hierarchical linkage clustering based on correlation in fractional occupancy κ using Ward's method [421].

4.2.6 Community Hub Selection

State hubs are the states most central to the dynamics of the model and facilitate the switching dynamics within each community. These are selected by maximising the community centrality z-score, $z(s)$, for each community $U \subset \mathcal{S}$ [351]. This score measures the within community degree centrality of a node relative to the mean community connectivity (see Supplementary Information D.0.3). Hubs are then analysed for their community structure, using the same Louvain algorithm as in 4.2.5 but this time on the directed brain state graph $G(s)$.

4.2.7 Identifying Functionally Important Spatial Communities

Not all detected communities are as relevant to a state's functional role as others. Performance of these roles requires both functional activation and coordination of brain regions. To discern which communities are the most functionally cohesive, we rank communities by comparing to samples of regional activity from the full multiplex graph model (see Figure 4.2C and D). We used random walks to sample plausible patterns of functional network activity and employ them as a benchmark against which to measure the level of coordination within spatial communities. Controlling for the local level of background activity in space and time allows for a more representative indication of functional cohesion within brain networks identified by community detection than naive comparison of communities by community mean functional activity.

We introduce to neuroimaging the Functional Homogeneity, FH, as our community coherence measure, a statistic derived from the mean activity $\mu(s)$ and $\Sigma(s)$ that is high when the community mean activity is most in agreement with the directions of maximum community functional connectivity and low otherwise. It is a measure of the alignment between the two key features of spatial communities, their level of shared information and activation. This measure is well suited for neuroimaging data, and is well established in computer vision and image classification where it is known as the covariance metric and measures the agreement between and within image classes [229]. The FH for a community C in a state s is

$$FH(s, C) = \mu(s)^{C^T} \Sigma(s)^C \mu(s)^C, \tag{4.5}$$

where the superscript C refers to the submatrix given by removal of all rows and columns not corresponding to regions in community C. This metric is key to the community ranking procedure which follows a six step process:

1. Given a community $C \subset V$ in state $G(s)$ we calculate $FH(s, C)$.

2. Sample a state s' from the stationary distribution π.

3. Select a region $x \in C$ and sample $|C|$ nodes from $G(s')$ starting at $x \in V$ in $G(s')$.

4. Repeat steps 2 and 3 to construct a representative sample of paired states and brain regions $(s_1, C_1), (s_2, C_2)..., (s_L, C_L)$

5. Calculate the T-score for functional cohesiveness of a subgraph

$$T(s, C) = \frac{1}{L} \sum_{l=1}^{L} I[FH(s, C) > FH(s_l, C_l)]$$

where I is the standard indicator function and rank the communities in s by decreasing T-score.

6. Determine whether the community represents a correlated or anticorrelated brain subgraph by the sign of $E_C[\mu(s)] = \sum_C \mu(s)^C$.

The T-scores of all the communities in a specific state can then be used to order the states in terms of which are most likely to contribute to the functional cohesion of the state. Note that $T(s, C)$ is a score between zero and one, with one implying that the community C is much more functionally cohesive than other comparable brain subgraphs in space and time. T-scores are not designed to be compared across states. These steps are summarised by steps E to F in Figure 4.2.

4.2.8 Analysis of States and Communities with *NeuroSynth*

NeuroSynth is a meta-analysis tool that takes in 3D images of brain activity (termed functional activity maps) in MNI152 standard space and returns a scored association (based on the Pearson correlation) between the activity maps and other images from published articles that directly reference a given term i [436]. We choose the six terms most clearly associated with resting state activity *default mode*, *salience*, *executive*, these are the resting state network terms and *sensorimotor*, *auditory* and *visual*, sensory network terms. We used these to characterise the mean activity of a given state s by projecting the activity pattern $\mu(s)$ back into 3D brain standard space (see Supplementary Figure D.2A) and inputting the resulting map into *NeuroSynth*.

The resulting score for a state s and term i is denoted $\theta_{i,s} \in [-1, 1]$, with 1 indicating perfect correlation between the state's mean functional activity map and i and -1 indicating perfectly anticorrelated activity. We note that these terms, while chosen to relate to known resting state patterns, are not equivalent. Due to the common usage of these terms in the neuroimaging literature, they can however be considered to suggest a global pattern of activity or its absence (see Figure 1.2 of an image suggestive of the Default Mode Network). We explore the activity

of actual networks in our spatial community analysis Section 4.3.5.

We propose that the global score θ can also be considered a dynamically changing property of the system. Given a score $\theta_{i,s}$ for a term i and state s, the one step ahead predicted score is

$$E_{t+1}[\theta_{i,s}] = \sum_{s' \in \mathcal{S}} P_{s,s'} \theta_{i,s'}. \tag{4.6}$$

We use this predicted score to examine the global properties of the activity observed after reaching a given state.

NeuroSynth can also be used in conjunction with the newly developed package *NiMARE* to directly calculate the posterior probability of terms from a large corpus of neuroimaging journal abstracts and images given a selection of brain voxels in standard space [1]. Due to the variability in brain region size, regions selected by community membership are downsampled by selecting 10,000 voxels with replacement from each community which was found to produce stable posterior probabilities up to the third decimal place.

We use NeuroSynth with *NiMARE* to determine a plausible function for each of our spatial brain region communities, selecting only those terms that are most a posteriori probable and which had a functional rather than anatomical interpretation (see Supplementary Figure D.2B). We pass each community from each hub state through our spatiotemporal community ranking method resulting in a ranked list of communities of brain regions per state and then pass each top ranked community through the *NiMARE/NeuroSynth* method to determine their most likely functional term associations. In order to be comparable with the global score θ, the *NeuroSynth* score is either a positive or negative association depending on the mean activity of the regions as suggested in Section 4.2.7.

4.3 Results

Results for our multisubject HMM model training and multiplex graph model analysis are given below.

4.3.1 Dimensionality Reduction

We select the appropriate number of principal components using the method of parallel analysis outlined in Section 4.2.1. This resulted in a reduced set of $d = 9$ dimensions that account for

roughly 75% of the total variance, which are then used in fitting the model. The validity of this method is explored in Supplementary Figure D.4 which shows that community detection performance is detrimentally affected by too high or too low a choice of dimension d.

4.3.2 Entropy Relates to Model Selection

Applying our cross validated maximum entropy Hidden Markov Model selection criteria by maximising the cross-validated entropy $H(k)$, we obtain an HMM with $K = 33$ initial states. Figure 4.3 shows that the entropy maximum also coincides with the maximisation of the cross-validated Bayesian log-likelihood, which is a general indicator of model fit. To further reduce the risk of overfitting, we exclude those states that occur in less than 25% of subjects and renormalise P so that the rows again sum to one. The resulting model has a total of $K = 27$ brain states.

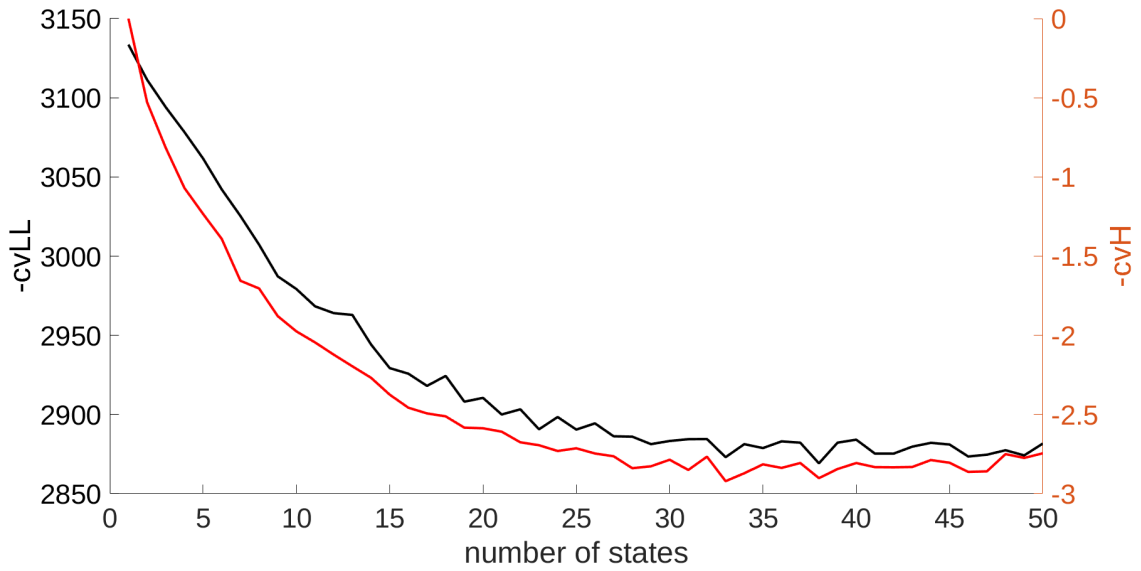


Figure 4.3: Selection of the number of hidden states by minimising the negative cross-validated entropy. The axes show the negative cross-validated log-likelihood −cvLL (left) and negative cross-validated entropy −cvH (right).. Qualitative similarities are evident between the two criteria suggesting deeper similarities between likelihood and entropy maximisation.

4.3.3 Network Dynamics Indicate Clustering of Spatiotemporal Activity

Table 4.1 shows that states positively correlated with resting state activity terms are significantly more likely to transition to states with similar associations and vice versa (see Supplementary Figure D.3 for linear model comparison). In contrast, states correlated with resting state terms

tended to transition to states that are negatively correlated with the sensory terms. This suggests that states associated with the former resting state networks tend to co-occur to the exclusion of sensory and sensorimotor patterns of activity. These results indicate a spatiotemporal separation between resting state network activity and sensory activity. The one-step ahead approach was chosen to examine close temporal relationships but higher time steps could reveal additional relationships. Given that many states have a high probability of self-transition, we examined these one-step ahead relationships when conditioned on transition to a different state and found no significant relationships after multiple testing correction. Conversely, significance is maintained when we consider the two step ahead relationships (given by substituting P^2 for P in Eq. 4.6). This may suggest that state persistence is an important part of maintaining these relationships.

States with high scores for sensory activity terms show a far weaker positive affinity for transition to each other than do the former resting state network terms. This suggests that concurrent activity in space and time is most likely between states with high resting state network activity. This pattern of concurrent activity is only weakly suggestive for sensory modes of activity. In contrast, robust mutually antagonistic spatiotemporal relationships between sensory and resting state network associations are present. We shall see in Section 4.3.5 this pattern of mutual exclusivity is mirrored by the most central states in the network or hub states at both the global (functional activity map) and the local (network community) levels. States show a general trend of transitioning from terms with one global activity association to another state that scores highly for the same association, suggesting some level of brain state inertia in the global pattern of functional activity.

	DM	S	E	SM	V	A
DM_{t+1}	0.9866 **	0.8201 **	0.4601 *	-0.4416 *	-0.3939 (*)	-0.5488 **
S_{t+1}	0.8253 **	0.9863 **	0.3550 (*)	-0.5526 **	-0.5302 *	-0.481 *
E_{t+1}	0.4519 *	0.3615 (*)	0.9826 **	-0.5203 *	-0.5241 *	-0.5753 **
SM_{t+1}	-0.4576 *	-0.5781 **	-0.5215 *	0.9878 **	0.1765	0.3128
V_{t+1}	-0.3718 (*)	-0.5085 *	-0.5092 *	0.1683	0.9885 **	0.1773
A_{t+1}	-0.4874 *	-0.4226 *	-0.5364 *	0.2657	0.1733	0.9842 **

Table 4.1: This table shows the relationships between *NeuroSynth* terms scores, calculated using the mean activity brain map for each state and the one step ahead projected score for each term according to the model (see Eq. (4.6)). Term scores for each state are correlated with the projected term scores one time step into the future (denoted by subscript $t+1$) from the current state (red is positive correlation, blue negative). False discovery rate corrected t-test significance is marked as ** ($p < 0.01$), * ($p < 0.05$) and (*) for marginal results ($p < 0.1$). The comparison between state scores for each term and the one step ahead predicted scores shows that there is a spatiotemporal relationship between resting state terms which are anticorrelated with sensory terms.

4.3.4 Evidence for Metatastate Structure in Wakefulness

In order to demonstrate the presence of temporal community structure, we performed hierarchical linkage clustering using the correlation in κ between subjects and states. Figure 4.4A suggests a temporally clustered pattern of state fractional occupancy in which certain states are more likely to co-occur in one subset of subjects than in the other. Figure 4.4B shows the transition probability matrix P organised into communities by Louvain community detection, where $\gamma = 0.48$ (as selected by Variation of Information minimisation) [215]. Temporal communities indicate modules of clustered state transitions. This temporal community partition was tested for robustness by comparing the Q modularity statistic to 10,000 random partitions with the same community labels (p =1e-4).

Each community, $U \subset \mathcal{S}$, is characterised by a hub state $h(U)$ determined by the state with the highest community degree z-score, a measure of state centrality to the temporal network (see Supplementary Figure D.1). Figure 4.4C, shows the long run probability of state s re-occurence π_s. Re-occurence and centrality to a community appear to be strongly correlated as states more central to their communities' according to the z-score, $z(s)$, also tended to have a higher stationary probability π_s, with correlation coefficient $\rho = 0.537$ ($p = 0.004$). As mediators of network dynamics, community hub states tend to re-occur.

4.3.5 Community Rankings Reveal Spatiotemporal Modules of Functional Activity

Louvain community detection was performed for each of the community hub state graphs $G(h(U))$, where we select the resolution as $\gamma = 2$. This was found to produce median spatial community network sizes that were sufficiently small on average for our community ranking method to efficiently sample the graph while also being large enough to detect functionally conserved brain state networks.

We perform *NeuroSynth* analysis by taking the mean functional activity maps generated for the hub states as input in combination with the resting state network terms *default mode*, *salience*, *executive* and the sensory network terms *sensorimotor*, *auditory* and *visual* (see Supplementary Figure D.2A for algorithmic explanation). The results in Table 4.2 suggests a separation between sensory and resting state activity in space and time with hub states scoring highly for either resting state or sensory terms but rarely both. Table 4.2 gives the highest ranked functional

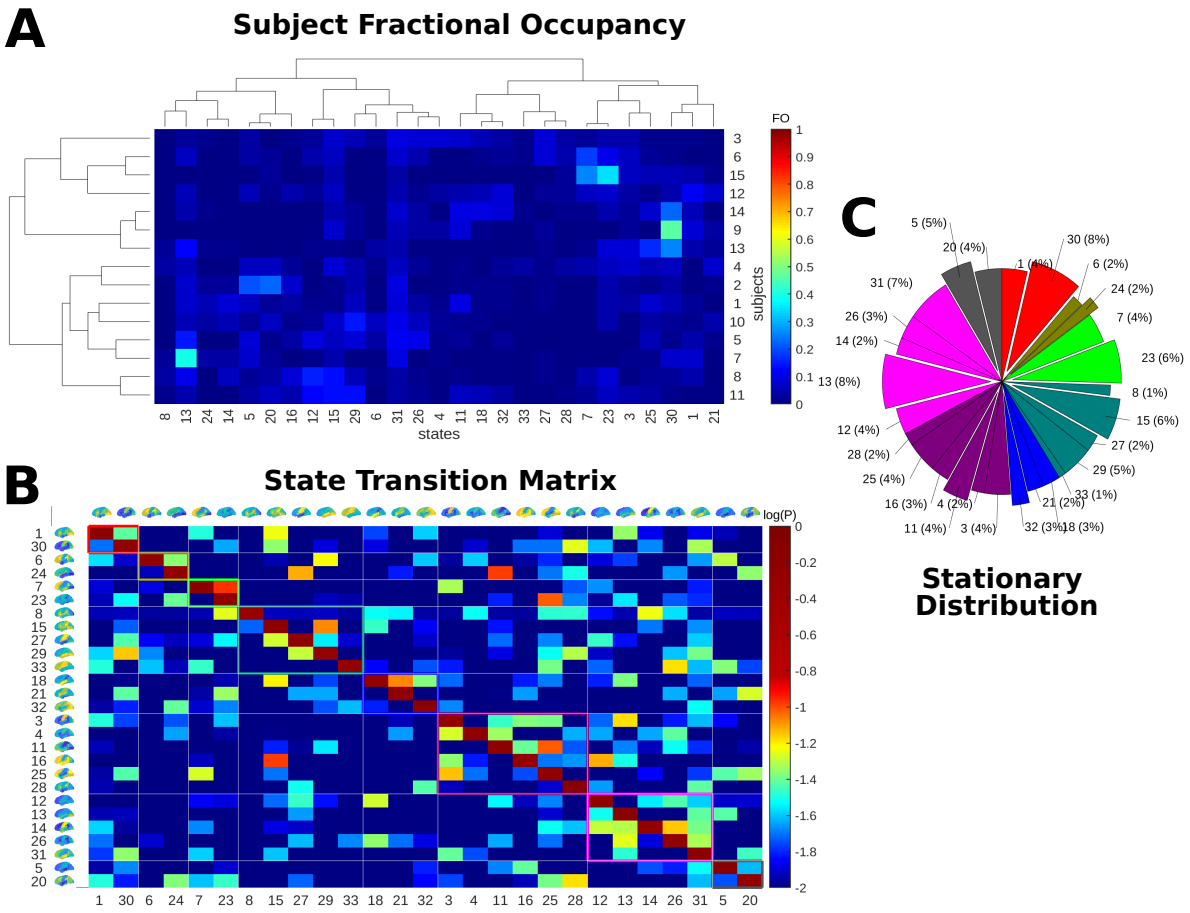


Figure 4.4: A summary of the subject state network dynamics. **(A)** Clustergram showing the relationships in Fractional Occupancy (FO), the proportion of time spent in a state clustered by subjects (vertically) and by states (horizontally). **(B)** The log of the state transition matrix P is shown, where states have been grouped along the diagonal, according to their community membership. **(C)** Pie chart showing the state occupancy at equilibrium (the probability of finding a subject in a state in the limit as time goes to infinity). Wedges in this pie chart are the individual hub states in each community according to the community z-score. These two scores share a significant 0.537 ($p = 0.004$), indicating the importance of community centrality to long run behaviour.

terms (filtering out purely anatomical terms) in each hub state for the top three ranked spatial network communities (using our ranking method). The top terms for each of the networks (communities) in the states largely coincide with the functional associations ascribed to each of the hub states themselves.

Exploring these relationships, we see that in some cases the connections between spatial community function and hubs are direct. State 23 shows a positive association with *observation* and *action* in dominant spatial communities and a strong association with all three sensory network terms. State 11 shows a clear association with auditory activity as well as a top ranked community association with the term *voice*. In state 15, which shows a strong correlation with

visual activity, the top ranked communities include positive associations with the *face* (a common object of visual processing).

In some states we see both strong positive and negative associations. Global negative asssociations are difficult to interpret in isolation as evidence of anticorrelated network behaviour within a state, however when paired with mesoscale information from the top ranked communities a stronger case is possible. State 32 appears mixed in activity but shows strong to moderate negative correlations with visual and auditory processing. The latter of these is corroborated by the anticorrelated speech network. State 30 is another state with mixed associations based purely on global functional activity, however, we see both moderate negative correlation globally with visual activity, and a specific negatively correlated community related to visual tasks or processing, suggesting a visual down state. A similar explanation can be used for state 5. State 23 is a sensory associated state with sensory associations at both the global and network scales. State 23 is negatively correlated with default mode activity. The default mode network is involved in language comprehension and reasoning, explaining the anticorrelated network associated with syntactic processing. Negatively associated communities may more generally suggest decreased metabolic or functional demand for these in networks leading to a coordinated down state.

term				**hub state**				
	13	11	15	32	23	5	24	30
DM	-0.0707	0.0528	-0.0094	0.1085	-0.3496	-0.0341	-0.095	-0.1918
S	-0.0924	0.0506	-0.1099	0.0714	-0.2039	0.0164	0.026	-0.0942
E	-0.2946	-0.1984	0.0534	0.1736	-0.085	0.1017	-0.0772	0.1024
SM	0.4245	-0.1793	-0.0535	0.1239	0.2502	-0.3186	0.2973	0.179
V	0.1544	-0.0055	0.4604	-0.536	0.2308	-0.289	-0.1644	-0.1168
A	0.1574	0.2959	-0.0946	-0.1542	0.2725	0.039	0.1534	0.0306
rank				**community**				
1	-reward	+voice	-incentive	-speaker	+action	-visual	-autobiographical	-basal
2	-theory of mind	+memory	+action	+autobiographical	+observation	+memory	-empathic	-memory
3	-language	-action	+face	+syntactic	-syntactic	+voice	-autonomic	-visual

Table 4.2: Summary of the *NeuroSynth* results for the hub states. The terms scored by *NeuroSynth* are the resting state terms *default mode* (DM), *salience* (S), *executive* (E) and sensory terms *sensorimotor* (SM), *visual* (V) and *auditory* (A). The first rows of the table under hub states show the *NeuroSynth* correlation score between each of the hub states' brain maps and the terms on the left (see Supplementary Figure D.2). The second section under terms shows the most probable terms associated with each of the top three communities identified by our ranking method, providing further information on the component functional communities of these states. The sign next to each term indicates whether the association is positive or negative (depending on the sign of the brain regions involved).

4.4 Discussion

In this paper we present a fully unsupervised pipeline for characterising the spatiotemporal activity of neuronal brain states in terms of a multiplex brain state graph model. This pipeline involves the training of an HMM in order to obtain a multiplex spatiotemporal directed brain state graph that represents the dynamics of subjects in resting wakefulness. We present a method for obtaining a set of states (layers) that generalises well over subjects and use this method to determine key states in the network dynamics. Lastly, we characterise the spatiotemporal components of the model that are most central and most functionally coherent, characterising these using metatextual image analysis of the neuroscience literature.

Our method reveals a rich array of complementary communities acting together to produce modes of neural behaviour during resting wakefulness. Crucially, we have shown that patterns of activity resembling the resting state networks tend to co-occur and that these patterns tend to preclude sensory and sensorimotor patterns of activity. This modularity of brain state function has been suggested by others [359, 409], but metaanalysis of terms associated with these functions allows us to characterise individual states and quantify their change in character through time.

Within each central hub brain state the division between functions was not clearly partitioned, with many terms featuring communities with *memory* or *autobiographical* associations, possibly suggesting an undercurrent of narrative thought which persists across numerous states. Alternatively, this may be due to artefacts caused by auditory memory-related tasks studies in the *NeuroSynth* database. It is also important to note that spatiotemporal state-based activity analysis is novel and so terms in the literature which derive from static models of activity may not map accurately onto dynamic patterns of activity. Notably, transient states may be smoothed out of these analyses meaning that new studies will need to be performed focusing on dynamic functional activity change at much shorter time scales. In addition, there are established biases towards known tasks or brain areas in the literature that may effect the distribution of metadata and results [195].

Some of the state global functional activity term associations, particularly negative ones, remain difficult to interpret. In state 13, there is a strong association with the term *sensorimotor*, however all of the top ranked communities for this state are negatively associated with functions

that may have a closer association to resting state activity. This could be due to putative link between the central executive activity and reward observed in primates [357], but may also be due to ranking error or noise in our graph model. However, the roles of many states become more clear when combining functional information from either anticorrelated or correlated mesoscale communities with global tendencies in functional activity. We hypothesize that strongly cohesive anticorrelated networks may be entering a coordinated down state due to changes in metabolic or functional demand [290, 382, 384].

In our future work we intend to investigate multilayer community detection approaches to look at dynamic changes in network membership. We plan to apply our multiplex analysis to conditions of altered consciousness in deep anaesthesia and determine novel spatiotemporal networks that characterise this condition with comparison to our current graph model for resting wakefulness. In this way we hope to elucidate the complex network dynamics underlying conscious brain activity [179].

Abbreviations

BOLD: Blood Oxygen Level Dependent signal

CEN: Central Executive Network

DMN: Default Mode Network

fMRI: functional MRI (Magnetic Resonance Imaging)

FO: Fractional Occupancy

PCA: Principal Component Analysis

HMM: Hidden Markov Model

ROI: Region of Interest

SM: Sensorimotor

SN: Salience Network

Availability of data and materials

Data, community ranking, and model selection code is available from the authors upon request.

Ethics approval and consent to participate

The study was approved by the Local Research Ethics Committee (Oxford Research Ethics Committee B, Oxford, UK) and data collection was performed between October and December 2009. The study was performed in line with the Declaration of Helsinki and all subjects gave written informed consent.

Competing interests

The authors declare that they have no competing interests.

Funding

JBW is supported by the Commonwealth Scholarship Commission UK and Ernest Oppenheimer Memorial Trust (South Africa), CEW is funded by MRC Development Pathway Funding Scheme (award reference MR/R006423/1), and GDR is partially supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grants EP/R018472/1 and EP/T018445/1. This research is funded in part by the Wellcome Trust (grant 203139/Z/16/Z). For the purpose of open access, the authors have applied a CC-BY public copyright license to any Author Accepted Manuscript version arising from this submission. Data collection was funded by the National Institute for Academic Anaesthesia, and the International Anaesthesia Research Society.

Authors' contributions

JW prepared the draft manuscript and developed the analysis methods. CW, CD and GR edited the manuscript. CW provided the raw data and interpretation of neuroscientific results. CD and GW supervised the analytical methods development. GR contributed to interpretation of model results. All authors read and approved the final manuscript.

Acknowledgements

We are grateful to the attendees and organisers of the Communities in Networks conference, where this work was originally presented, for the opportunity to contribute to this Special Issue. We are also grateful to Mark Woolrich, Angus Stevner, and the members of the Oxford

Anaesthesia Neuroimaging and Protein Informatics Groups, for their insightful questions and comments.

5 | Applying Hidden Markov Graph Models to Characterise Deep Anaesthesia

A contribution to our understanding of pharmacological unconsciousness

General anaesthesia is an invaluable surgical tool that reduces the risk of complications and shields patients from immediate pain while being generally safe and effective [394]. However, incorrect dosage and monitoring of anaesthesia risks psychological trauma through post-surgical recall of painful surgery, the feelings of extreme helplessness and anxiety that may occur, and distressing post-operative delirium that sometimes follows anaesthetic recovery [122, 338, 389]. These effects can be mitigated by monitoring of brain activity and preventing over anaesthetisation while minimising conscious awareness during surgery [146, 254]. Although many neurocorrelates of consciousness have been proposed [207], a quantifiable, computationally tractable or even agreed upon definition of consciousness remains elusive [80, 387]. By varying the effective concentration of anaesthetic, fine control of the subject's responsiveness, or more controversially level of consciousness, may be achieved [60]. Our study provides a first investigation into the changes in brain state, the combination of functional activity and connectivity, that differentiate deep anaesthesia induced unconsciousness from resting wakefulness.

Delirium affects nearly half of surgical patients, while conservative estimates of the incidence of intraoperative awareness in patients under anaesthesia are only around 2% the psychological trauma can be severe [122, 426]. Even so, this does not account for the potential hidden trauma as the amnesic effects of anaesthesia are still poorly understood [404]. EEG-based depth of anaesthesia monitors are available but they are not widely used by anaesthetists and the methods used to calculate the reported summary statistic are proprietary, making interpretation and

improvement difficult [162]. Previous EEG-based studies have revealed potential new electrophysiological measures of anaesthesia depth. These measures include anaesthesia-induced Slow Wave Activity (SWA), slow oscillations in the 0.5-2Hz band, which is reminiscent of activity also observed in deep sleep [271]. Slow Wave Saturation (SWAS) is defined as the dose at which power in the slow wave band reaches a maximum and plateaus [422]. Over-anaesthetisation beyond this point commonly leads to pathological burst suppressed activity, which is associated with protective activity after severe brain injury and post-operative delirium [142, 346]. SWAS has been put forward as a possible, interpretable alternative to proprietary measures of anaesthesia depth that tracks with measures of reduced consciousness such as neural complexity and entropy [101, 332].

Other markers of deep anaesthesia include high frontal alpha (8-12Hz) activity in EEG, thought to be generated by mechanisms deep in the brain (specifically in the thalamus). It is characteristic of deep anaesthesia but absent in deep sleep [410]. Altered thalamocortical connectivity has previously been observed in simultaneous fMRI recording and thalamic disconnection has been hypothesised to be an important part of the transition to anaesthetic induced unconsciousness [295]. Dose dependent cortical changes in response to pain and auditory stimuli are also observed in deep anaesthesia with the posterior prefrontal cortex retaining some level of activity, at moderate to low doses, in conjunction with the parietal lobe [262]. These areas are commonly thought to be involved in self-perception and awareness and may also remain partially active forming a possible conserved network in deep anaesthesia.

The so-called Resting State Networks (RSNs), notably the Default Mode Network (DMN) and Salience Network (SN), have also been shown to retain some activity at moderate anaesthetic concentrations, though with some loss of integration across brain regions as measured by the global clustering coefficient (see Section 2.4) [238]. This has been paired with a general decrease of the complexity of signals generated by the brain when anaesthetised. Signal complexity is of particular interest as it is thought to be a possible neural correlate of consciousness [58, 125, 252], though its accurate quantification in the context of real brain activity recordings remains difficult [344].

One particular measure, the Kolmogorov complexity, measures the algorithmic complexity required to generate a structure [210]. Algorithmic representation may be of particular inter-

est since conscious systems may construct generative models to make sense of their complex world [318], and many established neurocorrelates of consciousness actually approximate Kolmogorov complexity which is closely related to both entropy and compression measures of complexity [69]. In the special case of Markov chains, the sum of information required to encode that process' long run behaviour or Entropy Rate, is equivalent to the Kolmogrov complexity of that process [144]. This potentially means that entropy rate, if the system can be approximated by a Markov process, has advantages over the related Lempel-Ziv Complexity (LZC), a compression-based measure, as well as other measures which only approximate Kolmogrov complexity [144]. In addition, many of the existing measures are univariate and thus unable to capture the spatial multivariate integration of information that contributes to the complexity of brain activity. We present a novel method to measure information integration and complexity in the dynamic brain using the entropy rate of the underlying Markov chain. We believe this is the first time that the entropy rate has been proposed as a possible neurocorrelate of consciousness [344].

Cortical brain state dynamics under anaesthesia are well characterised by EEG studies in humans, but little is understood about the effects of anaesthesia during the transition to unconsciousness at the whole brain level, a transition that may be affected by white and grey matter distribution, age and other factors [262, 422]. In particular, as we age, our concentration of grey matter diminishes, which may have an effect on how we respond to anaesthesia (anaesthetic susceptibility), having deep clinical implications for correct dosage for patients. We investigate the possible connections between susceptibility, age and anaesthesia depth as measured by SWAS.

Electroencephalographic (EEG) and fMRI recordings allow for cross comparison of known brain state markers. We analyse data from a three phase experiment in which simultaneous EEG/fMRI was recorded during an initial resting wakefulness phase, a super slow propofol anaesthesia induction phase and a post-induction, stable anaesthesia phase under which deep anaesthesia was maintained at a fixed concentration [262]. This allows for the examination of changing brain state dynamics over a long period but we focus primarily on the two static periods, when propofol concentration is held constant before (at zero) and after induction. This was thought to be when activity would be at its most stable. Stability is important as models generally make assumptions of some kind of weak stationarity, thus it can be difficult to study dynamics in a non-stationary context where experimental conditions are changing and brain dynamics are

radically perturbed.

We develop the Hidden Markov Model (HMM) framework put forward in Chapter 4, where we developed the Hidden Markov Graph Model (HMGM) as a way of studying the relationship between regional and dynamic brain state activity. HMMs have previously been applied to conditions of consciousness including sleep (using fMRI) and most recently to Disorders of Consciousness (DOCs) with EEG [30, 369]. These studies demonstrate large scale changes in the brain state dynamics, notably reduced functional activity and connectivity in brain states during altered and reduced consciousness when compared to healthy resting wakefulness. However, any changes to the modular structure of brain state functional connectivity have not been adequately addressed, especially in anaesthesia induced unconsciousness.

Multiplex community structure under anaesthesia has recently been studied in light microscopy of neurons in the mouse sensorimotor cortex, where anaesthesia induced inhibition of neurons was found to fragment the communities [395]. In order to study this phenomenon at the whole brain level, we develop new methods to analyse community structure in HMGMs, using consensus community detection. We apply this first to the temporal state transition network as in Section 4.2.5, and then to the spatiotemporal multiplex network of regional activity changing probabilistically with time. Our new multi-resolution consensus community detection method is overall less vulnerable to the resolution limits of standard community detection methods and more robust to noise [136].

This chapter is broken up into three sections: a methods section, Section 5.1 that focuses on the new methods and data generated to analyse altered consciousness in anaesthesia, a results section, Section 5.2, detailing novel findings comparing resting wakefulness and anaesthesia as fully dynamic processes with multiple layers of functional activity. Lastly, in the discussion, Section 5.3, we integrate our findings and suggest possible states of reduced connectivity in deep anaesthesia, networks of retained neural activity, new neural and biological correlates of consciousness and anaesthesia susceptibility.

5.1 Materials and Methods

Chapter 4 presents a model of 15 wakeful subjects using an HMGM. Although we include both a model of subjects in deep anaesthesia as well as a model of subjects in the wakeful condition,

we focus mostly on conclusions drawn from the former and use the wakeful model largely for comparison. Supplementary Table E.1 shows which subjects were included in the analysis for this chapter.

In this section, we detail the broader experimental context in which the data from Chapter 4 was collected as this relates to certain markers of deep anaesthesia established in previous studies by our collaborators in [262, 422], involving ultra slow induction of propofol anaesthesia. Furthermore, we generalise the clustering procedures in Section 4.2.7 by finding a set of consensus communities over a range of resolution parameters in the transition network. We follow this analysis up with a spatial analysis of brain state regional multiplex communities that integrates temporal relationships in regional activity.

5.1.1 Data Acquisition

EEG and fMRI data was acquired over the course of a four phase experiment in which 16 volunteer subjects (8 female, 8 male) received a phase-specific dose of the GABA-ergic general anaesthetic propofol [262]. In the first phase, baseline wakeful activity was recorded for 10 minutes. Subjects were asked to remain still with eyes closed. In the second phase, propofol was gradually infused to a maximum Effective Concentration (ESC) of 4μg/ml, which is well above the minimal dose previously found to be effective in inducing loss of consciousness in subjects and is sufficient for inducing sleep-like slow waves (see Figure E.1) [133]. In phase three the maximum ESC was maintained for 10 minutes leading to a stable deep anaesthesia state. Lastly, a slow recovery period was recorded in which patients regained full responsiveness.

The entire protocol was repeated twice, once for EEG recording only and again for simultaneous EEG/fMRI recording in the fMRI scanner. Our analysis here focuses on the use of the 10 minutes before induction (phase 1) and the 10 minutes post induction recordings (phase 3), when subjects are held at high propofol anaesthetic dose of $4\mu g/ml$, resulting in a condition of deep anaesthesia and non-responsiveness to conventional rousing stimuli.

fMRI data was acquired at a frequency of one 3D volume scan of the head every three seconds at a 3D voxel resolution of 2mm^3. All brain volumes are artefact corrected and registered to a single 1mm^3, subject-specific high-resolution volume-scan in FSL, a freely available software package for fMRI pre-processing and analysis [193]. The resulting volumes are then mapped into fMRI

standard space MNI152 [257]. See Section 4.2.1 for more details. EEG was acquired on a 32 channel set-up (see Table E.2 and Figure E.2). P_{SWAS}, C_{SWAS} and T_{SWAS} measurements were acquired during induction in [262] and made at the F_Z electrode using the EEG only recording. EEG data was downsampled to 250Hz, notch and band filtered, see [262] for further details.

5.1.2 Subject Exclusion

The association of burst suppression with post-anaesthesia delirium suggests a non-standard pathway to unconsciousness and so is outside the scope of this study. Burst suppression is characterised by periods of low amplitude activity, alternating every few seconds with bursts of high frequency, high amplitude activity (see Supplementary Figure E.3 for example). We thus excluded subjects on the basis of pervasive burst suppression (>2 minutes) in either the EEG or EEG/fMRI recordings as evaluated by an expert (Marco Fabus). Recordings of subjects exhibiting brief periods of burst suppression (less than a minute in total duration) were not altered or removed. One subject was removed due to an incomplete recording, two more subjects were removed on the basis of extensive burst suppression. This resulted in a reduced set of 13 complete subject recordings (see Supplementary Table E.1) of ten minutes deeply anaesthetised and ten minutes awake which were used in the construction of new wakefulness and anaesthesia HMGMs.

5.1.3 Preprocessing, Model Selection and Training

Model selection and training for both anaesthesia and wakefulness is carried out as in Chapter 4 using the maximum entropy approach and the set of Harvard-Oxford parcellation regions V, with $D = 63$ total brain regions. As in Section 4.2.2, the number of initial states are selected by maximum entropy after dimensionality reduction using Parallel Analysis/Principal Component Analysis. States were then removed if they appeared in less than 20% of subjects and transition probabilities were then re-normalised. The Temporal Occupancy (TO) is defined as the posterior probability of a state at each point in time along a recording and is inferred from the data using the HMM model. Following pruning of non-general brain states, state trajectories, as defined by the Viterbi path (see Appendix C.1) and TO, were recalculated under these new transition relationships. Lastly, mean and covariance relationships were calculated by backprojecting into the original $D = 63$-dimensional Harvard-Oxford parcellation space, resulting in a model $\mathcal{M}$, with state set S, with number of states $|S| = K$, brain regions in V, mean $\mu(s)$ and covariance

$\Sigma(s)$ for each state $s \in S$ (see Supplementary Figure C.1).

5.1.4 Entropy Rate as a Measure of Dynamic Brain State Complexity

The Entropy Rate (ER) is a measure of the amount of information per transition that would be required on average to encode the long-run state dynamics of a Markov chain. Given a Markov transition matrix P that describes the brain state transition dynamics under a certain condition, the entropy rate is given by

$$ER = -\sum_{s,s' \in S} \pi_s P_{s,s'} \log(P_{s,s'}). \tag{5.1}$$

where π_s is the stationary distribution of P.

5.1.5 Measurement of Functional Connectivity

As a notable difference from Chapter 4, in this chapter, we use as a measure of functional connectivity, the conventional connectivity measure in fMRI studies (as opposed to the directed, Markov connectivity presented in Section Section 4.2.4) based on the Pearson correlation. This was chosen to simplify the interpretation and allow for comparison with previous results. More formally, we measure the functional connectivity $W(s)^{x,y} = |\rho_{x,y}(s)|$, the absolute value of the Pearson correlation coefficient calculated from $\sigma(s)$ for the regions $x, y \in V$. This is a more conventional but equally valid choice for $W(s)$ which is common across many studies of functional connectivity [240].

5.1.6 Consensus Louvain Community Detection

Temporal consensus community detection is performed over the state transition network $G = (S, P)$ (with states S and weighted adjacency P). This method improves on standard Louvain modularity detection as detailed in Section 2.2 by considering a plausible range of resolution parameters $\{\gamma_1, \gamma_2, ..., \gamma_g\} = \Gamma$ and integrating the results of multiple realisations of the stochastic Louvain algorithm. In Louvain modularity optimisation the same modularity function $Q(\mathbf{C})$ (see Eq. (2.7)) is optimised $R = 1000$ times resulting in a partition of S, $\mathbf{C}_{\gamma,r}$. The consensus

matrix

$$F_{s,s'} = \begin{cases} \frac{1}{Rg} \sum_{\gamma \in \Gamma} \sum_{r=1}^{R} \delta_{\mathbf{C}_{\gamma,r}}(s, s') & \text{if } s \neq s' \\ 0 & \text{otherwise} \end{cases},$$

depends on

$$\delta_{\mathbf{C}_{\gamma,r}}(s, s') = \begin{cases} & \text{if } \exists C \in \mathbf{C}_{\gamma,r} \text{ s.t. } s, s' \in C \\ & \text{otherwise,} \end{cases}$$

so that $F_{s,s'} \in [0, 1]$. The final consensus community is then obtained by performing Louvain community detection on the consensus matrix $F_{s,s'}$ with resolution $\gamma = 1$. The resulting partition is generally robust to repeated applications of the Louvain algorithm. We select the values of $\gamma \in [0.1, 1.5]$ in increments of 0.1. This range was chosen as for $\gamma > 1.5$, only the trivial partition was produced for all transition networks tested. This method is based on a simple method for consensus clustering in [216].

Multiplex Graph Model and Consensus Community Detection

In Chapter 4, we introduced the HMGM as a multiplex graph model with intralayer edges weighted by the transition probabilities P and interlayer edges determined by the functional connectivity $W(s)$ between brain regions in state s, $G(\mathcal{M}) = (S \times V, \mathcal{A}(W, P))$. The adjacency matrix $\mathcal{A}(W, P)$ is

$$\mathcal{A}(W, P)_{(s,x),(s',y)} = \begin{cases} P_{s,s'} & \text{if } x = y, s \neq s' \\ W(s)^{x,y} & \text{if } s = s' \\ 0 & \text{otherwise.} \end{cases}$$

For a partition $\mathbf{C}$ of $S \times V$, the multiplex consensus community detection framework repeatedly optimises the symmetric version of the multiplex modularity function

$$Q^*(\mathbf{C}) = \sum_{x,y \in V} \sum_{s,s' \in S} W(s)^{x,y} - \gamma \frac{d(s)^x d(s)^y}{\sum_{u,v \in V} W(s)^{u,v}} + \omega P'_{s,s'} \delta_{\mathbf{C}}((x, s), (y, s')),$$

where $P' = (P + P^T)/2$, and $d(s)^x$, $d(s)^y$ are the degrees of x and y in the functional connectivity matrix $W(s)$ and $\omega > 0$ is the temporal resolution which encourages shared community membership between regions in different states with strong community membership.

This definition of Q^* is a modified form of the standard multiplex modularity, which generally cannot take into account weighted or directed intralayer relationships. Such a constant term has been found not to be able to recover the many simple multiplex community structure [165]. The addition of the non-constant $P'_{s,s'}$ then is meant to allow for more complex intralayer relationships and thus potentially discover more complex spatiotemporal relationships. So far as we are aware, this is the first instance of leveraging the multiplex information embedded in multivariate HMMs to learn higher order community network structure from data.

The resulting multiplex community consensus matrix F^*, for $\gamma \in \Gamma = \{\gamma_1, .., \gamma_g\}$ and $\omega \in \Omega = \{\omega_1, ..., \omega_o\}$, has $DK \times DK$ entries given by

$$F^*_{(s,x),(s',y)} = \begin{cases} \frac{1}{Rgo} \sum_{\omega \in \Omega} \sum_{\gamma \in \Gamma} \sum_{r=1}^{R} \delta_{\mathbf{C}_{\gamma,\omega,r}}((s,x),(s',y)) & \text{if } s' \neq s \text{ or } x \neq y \\ 0 & \text{otherwise} \end{cases}$$

The resulting multiplex spatiotemporal consensus communities are then given by applying the standard Louvain algorithm to F^* with $\gamma^* = 1$.

We chose the temporal resolution parameters $\omega \in [0.01, 0.1, 1]$ and spatial resolution $\gamma = [1, 2, 3]$, with $R = 1000$.

5.1.7 Slow Wave Activity as an EEG Marker of Deep Anaesthesia

Slow Wave Activity (SWA) is characterised by an increase in slow wave or low δ activity (0.5-2Hz), as a proportion of total spectral power. This measure has been shown to correlate strongly with altered response to to nososensory (pain) and auditory (word) stimulation under increasing anaesthesia dose [262, 422]. SWA, as a possible measure of anaesthesia depth has been assessed in terms of its saturation at high doses. The saturation point can be measured in experiment and the timing T_{SWAS}, total power slow wave power P_{SWAS} and concentration at SWAS (C_{SWAS}) were all recorded in the first EEG experiment for each subject, which was published in [262] along with subject age and time of Loss Of Behavioural Responsiveness (LOBR), or the time at which the subject stops responding to stimuli.

5.2 Results

In this section we observe key differences in static activity between wakefulness and anaesthesia conditions. We then train and analyse two models using data from each condition, including model selection and community analysis methods derived in Chapter 4. The deep anaesthesia model $\mathcal{M}_{anaes} = (V, S_{anaes}, V, P_{anaes}, \mu_{anaes}, \Sigma_{anaes})$ which we compare to an awake model $\mathcal{M}_{wake} = (V, S_{wake}, P_{wake}, \mu_{wake}, \Sigma_{wake})$ of fMRI activity from the same subjects, with brain V (as defined by the Harvard-Oxford parcellation) and states $|S_{anaes}| = K_{anaes}$ and $|S_{wake}| = K_{wake}$.

5.2.1 Comparing Subject-wise Static Activity

Before performing normalisation or dimensionality reduction, we characterised static activity by comparing functional activity and connectivity across subjects in anaesthesia and wakefulness. On a subject level there is a high degree of correlation between mean functional activity in wakeful and anaesthetised conditions ($r = 0.8760$, $p < 0.001$) as measured by subject mean BOLD signal. This is contrasted against a much lower, but still significant, level of correlation at the level of functional connectivity ($r = 0.3908$, $p < 0.001$) as measured from region-wise absolute cross-correlation (see Section 5.1.5) in BOLD signal.

Figure 5.1 shows the static differences in BOLD signal activity between anaesthesia and wakefulness prior to normalisation or dimensionality reduction. This figure shows clearly that frontoparietal regions are overall most active in deep anaesthesia, occurring primarily in the nucleus accumbens, frontal, temporal and parietal regions. Overall mean thalamic activity in anaesthesia is lower than in wakefulness with the greatest decreases occurring in the occipital pole.

5.2.2 Selection of Models

We select the number of states for both models based on the maximum entropy of fractional occupancy (minimum negative entropy). This is complicated by the entropy for each condition scaling differently with the number of states between conditions (see Figure 5.2). We thus choose different numbers of states $K_{wake} = 33$ for wakefulness and $K_{anaes} = 19$ for anaesthesia. Note that we have been somewhat parsimonious in our choice of the optimal number of states for

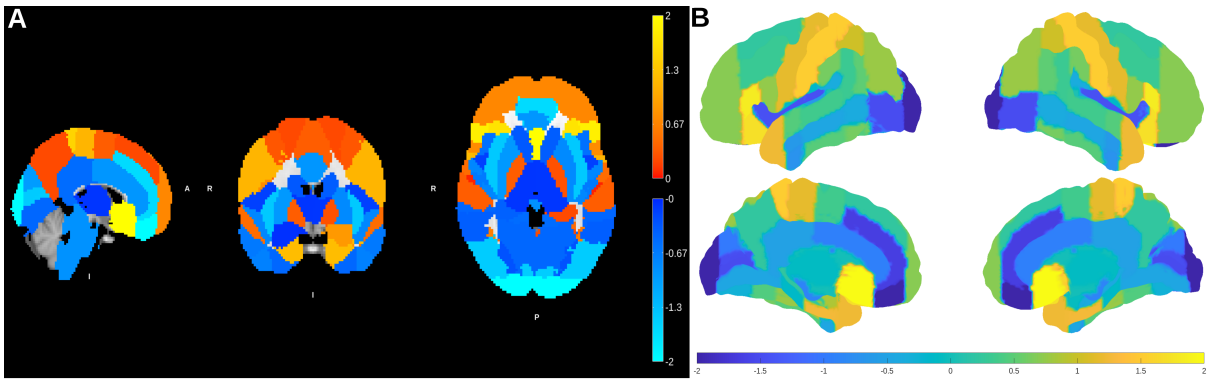


Figure 5.1: This figure shows the mean differences in static, regional functional activity between awake and anaesthetised conditions. This is calculated by taking the between subject differences in mean regional activity (normalised by mean and standard deviation). (**A**) This image shows a central anterior-posterior sagital cross-section (left), a right-left sagital slice (centre) and an anterior-posterior axial slice. Colour indicates activity above (red-yellow) or below (blue-cyan) wakeful activity. (**B**) Surface plot of functional activity differences for both left and right hemispheres. The top two images are of the right and left exterior surface and the bottom two are of medial activity (along the longitudinal fissure). Higher deep anaesthesia activity is clearly evident in frontal and temporal and parietal areas.

each condition. This choice is made difficult to determine from a limited set of subjects due to intersubject variability in anaesthesia response as determined by a range of psychological and physiological factors.

After exclusion of sporadic states that did not feature in more than 25% of subjects, the resulting models had $K_{wake} = 32$ and $K_{anaes} = 18$ states. In addition, we trained one combined model using both anaesthesia and wakeful data concatenated to produce single model with the same number of states as wakefulness $K_{comb} = 33$. This was done for validation purposes as training on such different data violates the ergodicity assumptions of the HMM (see Supplementary Information, Figure E.4). es.

5.2.3 Deep Anaesthesia Brain State Dynamics

Figure 5.3A shows the log transition matrices and dynamic state graphs for both anaesthesia and wakefulness. Note that the wakefulness model has a similar number of states and communities to the model in Chapter 4, while the anaesthesia model has a reduced number of states and communities by comparison with only two consensus communities. The single state, state 17, in the second community has a much higher long run probability of occurrence π_{17} (indicated by node size) than any other state in the community. This state also has the highest module degree z-score (see Appendix D.0.3), with $z(17) = 2.74$, which is much higher than the next most central state 12 in community one with $z(12) = 1.43$. These two states are the state hubs,

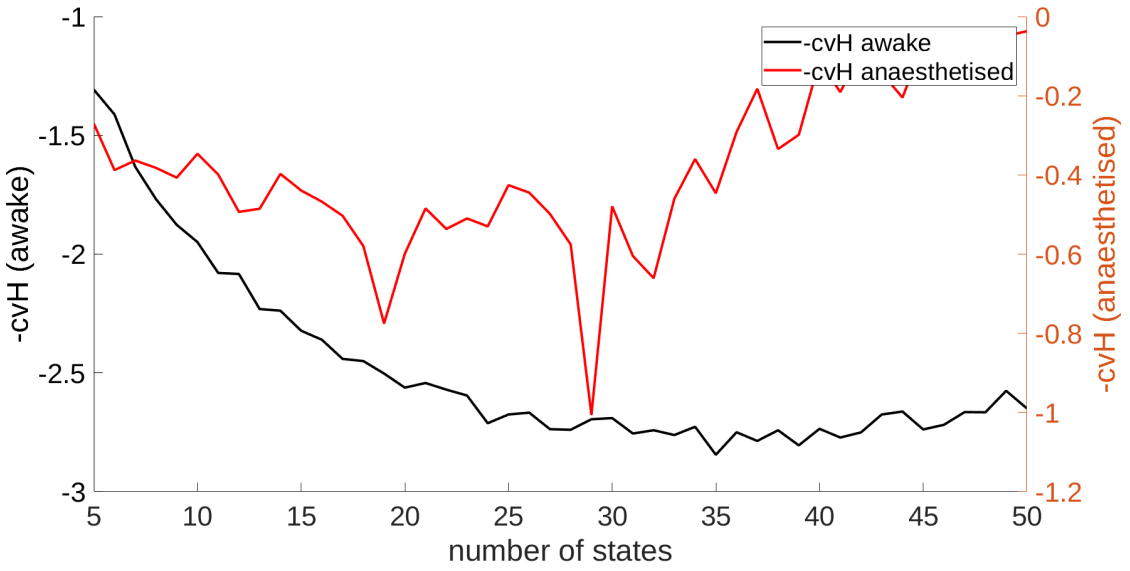


Figure 5.2: This plot shows the values of the negative cross-validated entropy (-cvH) selection criterion for both wakefulness (scale on left axis) and anaesthesia (scale on right axis). The two conditions show that entropy of fractional occupancy reaches a minimum roughly within the range 19 to 29 for anaesthesia and 33 to 40 for wakefulness.]

important states that are central to anaesthesia dynamics.

From the state dynamics graph 5.3B we see that these most central hub states have many high probability in-edges and few out-edges. In particular, state 17 is reachable with relatively high probability from states within its own community. It is unlikely to be transitioned to from external states, except state 13 which acts as a bridge from the other consensus community. Similarly, transitions to state 12 generally occur from within its own community but there are a few states in the green community that transition to state 12. Overall, the simpler modular structure of the anaesthesia model suggests lower complexity dynamics and indeed the anaesthesia model's entropy rate $ER(M_{anaes}) = 0.940$ is lower than the wakeful model with $ER(M_{wake}) = 1.82$. Supplementary Figure E.5 shows that this difference in complexity is not simply due to the difference in number of states as even when $K_{anaes} = K_{wake}$, $ER(M_{wake}) \leq ER(M_{anaes})$. Lower temporal complexity in deep anaesthesia has deep implications for the capacity for complex cognitive processes and conscious awareness under anaesthesia.

While there is some overlap in Fractional Occupancy (FO), as defined in Eq. (4.2), between subjects, Figure 5.4 shows that the two hub states separate subjects into two groups depending on which one is dominant. The majority of subjects are dominated by state 17, the remainder have a high state 12 occupancy. While it is not clear why subjects 10 and 4, are primarily

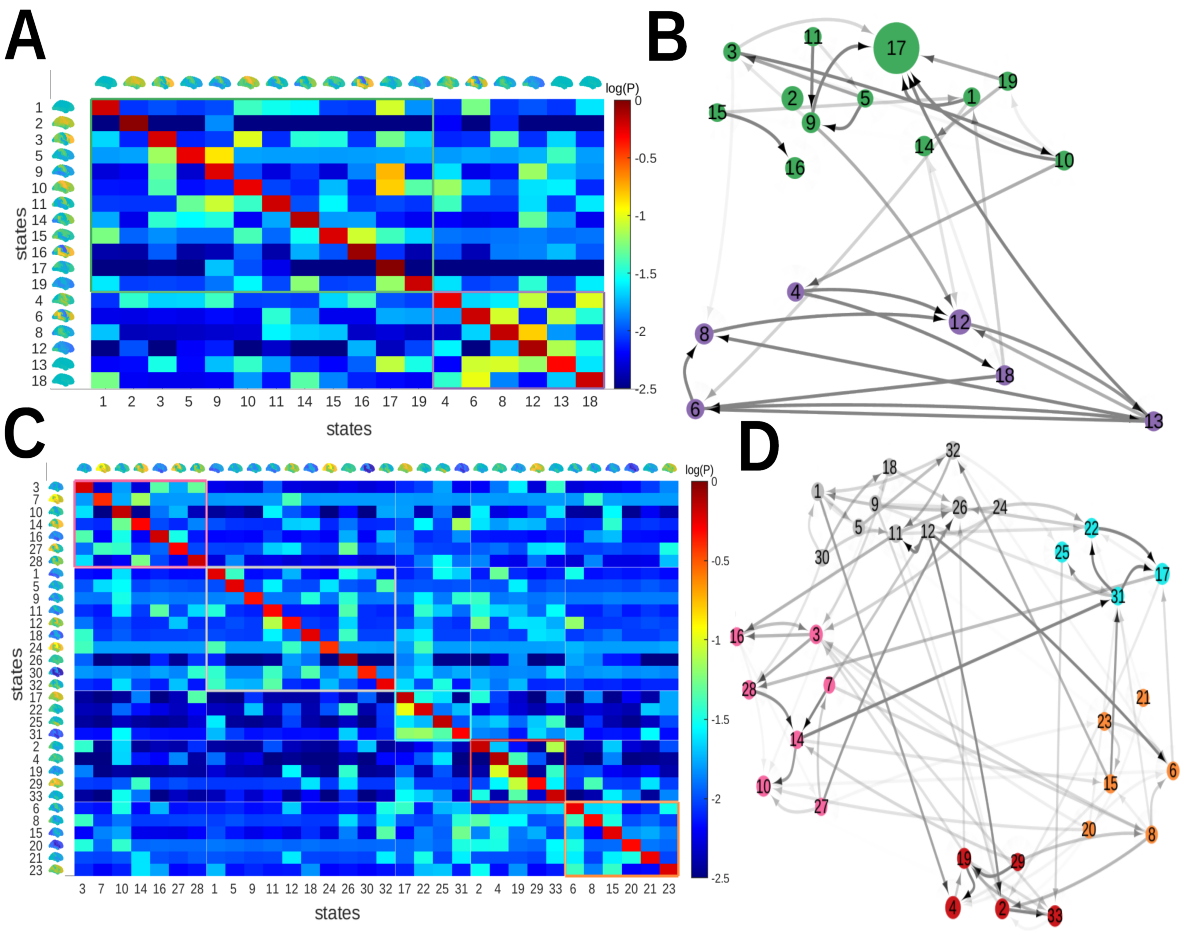


Figure 5.3: This figure presents the state dynamics graph and community composition of the two models in anaesthesia (top) and wakefulness (bottom). **(A)** Shows the state log transition probabilities (base 10) in anaesthesia as well as the right hemisphere activity for each state. **(B)** Shows the state dynamics graph (interlayer activity) in anaesthesia with edge transparency determined by the transition probability and node size determined by the stationary distribution of the Markov transition matrix P_{anaes}, and nodes coloured by their community membership. **(C)** As in **(A)** this figure shows the log transition probabilities and community memberships in wakefulness. **(D)** Shows the state dynamics graph for the reduced wakefulness model.

dominated by non-hub states, this may be due to the fact that the two subjects are the youngest in the study, where age is known to modulate the effect of propofol on cognitive processing in anaesthesia [262]. Together states 12 and 17 account for $> 50\%$ of the mean Fractional Occupancy (FO) across all subjects in deep anaesthesia, indicating that a majority of subject time is spent in one of these two states. This is in comparison to wakefulness in which no two states can account for more than 20% of average FO. These observations are true also for the validation combined and 20 state anaesthesia models. We see a few states emerge to dominate anaesthesia both the 20 state and in the combined model, with a clear separation between wakeful and anaesthetised subject trials emerging in the combined model (see Supplementary Figures E.6A

and E.4A).

Figure 5.5 shows a statistical breakdown of the state switching dynamics. Figure 5.5A and B

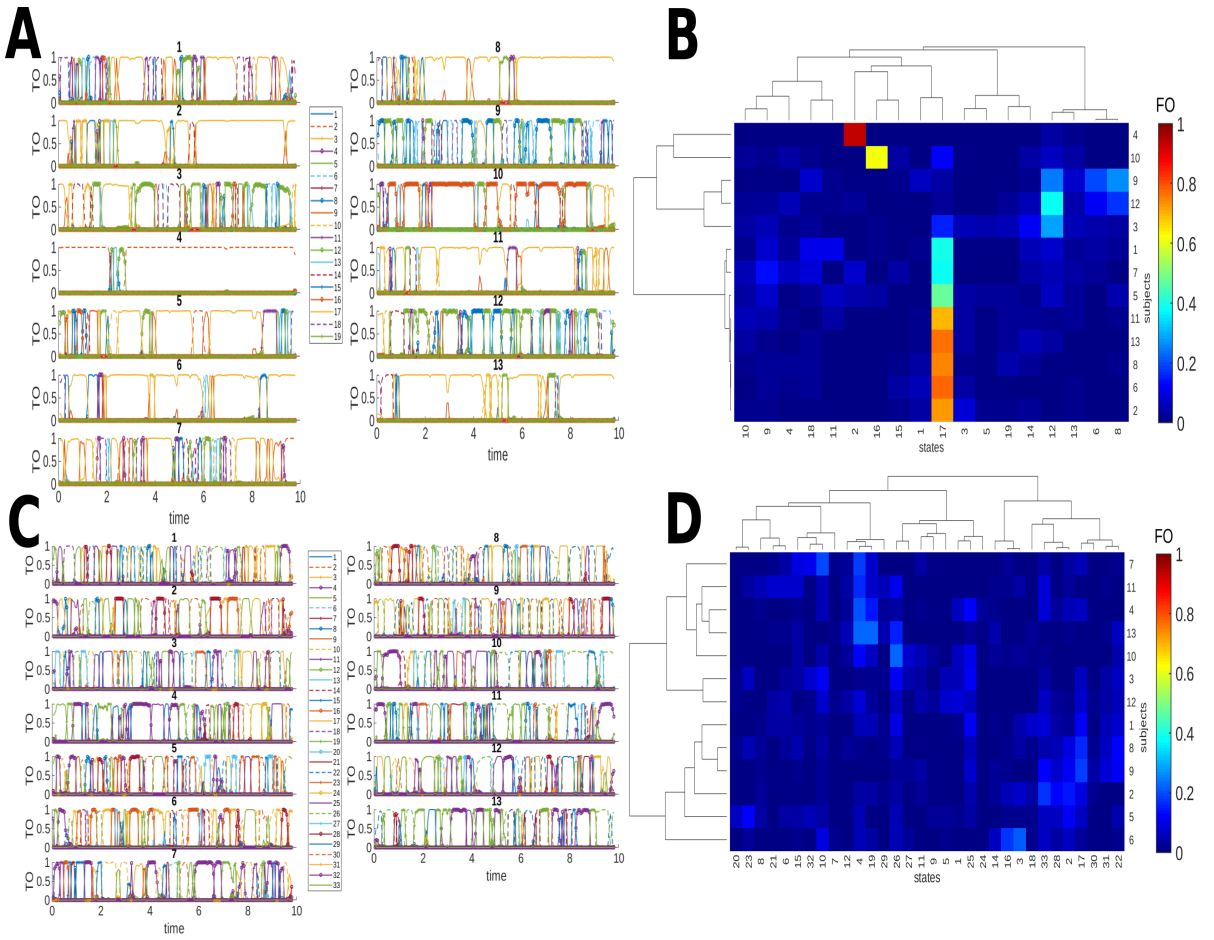


Figure 5.4: This figure summarises the subject state dynamics graph and community composition of the two models in anaesthetised (top) and wakeful (bottom) subjects. **(A)** Shows the probabilistic state trajectories of each state during anaesthesia recordings for each subject numbered one to 13. Probabilistic trajectories are determined by the Temporal Occupancy (TO) of each state, the posterior probability of a subject occupying a state at each point in time given the model and data. Hub state 17 dominates in many of these subjects' recordings. **(B)** Shows the state Fractional Occupancy (FO) hierarchically clustered by subject and by state using correlation with Ward's algorithm as in Section 4.2.5. The majority of anaesthetised subjects have very similar dynamics characterised by a large FO in state 17, while the other branch of the dendrogram spends a large proportion of time in state 12. **(C)** As in **(A)** this figure shows the posterior state probabilities, TO, over the ten minute wakeful recording. **(D)** Shows FO clustered as in the anaesthesia model.

show that while all states have a median lifetime below twenty seconds in wakefulness, as do most in anaesthesia, a small minority of states in anaesthesia have median lifetimes exceeding 20 or even 30 seconds (in the case of state 17). State 2 in Figure 5.5A has a large interquartile range, with an upper quartile of 126 seconds and is the state that dominates activity in subject 4. These long lived states are observed on a subject level as a lower rate of dynamic switching between states. When comparing across conditions and models, subjects in anaesthesia had a

significantly lower rate of state switching than in wakefulness ($p < 0.001$, Wilcoxon Signed-rank test) and a lower state diversity, as measured by the normalised entropy of FO ($p < 0.001$,, Wilcoxon Signed-rank test). Together these results suggest that dynamics under deep anaesthesia are not only simpler but less dynamic with a lower fraction of states participating in the dynamics of most subjects. However, there is also some evidence that subjects are more variable, likely due to the individual subject's depth of anaesthesia, which varies from subject to subject at the same ESC.

Dynamic BOLD and state activity under anaesthesia for a single subject example (subject 5) is shown in Figure 5.6A and wakeful activity in the same subject is shown for comparison in Figure 5.6B. BOLD signal activity is shown for two representative brain regions (see Figure 5.6C). The state dynamics shown here illustrate how activity under anaesthesia tends to include longer periods of apparent depression in activity (periods of lower than average mean activity) dominated by a single state (here state 17), interspersed with sporadic state changes and increases in activity. This is contrasted against wakeful activity which is more dynamic with greater state and activity pattern changes over the same time period. No easily discernible differences are seen in the representative cortical and subcortical time series either between states or within conditions. The lack of easily discernible features differentiating cortical and subcortical activity further motivates the use of a region-specific characterisation to tease out these differences.

5.2.4 Hub State Functional Activity Under Anaesthesia

Hub state 12 and 17 functional activity is shown in Figure 5.7. Activity across most regions in both states is below the mean activity in anaesthesia baseline. However, in state 12, certain regions are marginally above baseline activity. These include thalamic, occipital and dorsal prefrontal areas (see Figure 5.7A), as well as supplementary somatosensory and posterior prefrontal areas that are visible in Figure 5.7B. Activity is even more subdued in state 17, where no region is above mean activity. Activity is most depressed in dorsal prefrontal areas as well as brain stem and thalamus (see Figure 5.7C), and closest to baseline in the basal ganglia, temporal, parietal and supplementary motor areas. This is most evident in Figure 5.7B and D where surface activity seems superficially similar across the two states, particularly in sensorimotor and temporal areas but can also be seen in the parietal areas of A and C.

The two states are weakly correlated in spatial activity with a Pearson correlation of $r = 0.2136$

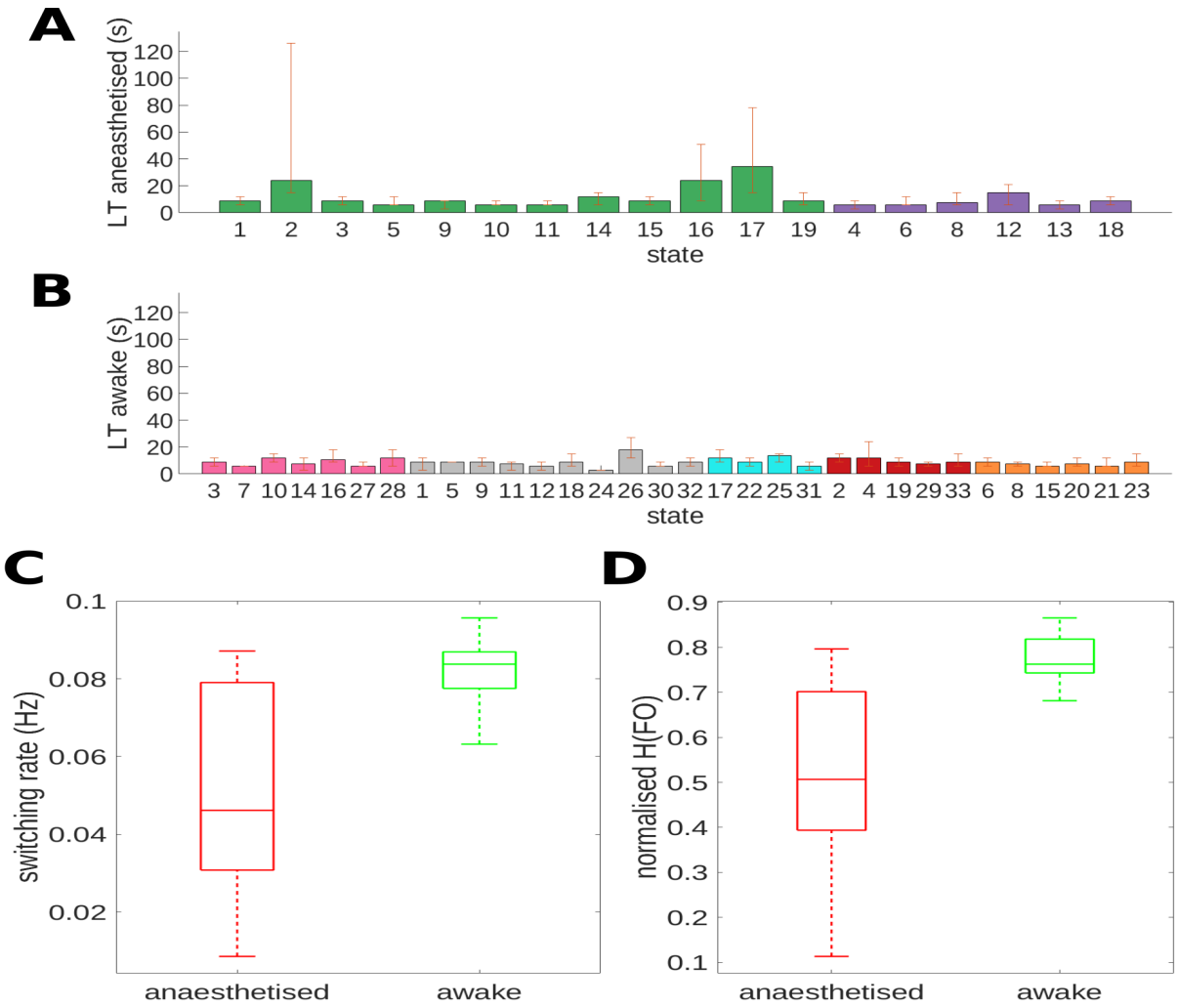


Figure 5.5: This figure summarizes the temporal statistics of states across subjects. **(A)** This bar plot shows the median Life Times (LT) of states in seconds (calculated from the Viterbi path), with error bars given by the interquartile range and colours determined by the state's consensus community membership. The interquartile range was chosen over the standard deviation due to the non-normality of life time data. **(B)** Shows the same state life time plot for states in the wakefulness HMGM.**(C)** These boxplots show the state switching rates (Hz) for subjects in anaesthesia and wakefulness, with the wakeful rate being significantly higher than in deep anaesthesia ($p < 0.001$, Wilcoxon Signed-rank test). **(C)** These boxplots show the normalised FO entropy (H(FO)), which is the entropy of the subjectwise fractional occupancy divided by the information length $\log(K)$ (for K the number of states in the model). Again, wakeful state diversity is significantly higher than in deep anaesthesia ($p < 0.001$, Wilcoxon Signed-rank test).

across brain voxels. There is no evidence of direct linear correlation between anaesthesia hub state occupancy and wakeful dynamics, either in terms of switching rate ($r = 0.1028$, $p = 0.7383$) or entropy of state occupancy ($r = -0.0123$, $p = 0.9682$).

5.2.5 Stratification of Functional Connectivity in Anaesthesia

Spatiotemporal consensus clustering combining functional and temporal relationships (see Figure 5.8), reveals strikingly different results when applied to anaesthesia and wakefulness graph

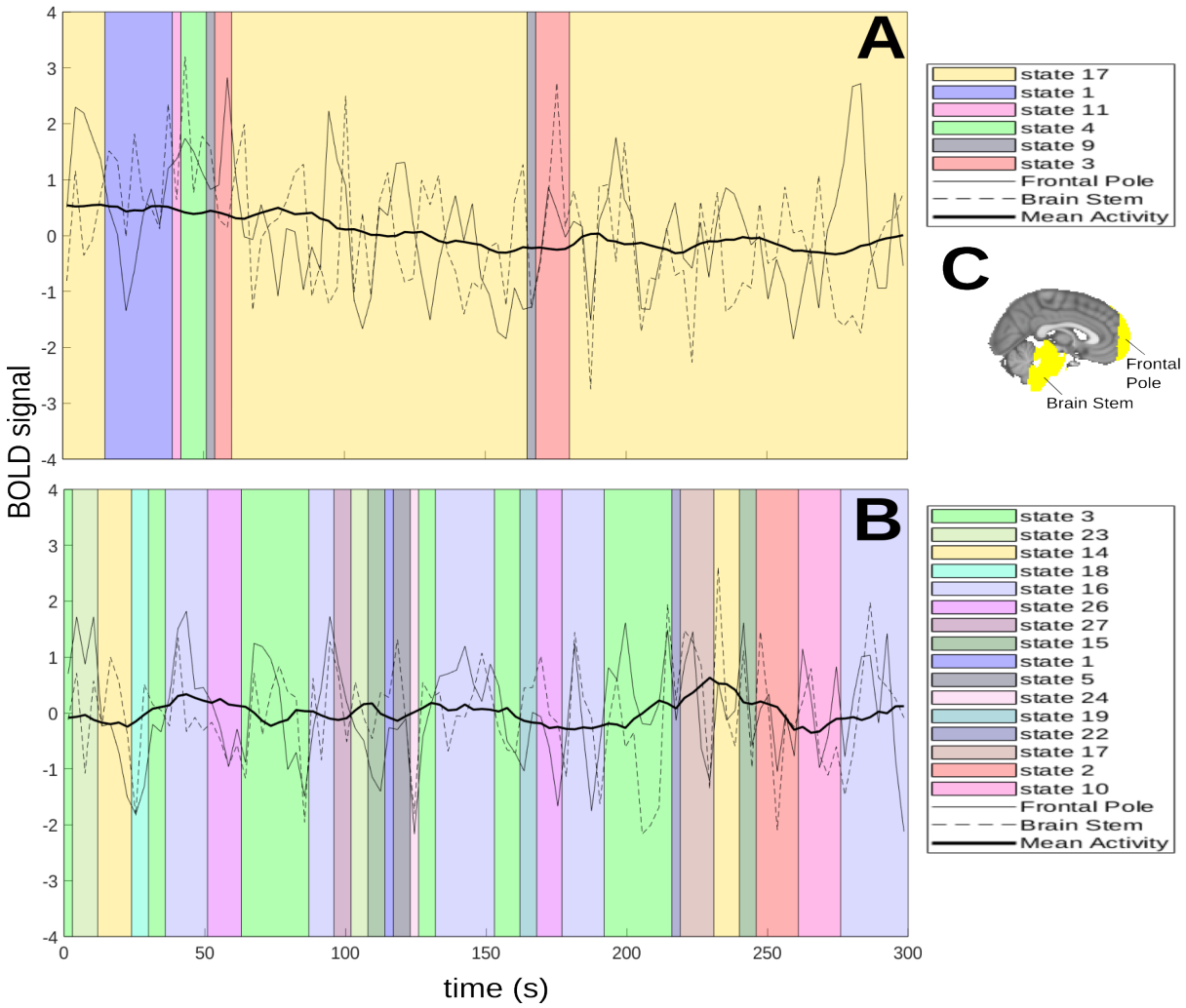


Figure 5.6: This plot shows BOLD and state activity time series from mean and representative cortical (frontal pole), subcortical (brain stem) regions during a 5 minute recording from a representative subject (subject 5) under both anaesthesia and wakeful conditions. The most likely (Viterbi) state trajectory is shown in shaded segments. The horizontal length of the segment indicates the expected time spent in that state. (**A**) This plot shows activity under anaesthesia. Here, state activity is less diverse when compared to wakefulness and includes long segments of persistent activity, particularly in state 17. (**B**) This plot shows state and BOLD activity in wakefulness which is characterised by similarly short intervals across a large proportion of wakeful states. (**C**) This image shows a medial view of the brain with labels for the brain stem and frontal pole as defined by the combined cortical-subcortical HO atlas.

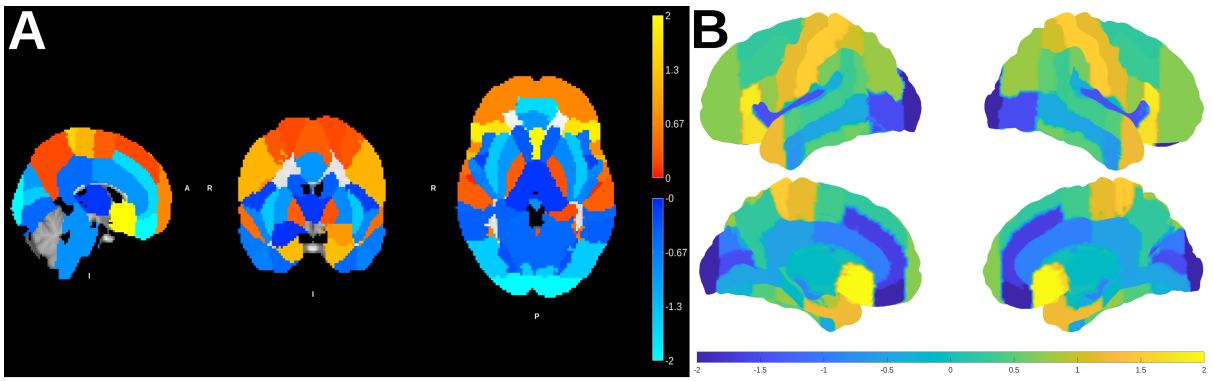


Figure 5.7: This figure shows functional activity brain maps in each of the two central hub states, state 12 and 17. **(A)** This image shows a central anterior-posterior sagital cross-section (left), a right-left sagital slice (centre) and an anterior-posterior axial slice. Colour indicates activity above (red-yellow) or below (blue-cyan) average activity baseline. Activity in state 12 are generally below baseline with some positive activity in the dorsal frontal, thalamic and somatosensory areas. **(B)** Surface plot of functional activity in state 12 for both left and right hemispheres. Activity is clearly visible in supplementary motor, occipital and posterior prefrontal areas. **(C)** Cross-section of functional activity maps in state 17 show a global activity decrease in this state. The top two images are of the right and left exterior surface and the bottom two are of medial activity (along the longitudinal fissure). **(D)** Similar surface plot showing functional activity in hub state 17. Functional activity is slightly up in basal ganglia, temporal, parietal and supplementary motor areas but still lower than baseline. The voxel-wise spatial (Pearson) correlation between states 12 and 17 is 0.2136.

models. In the spatiotemporal case there is far less changing of community membership of brain regions in anaesthesia than in wakefulness. This is reflective of the temporal dynamics of the anaesthetised in which we see far less switching between states. Even though we do see less switching in anaesthesia dynamics, sporadic co-recruitment of regions commonly associated with the so-called Resting State Networks (RSNs, see Appendix A.1), does occur for example in state 16 (exhibiting Default Mode-like activity, community 3).

Furthermore, the consensus clustering suggests that subcortical areas, most notably the thalamus and neighbouring regions (including the caudate), generally thought of as an information bottleneck for sensory inputs reaching the brain [15], are isolated from cortical brain regions in almost all states in anaesthesia. This is not seen in wakefulness in which thalamic regions share and switch membership with multiple cortical regions depending on state suggesting active involvement of the thalamus in brain state activity.

The Average Local Clustering (ALC) coefficient (see Eq. (2.14)) is a common measure of efficiency of communication and integration of information within the brain [238]. Figure 5.9A shows that greater stratification of regional brain activity in deep anaesthesia may in part be due to a lower level co-clustering (ALC) in the brain state functional connectivity networks of

the anaesthetised when compared to wakefulness (p <1e-5, Mann-Whitney U Test). In particular, state 17 exhibits the lowest ALC of all states in anaesthesia or wakefulness while state 12 exhibits the next lowest, suggesting that subjects are drawn to states of minimal information integration in deep anaesthesia.

Figure 5.9B shows that in general deep anaesthesia brain states have an absolute activity (mean of the absolute value of expected brain activity across regions) that is far lower than states in wakefulness ($p < 0.001$, Mann-Whitney U Test). It should be noted that since the data has been normalised, the result shows the degree of deviation from an anaesthesia and wakeful static baseline (as presented in Section 5.2.1). Overall, depression of both connectivity and activity is evident during deep anaesthesia, driven in part by hub states.

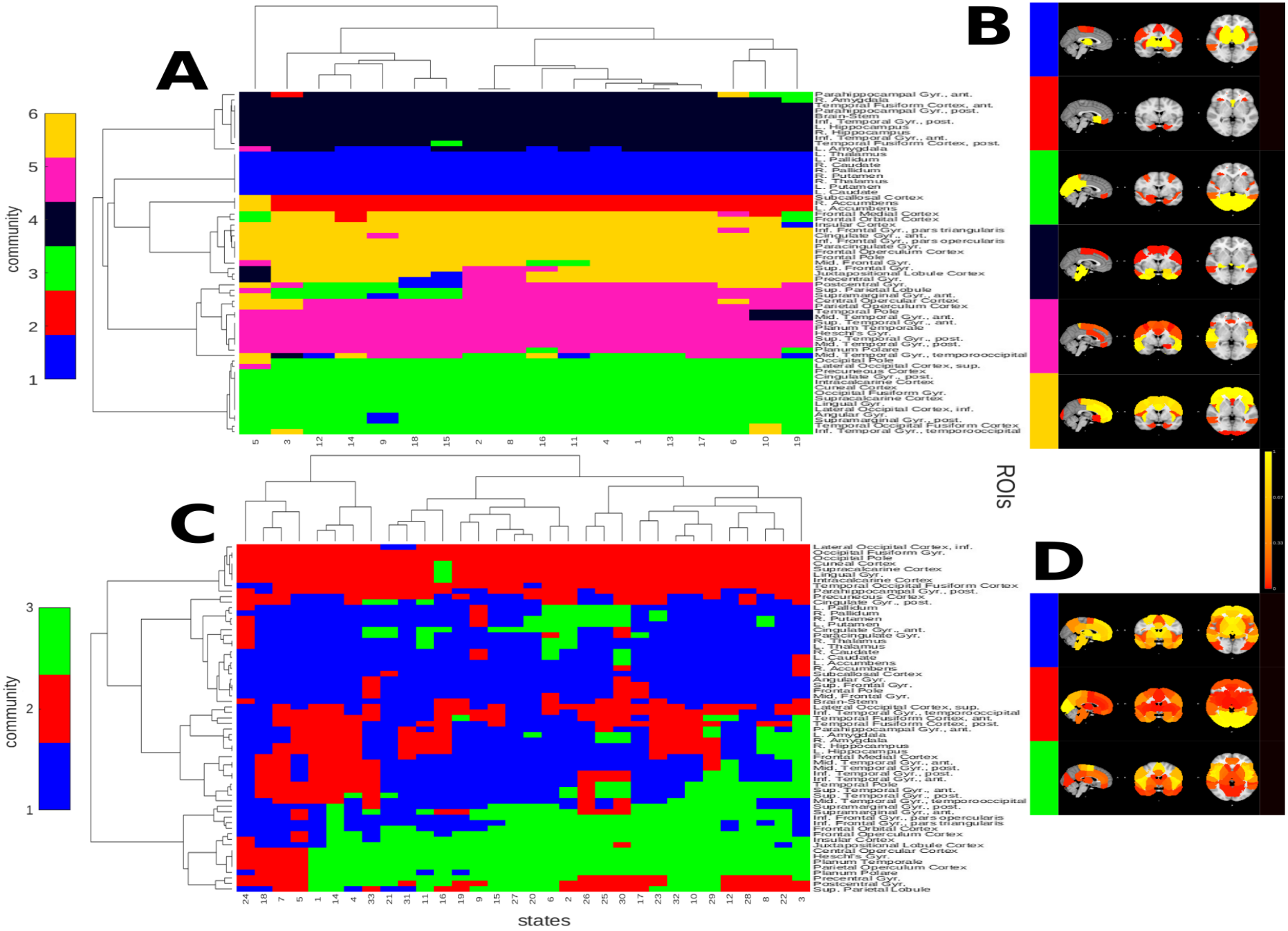


Figure 5.8: This figure summarises the results of multiplex consensus community detection across states and regions (ROIs) with hierarchical clustering over the spatial resolution $\gamma = [1, 2, 3]$ and $\omega = [0.01, 0.1, 1]$ for both deep anaesthesia and wakefulness. Hierarchical clustering was performed using Ward's algorithm, with Jaccard index similarity. **(A)** Shows the results of community detection for the deep anaesthesia condition. Brain regions are largely static but separated into a larger number of communities than in wakefulness. **(B)** Plots of brain regions involved in each anaesthesia community. **(C)** The clustergram of wakeful communities show a much lower level of community consistency than in anaesthesia. **(D)** Regional activity is far more general in wakefulness. The colour bar indicates the fraction of states which include a specific region in a community.

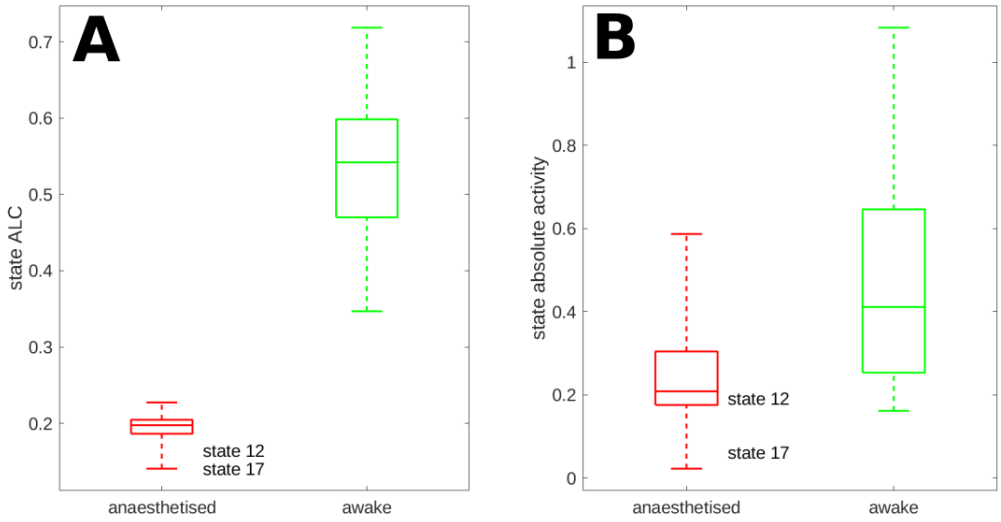


Figure 5.9: These boxplots provides a summary of mean connectivity efficiency and activity levels of the brain states during deep anaesthesia when compared to wakefulness. **(A)** Average Local Clustering (ALC) coefficient as a measure of network efficiency of each brain state (with labels for the two hub states). A significantly higher efficiency for brain states in wakefulness when compared to anaesthesia is evident (p <1e-5, Mann-Whitney U), with hub state 17, followed by 12, having the lowest ALCs of all states in anaesthesia. **(B)** Boxplots comparing the mean level of absolute activity (absolute value of mean regional activity) between deep anaesthesia and wakefulness. Again, the mean activity of anaesthesia states deviants significantly less from baseline activity than in wakefulness ($p < 0.001$, Mann-Whitney U Test).

5.2.6 Hub State Functional Connectivity and Homogeneity

Plots of the functional connectome of each of these states in Figure 5.10. This plot shows the top 5% strongest (by magnitude) functional connections between brain regions. Few strong functional relationships outside of the brain region spatiotemporal communities are present as suggested by the low level of neural integration (see Section 5.2.5), of these strong connections between regions of the parietal lobe as well as somatosensory regions.

Using the Functional Homogeneity (FH) measure, that takes into account local spatiotemporal brain activity outlined in Section 4.2.7, we see that the top communities by T-score are communities 2, 6 and 1 in state 17, and communities 3, 5 and 6 in state 12. Both states share conserved activity in community 6 which is of prefrontal, insular and cingulate regions involved in interoception. On the other hand state 17 shows strong conservation of community 1, a subcortical community composed of the thalamus and neighbouring structures, with functional connections to the brain stem and parietal lobule in this state. The thalamus is a bottleneck for sensory inputs and outputs and its disconnection from cortical brain regions as occurs in both states,

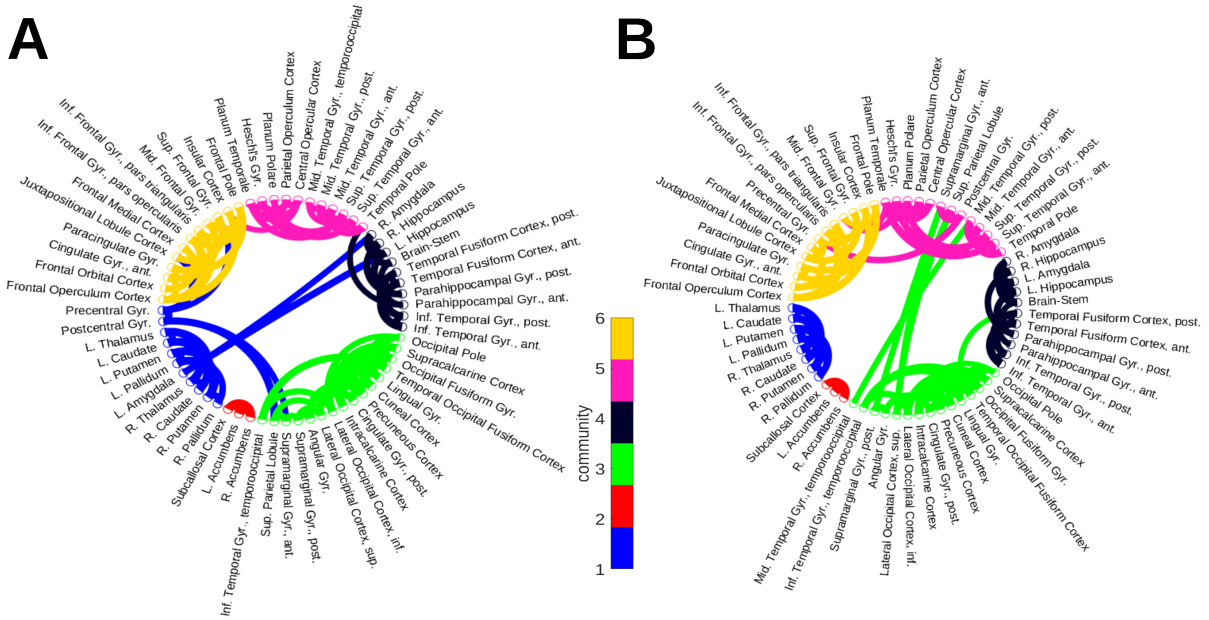


Figure 5.10: This figure illustrates the functional connectivity relationships in each of the two central hub states, state 12 and 17. **(A)** This figure shows the top 5% strongest (by magnitude) connections between and within spatiotemporal consensus communities (colour legend) in state 12. The state has mostly strong in state communities. **(B)** Top 5% strongest connections between brain regions within state 17, with regions divided into the same 6 consensus communities, showing a similar propensity for within community connection. The correlation in functional connectivity is the pairwise highest across all states indicating that state 12 and 17 have a lot of shared connectivity.

has been noted in previous studies as one of the possible ways in which nososensory and other stimuli fail to reach the rest of the brain during anaesthesia as well as in sleep [262, 295].

5.2.7 Deep Anaesthesia Brain States and Anaesthetic Susceptibility

We model the probability of hub states 12 and 17 as a function of two factors, C_{SWAS} and P_{SWAS} (that were calculated from previous experiments on the same subjects [262]) with subject age included as a confounding variable. In order to know which of C_{SWAS} and P_{SWAS} had the most important effects on brain state over and above the effects of age, we fitted two multiple linear regression models with the fractional occupancy of hub states 12 and 17 as their respective outcomes. Single variable correlation analysis yields that all three factors show some moderate to strong correlation with the probability of state occurrence, but with generally opposing relationships. The probability of state 12 is negatively-correlated with age ($r = -0.537$), while being positively correlated with both C_{SWAS} ($r = 0.575$) and P_{SWAS} ($r = 0.411$). In contrast, probability of state 17 is strongly positively correlated with age ($r = 0.727$), and negatively correlated with C_{SWAS} ($r = -0.463$) and P_{SWAS} ($r = -0.551$). The model residuals were then

tested for heteroscedasticity using the Breusch-Pagan test [63], which was significant for the state 12 model indicating potential heteroscedasticity ($p = 0.0288$) but not state 17 ($p = 0.641$). Given the small sample size, we used a standard normal Kolmogrov-Sminov test on the residuals of both models which exhibited no evidence of non-normality ($p > 0.1$).

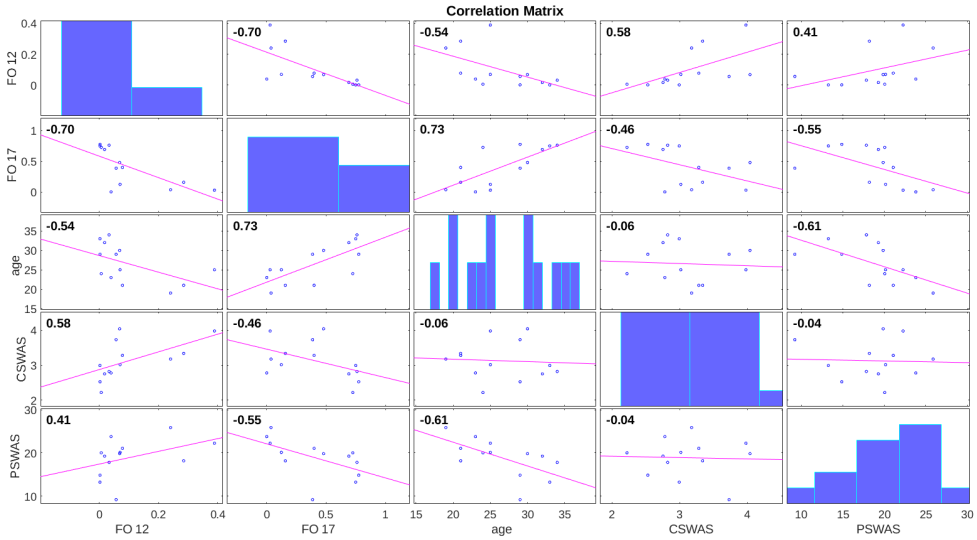


Figure 5.11: This figure summarises the linear relationships between each of the hub state probabilities (states 12 and 17), given by the Fractional Occupancy (FO), and the age, Slow WAve Saturation (SWAS) Concentration (C_{SWAS}) and Power at SWAS P_{SWAS}. These measures were recorded from EEG data in a previous study [262]. The off-diagonal plots are of single linear fits between the named variables. The number in the corner indicates the Pearson correlation. The diagonal plots are histograms of each variable.

Multilinear regression models were fitted and the coefficient and error estimates for state 12 were heteroscedasticity corrected. The model fit results (see Figure E.7 and E.8 for full models) show that when age is taken into account, C_{SWAS} has a significantly negative effect on the probability of state 17 ($p = 0.0328$, Student's t-test). There is no significant positive effect of P_{SWAS} on this probability over and above age, which had a significant effect ($p = 0.0297$, Student's t-test). This pattern suggests that age and anaesthesia susceptibility (as measured by C_{SWAS}) not slow wave power are the major determiners of the dominance of this hub state. Conversely for state 12, only C_{SWAS} has a significant positive effect on the FO ($p = 0.0323$, Student's t-test), indicating that subjects that required a higher dose to attain SWAS tended to have a higher occurrence of state 12 (thus a lower susceptibility).

Combining these results with the correlation analysis we can interpret state 12 and 17 as dividing subjects into two groups. Older subjects who may be more susceptible to anaesthesia

tend towards high state 17 activity. While less susceptible subjects may tend towards state 12. This can be interpreted as state 12 representing a brain state of slightly lower anaesthesia depth that is occupied by subjects who may be at a lower effective depth of anaesthesia due to their relative youth. These subjects tend to require a higher anaesthesia dose to enter SWAS and may generate a greater power at SWAS due to their high relative grey matter volume (which decreases with age). In support of the low (state 12) and high (state 17) susceptibility hypothesis, we found that time at LOBR is lower when FO of state 12 is high ($r = -0.328$) and early LOBR often leads to high FO of state 17 ($r = 0.440$). While neither correlation coefficient is significant, $p > 0.1$, these results corroborate this explanation and provide testable hypotheses for future large studies.

5.3 Discussion

Our results show clear differences in brain activity between deep anaesthesia and wakefulness at the static level that reveals clear links between functional connectivity in wakefulness and in anaesthesia but suggests greater parietal-cingulate-prefrontal regional activity in deep anaesthesia with overall decreases in thalamic activity. We find a weaker overall similarity in functional connectivity between the two conditions. The results suggest a high level of within subject similarity. However, this direct approach is unable to determine any dynamic switching in activity that occurs within conditions.

In order to determine the dynamic features of activity during deep anaesthesia, we then apply the HMGM and functional homogeneity analysis framework to deep anaesthesia with comparison to a model of resting wakefulness. We also introduce new measures of dynamic brain state complexity and employ new methods to uncover consensus community structure. The results show that deep anaesthesia dynamics are less complex and dominated by a small set of central 'hub' states. We again find, a conserved parietal-cingulate-prefrontal network emerged as highly functionally homogeneous in both states. The states themselves are largely mutually exclusive in activity and act as subject specific sinks with high transition in-degree. The dominance of each hub is modulated by a number of neurobiological factors, including age and susceptibility to the anaesthesia dosage with state 17 arising as the 'deeper' state that is occupied by subjects which are more susceptible to anaesthesia and thus likely to be at a greater anaesthetic depth at the same dosage. While higher anaesthetic depth in state 17 is evidenced by lower connectivity,

complexity and overall activity, functional connectivity across hub states was found to be very similar, while overall activity was less conserved.

With regards to functional activity, anaesthetised subjects saw an overall decrease in activity different from baseline. This is particularly evident in hub states, in particular hub state 17. Prefrontal, sensorimotor, cingulate and parietal areas which were active in state 12, were all below baseline activity in state 17 which is a global down state with all regions below the anaesthesia baseline. Of these temporal, parietal and sensorimotor areas were less affected while basal ganglia activity was nearest to baseline. Increased posterior frontal and motor activity in state 12, are generally consistent with previous findings for subjects receiving a moderate anaesthesia dose during the transition to loss of consciousness [262]. The thalamus appears to be affected in a dosage dependent manner as well with its activation or inactivation depending on which hub state is dominant. A possible explanation is that as the thalamus becomes more isolated in the deeper state 17, thalamic metabolic demand and thus activity decreases. This decrease was not evident from the changes in static activity during anaesthesia alone but were a result of subject and time-specific state change. This is evidence for the thalamocortical disconnection in deep anaesthesia and possibly unconsciousness more broadly. In general, the presence of low activity dominant states is also seen in HMM models of sleep with low activity states emerging to dominate more than 50% of fractional occupancy in the deepest N3 stage of sleep where subjects are most likely to be unconscious [369].

Our characterisation of brain state functional connectivity suggests that anaesthetised subjects have a poor capacity for efficient communication between brain regions. Not only this but anaesthesia hub states were particularly poorly suited to efficient communication, suggesting a default state of inefficient communication in anaesthesia. Reduced connectivity is a pattern seen in brain state HMMs of minimally conscious and vegetative subjects [30], suggesting disintegration of information integration is a common feature of both pharmacological and pathological unconsciousness. These results are supported by static analysis of connectivity in the anaesthetised brain where decreased capacity for information transfer is seen in functional connectivity networks of anaesthetised subjects [238].

Reduced integration of state functional connectivity is accompanied by a reduction in the dynamic complexity of state change as measured by the state diversity, switching and entropy

rates. In particular, we show that the ER is anticorrelated with unconsciousness. While these results corroborate previous results of the reduced complexity in raw EEG and fMRI time series under anaesthesia when compared to wakefulness, the new ER measure provides a novel approach that measures the complexity of whole brain state change. Although these early results are promising, a more sophisticated measure including state functional connectivity and activity information into the ER measure will require methodological development in the field. However, recent results suggest that these advancements are well underway [199].

A central question to our understanding of anaesthesia induced unconsciousness is whether functional connectivity decreases uniformly between brain regions with depth, or whether certain connections are preferentially conserved, with evidence from static models supporting both hypotheses [25, 227]. In support of the latter, our analysis suggests stratification of regional connectivity into mostly isolated communities occurs in a mostly state-independent way across dynamic brain states in deep anaesthesia, while dynamic shifts in connectivity and community membership occur throughout wakefulness.

In addition to being highly correlated in functional connectivity, the two hub states 12 and 17 share a conserved community, community 6. This parietal-prefrontal-cingulate community shares key brain regions with the so-called salience network, a basal network in wakeful brain activity (see Table A.1). A community involving the precuneus, a parietal region previously suggested to be less affected by anaesthesia dose, community 5, is implicated in state 12 as being functionally homogeneous but not state 17. This may be due to dosage dependent reduction in the coordination of these regions which is no longer present in the 'deeper' (high susceptibility) anaesthesia state 17. It is notable however that parietal areas remain connected to each other and to sensorimotor regions in both states, while not belonging to the same community. It is possible that performing consensus clustering over a higher resolution range would pick up possibly conserved core network regions in even the deeper state. Overall, our analysis corroborates findings of a low activity, coordinated network resembling a frontal-parietal-cingulate or salience network that persists in deep anaesthesia dynamics.

Our findings regarding the anaesthesia dose-state probability relationship rely on data that was analysed from a previous EEG only experiment and was not simultaneously recorded along with fMRI. This was due to lack of availability of clean simultaneous EEG data for all subjects from

which to obtain accurate P_{SWAS} and C_{SWAS} estimates. As there is some variability in the exact time and concentration at which subjects lose consciousness, this could contribute to noise in the apparent relationship between EEG markers and anaesthesia state probabilities. Later unpublished experiments which included repeated measures from the same subjects suggest that such markers remain largely consistent. In addition, heteroscedasticity in one model limits the interpretability of the significance and magnitude of some of the coefficients, non-linear models, such as logistic regression could improve the results.

It may be that rather than being markers of anaesthetic depth, different hub state dominance in each subject reflects fundamentally different pathways to unconsciousness that are specific to that subjects' physical and mental state at the time of the trial, as factors such as sleep quality and state anxiety are known to have quantifiable effects on anaesthesia effectiveness, at least at lower doses [392]. So far, the evidence we provide for a similar unconscious, sink-like brain state in humans rests mostly on how large a fraction of our subjects are dominated by only a single brain state and the similarities in connectivity across these two states. Though hub state occupancy appears somewhat subject dependent, we found no direct connection between wakeful and anaesthetised hub state dynamics on a subject level. This inability to pick up subject-level effects may be a limitation of the small study size. Future studies that focus on the transition both into and out of unconsciousness during deep anaesthetisation should be able to more clearly ascertain if the difference in hub state occupancy is due to a difference in pathway or merely a difference in the subjects' position along the same trajectory to a general unconscious brain state. Regardless of which of the two bares out in the end, mapping the network pathways to unconsciousness continues to be a deep and unique mystery with important surgical and philosophical implications.

6 | Modelling and Model Selection for Neuronal Calcium Dynamics

Towards understanding neuronal integration and synaptic plasticity

In this chapter we examine a different kind of Long Range Dependence (LRD) in the 3D structural images of neurons and image time series of neuronal activity. Like in whole brain macro activity, LRD in temporal and spatial networks plays a role at both micro and mesoscales of neuronal activity. On the mesoscale of interneuronal network dynamics, the interplay between neuronal ions, in particular calcium ions (Ca^{2+}) and network topology play a well-established role in the dynamics of learning, memory and behaviour [376, 442]. However, the complex microscale topology of the individual neurons that constitute these mesoscale is often overlooked. These subcellular networks plan an important but poorly understood role in neuronal integration, the processing of incoming signals within each neuron. Understanding the process of integration within neurons importance both to our understanding the mechanics of neural processing but also to the mesoscale and macroscale processes of learning and memory.

New light microscopy techniques and technologies can help us discover more about the relationship between neuronal topology and local neuronal Ca^{2+} fluctuations, termed Ca^{2+} events (or transients) [167]. These events are the markers of signals being integrated by the neuron, small fluctuations that start from synapses and propagate locally along the complex tree-like topology of the neuron. Given the important role that topology plays, we put forward a new pipeline for modelling and model selection of Ca^{2+} events on neurons to aid neuronal integration researchers in making important experimental and analysis decisions.

Synaptic plasticity is the change of size and strength of synapses which can occur in response to a large cell-wide excitation called an Action Potential (AP). APs result from many concurrent ion fluctuations termed Excitatory Postsynaptic Potentials (EPSPs) across the neuron's surface,

which can be observed through microscope-based Ca^{2+} imaging. This summation of EPSPs to produce an AP (or not) is the final output of all the subcellular integration activity, ultimately resulting in the macroscopic activity of the brain and its behaviours [253, 379]. Although important to our understanding of how APs originate, integration is not as well studied in the sub-threshold regime (when the events are insufficient to produce an AP). Studying neuronal activity in this regime is particularly difficult due to the high spatial and temporal resolution required to resolve events across a neuron and the relatively low signal to noise ratio of the localised, low amplitude spikes observed. We analyse pilot data from two Light Sheet Microscope (LSM) set-ups for imaging single neuron Ca^{2+} activity and generate a simulation that accurately captures these real-world constraints. From this simulation and our analysis tools we determine which of a variety of methods may be best suited to detect Ca^{2+} events in the sub-threshold regime from future microscopy data. The pipeline we propose follows our experimental timeline but may be adapted to other imaging set-ups and be used in whole or in part to inform research at many points during the neuronal imaging and experimentation process.

LSM is a relatively new technology in microscopy with enormous potential for fast imaging of large (on the microscale) sections of neurons or even whole 3D neuronal activity at speeds fast enough to detect Ca^{2+} event timing and location [182]. The Field Of View (FOV), size of the imaging plane, available with LSM is wide. This allows for potentially high quality structural images of whole neurons to be captured which can be used for moderate to high fidelity reconstruction of neuronal topology, a process called neuron tracing [432]. Light sheet imaging is an appealing alternative to multi-photon and confocal microscopy that lack the temporal resolution necessary to image Ca^{2+} as rapid imaging risks toxic over-exposure of the cell to light, so-called phototoxicity [330]. However, current 3D light sheet imaging relies on curved beam set-ups that requires complex post-processing by deconvolution which have to be calibrated to the particular microscopy set-up and does not always work to remove imaging artefacts [405]. These inhomogeneities are made worse in the case of dynamic 2D imaging where deconvolution to correct curvature in the laser beam cannot be performed worsening the problem of accurate automated event detection.

Interrogation of single cell neuronal Ca^{2+} activity is in its infancy and there are as of yet, no methods designed specifically to detect transient Ca^{2+} events (i.e. not APs) in the sub-threshold regime [139, 150, 300]. Manual methods, such as those discussed in Section 6.2.5, are difficult

to apply consistently or on a large scale as they are highly dependent on the choice of Region of Interest (ROI) made by the experimenter. Of the methods available to automatically detect Ca^{2+} events, the simplest class use the cumulative summation (CUSUM) algorithm [283]. These have been employed for detection of multiple successive APs at lower temporal resolutions than is required for Ca^{2+} event detection and provide a coarse grain all-or-nothing approach that is generally conservative. CUSUM methods follow a simple summation and thresholding rule that produces a binary series indicating the presence or absence of events [155]. We use a modified CUSUM method to determine a baseline signal to noise ratio from our LSM pilot data. Of the other methods available, many of the most widely used for spike-like event detection are deconvolution methods which decompose the Ca^{2+} signal into a train of discrete spikes and background signal. These methods were generally designed to detect APs in images with multiple cells or circuits as they operate at the lower spatial resolution of whole neurons. However, they may be repurposed to detect Ca^{2+} events which are also discrete spike-like transient events in the neuronal Ca^{2+} signal. Smoothing is a necessary step when comparing events on a large neuron across space and time due to the sparseness of both automatically and expert identified events.

We use 3D structural image data to extract a realistic neuronal topology and 2D videos of Ca^{2+} activity captured using LSM to motivate a model for Ca^{2+} dynamics [167, 432]. Neuron tracing is used to extract the topology of the neuron while new methods are proposed to remove artefacts and segment the neuron into its anatomical components. We perform automated tracing of the neuron. This method is more error-prone but more practical as it is far less labour intensive than expert directed manual tracing of the neuron using image analysis software [4, 333].

Development of novel microscopy methods for Ca^{2+} event detection requires constant refinement of the experimental protocols and analysis tools. The pipeline we propose involves several steps, alternating between data acquisition and analysis in a collaborative fashion with experimenters. Each step refines the analysis and model choices to arrive at an optimal combination of detection method and experimental protocol for future experiments. Unfortunately these experiments were not conducted due to pandemic related funding and lab-access restrictions. Figure 6.1 shows the full pipeline for Ca^{2+} event detection and refinement. The pipeline combines real pilot or simulated data to inform further choices. The pipeline involves analysing real data from signal-close-to-noise Ca^{2+} light microscopy as well as simulated neuronal Ca^{2+} activity. The pipeline allows for appropriate suggestions to the experimenter which inform the modelling and

experimental choices for follow up experiments. Model selection techniques are based on the free energy principle, a model selection framework in neuroscience and statistics [140]. We also develop novel graph diffusion-based smoothing techniques that make accurate model selection possible given the unique topology of each neuron and use the criterion and smoothing techniques to select models for Ca^{2+} event detection.

We find corroborating evidence that information propagation can be detected in neuronal Ca^{2+} intensity time series in a directional and topologically dependent way [61,124]. We show that neuronal Ca^{2+} activity as measured by LSM is dominated by coloured noise, a form of Long Range Dependent signal introduced in Section 1.1. We demonstrate that a slightly modified version of CUSUM event detection, which is robust to many forms of inhomogeneity in background activity, can detect event locations with reasonable accuracy in a limited proof of concept dataset and outperforms more sophisticated deconvolutional approaches that assume a strongly stationary background process in simulation. We develop a computationally efficient heat kernel-based diffusion smoothing technique. Our empirical results suggest that free energy with diffusion-based kernel smoothing is a consistent model selection criterion, selecting the correct model parameters in a simulated Ca^{2+} experiment. Previous methods for event detection model selection on neurons have largely ignored both the effects of neuron topology and the benefits of smoothing in favour of using coarser temporal resolution that may miss events [326]. Lastly, having shown the validity of the model selection approach, we then use the free energy selection criterion to make experimental recommendations based on further hypothesis-driven simulations.

Distinct from the previous chapters, the structure of this chapter roughly follows the progression of the experimental and simulation phases in our pipeline. The first section provides background on the necessary neurobiology and microscopy understanding (including classical supervised event detection techniques). The second section details classical microscopy analysis methods and terminology for a pilot study based on a previous LSM set-up detailed in Appendix F.1. The third section presents results based on this classical analysis including more statistically motivated time series analysis of the dendrites and neuronal background. Section 6.4 details an analysis pipeline for extracting neurons from 3D structural images and detecting Ca^{2+} events from the extracted neuron in an unsupervised way. Section 6.5 details the simulation pipeline used in our analysis, which is based on the framework in [361], with novel improvements to increase realism of the Ca^{2+} dynamics including fluorescence decay, silent synapses and techni-

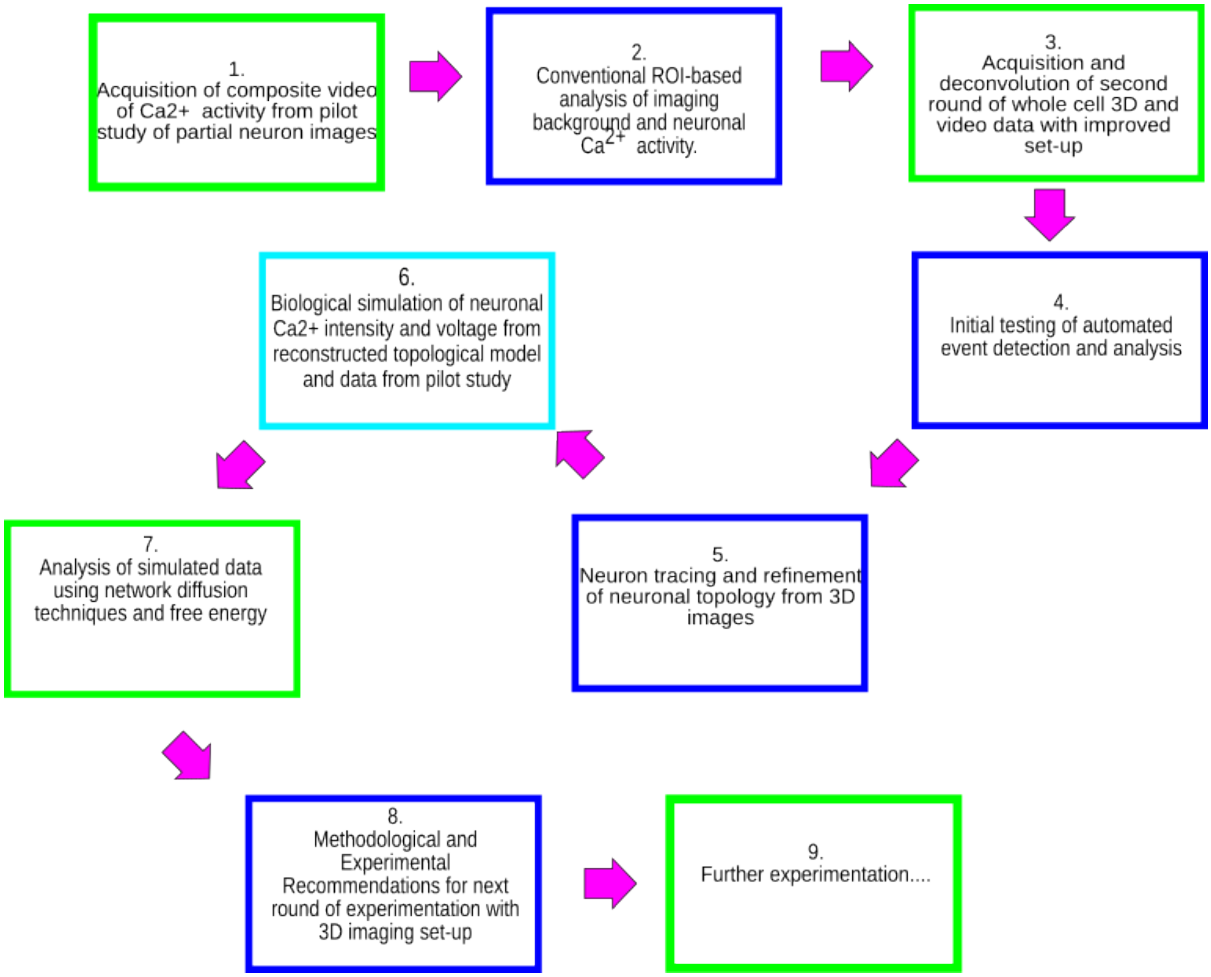


Figure 6.1: This flow chart shows the pipeline that was followed for Ca^{2+} event detection and experimental refinement. The pipeline provides a guide for future experimentation which involves refinement through alternating experimental and analysis phases. Purple arrows show the order of steps in the pipeline. Green boxes show data acquisition steps in which data is generated for analysis or refinement. Cyan boxes combine data generation with analysis (such as for simulated data). Blue steps are purely for analysis of the experimental or simulation data. The pipeline starts with a first round of pilot data acquisition of video Ca^{2+} data. These videos are generated from maximum intensity projection of the only 3D dynamic data in the study. In the next step we perform conventional microscopy analysis similar to [167] along with time series analysis in order to understand the Ca^{2+} activity background and manual event detection procedures. In the second phase, collaborators generate a new pilot data set-up based on a new LSM set-up based on the original which included dynamic 2D video of Ca^{2+} events as well as 3D structural images. We then test a simple automated event detection method and perform neuron tracing to extract the topology. Neuronal topology and pilot data analysis informs a biological simulation of Ca^{2+} and voltage dynamics in a biophysically realistic neuron simulation and the simulation data is used to inform experimental and event detection model choices for future rounds of experimentation. Unfortunately, future experiments were not able to be completed due to COVID-19-related funding and lab restrictions.

cal noise. Section 6.6 presents novel mathematical and computational tools for detecting Ca^{2+} events and comparing the performance of different methods based on an information theoretic framework and kernelised smoothing approach. This is followed by simulation results and final discussion sections.

6.1 Background

This section relates the structure and function of dendrites and dendritic spines and the role played by Ca^{2+} in their modulation to the types of experiments that have and will be performed with the fast volume-scanning, Light Sheet Microscope (LSM) imaging techniques.

6.1.1 Dendritic Structure and Function in Context

Neuronal dynamics occur at the slower level of re-wiring, and at the much faster level of ion-mediated voltage signals passed between and within these wires. Individual neurons are geometrically complex with multiple long input (dendrite) and output (axon) filaments collectively called neurites (see Figure 6.2, left). Neurites connect to other neurons at junctions called synapses. Neurites may have multiple branches but all neurites belonging to the same cell meet at the soma. Neurites are biochemically complex, containing extensive networks of cytoskeletal and scaffolding proteins, as well as protein synthesis machinery [176]. This machinery allows dendrites and axons to grow and branch to form complex arborisations and respond to signals by ion and neurotransmitter release [372]. We focus largely on the biochemical activity of the former as these are where the synaptic inputs are located and thus where Ca^{2+} events originate.

Dendrites and Neuronal Integration

Dendrites are branched tree-like protrusions from the main body of the neuron. The unique and complex branching pattern formed by dendrites, termed the dendritic arbour, change during development and in learning and is notably more complex than the axonal branching pattern in most cell types. Dendritic re-wiring defines one form of neural plasticity which occurs over a longer time frame than synaptic plasticity [178, 209, 248, 304]. Dendrites integrate the inputs they receive from pre-synaptic cells by electrochemical propagation of signals along the dendritic arbour which acts like a potentially non-linear filter on the signal [390]. Propagation occurs as a result of protein pores or channels in the surface of the dendrite that modulate the exchange of ions with the cellular medium. When a sufficiently large voltage potential reaches the soma an AP is produced which propagates to downstream cells via the axon. A simultaneous global event known as a backpropgating Action Potential (bAP) leads to a transient cell-wide increase in voltage and corresponding global uptake of Ca^{2+}. The AP threshold is known as the firing

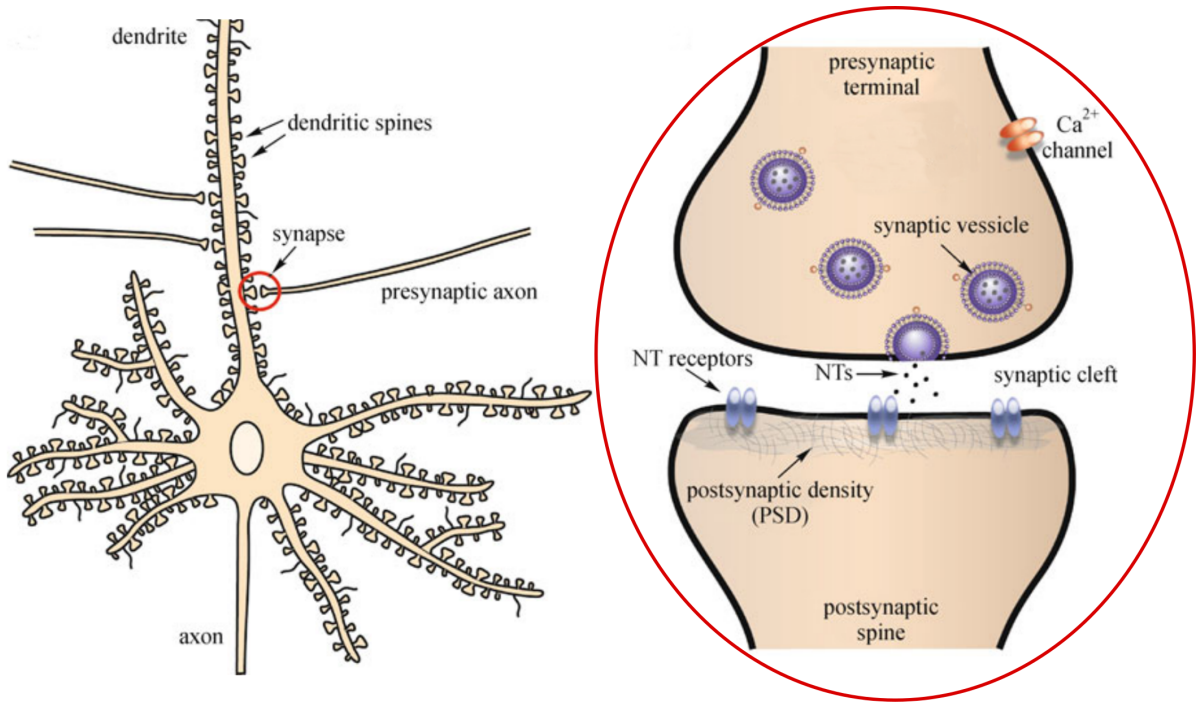


Figure 6.2: A number of axons are shown connecting with a typical postsynaptic neuron of the central nervous system (left). The neuron has a single smooth axon and multiple dendrites with many nodules (spines) that form synapses with the axons of other neurons (not all of which are shown). The red circle around one of the synapses (right) gives a more detailed illustration of the components of a typical synaptic connection between an axon and dendrite (right). The presynaptic neuron releases neural transmitters (NTs) from vesicles at the cell surface. These transmitters bind to ion channel receptors on the postsynaptic spine surface, triggering local ion flux and voltage changes through these and neighbouring channels (including Ca^{2+} channels). This figure is adapted from Smrt *et al.* 2010 [364].

threshold and is a biophysical property of the cell. In rat hippocampal excitatory pyramidal cells, the cell-type we have access to, this is roughly -45mV [138]. In these experiments, membrane voltage excitations were kept well below the -45mV threshold to stay within the sub-threshold regime.

Neuron Topology and Non-Linear Processing in Neurons

The topology of the dendritic tree and protein composition of the dendrites have a great influence on the brain's computational complexity and are strongly implicated in computer simulations to influence signal integration in a non-linear way [203, 259]. Difference in arborisation (the branching pattern of the tree) is a major source of variation between neuronal cell types in different parts of the brain [305], suggesting a connection between neuronal function and the architecture of dendrites that may increase the overall computational complexity of neurons. Specifically, when compared to their artificial counterpart, a single layer perceptron which is limited by its ability to only encode linearly separable data [312], a neuron's dendrites can act as

non-linear transformations on the inputs, while branches in the neuronal tree act like secondary outputs leading a single neuron to have a theoretical functional expressibility more similar to a two-layer perceptron network [298]. Neuron topology thus potentially increases the brain's overall capacity to encode complex algorithms that generate a complex internal representation (see Chapter 5 and Section 5.1.4).

In addition, electrophysiological properties such as impedance and conductance, have a non-linear effect on signal strength depending on event location, with event magnitude increasing towards the distal tips of the neuron, far from the soma [19]. Pyramidal neurons (the type in our imaging data sets) typically have one long apical dendrite at the top of a pyramid-like soma and two or more basal dendrites [44]. Neurons can thus be viewed as a tree with the soma as root node [88]. This views of neurons as tree graphs is key to our analysis in Section 6.6.

6.1.2 Synaptic Spines and Ion Channels

The synapse is crucial to information transfer, learning and memory encoding in the brain. The synapse is the physical junction that facilitates the directed transfer of information from the presynpatic to the postsynaptic neuron and is often visible on the postsynaptic side as a protruding mass in high resolution microscopy images of neurons. Information is transferred when a pre-synaptic neuron excites a post-synaptic neuron resulting in exchange of ions with the surrounding extracellular medium and a change in membrane potential. Using neuronal Ca^{2+} imaging, incoming signals from other neurons can be detected in the post-synaptic neuron as a transient influx of Ca^{2+} into the cell via the synapses and nearby channels in the cell membrane. These transient events are termed Ca^{2+} events and are accompanied by a EPSP, a change in local membrane voltage. Rapid accumulation of ionic potentials, including Ca^{2+} from repeated events plays a vital role in learning and memory formation through the process known as synaptic plasticity [120]. The structure of the neuron and synapse is shown in Figure 6.2 and is expanded upon in Section 6.1.

The cell body and each branch in the dendritic arbour, itself only a few hundred micrometres in length, may have hundreds of tiny nodules. These so-called synaptic spines protrude from the cell membrane and are the input end of synaptic junctions (see Figure 6.2, right) [444]. Each tiny spine ($0.001 - 1\mu m^3$) is a connection point, one input terminal of a synapse. A spine receives excitatory neurotransmitters from an outputting axon and responds by opening channels

(voltage-dependent, ion permeable pores). These channels are embedded most prominently in the postsynaptic density which is a receptor and channel dense region on the surface of dendritic spines.

The opening of different channels allow for the influx of ions. This flow of ions alters the voltage across the membrane and leads to the opening of further, voltage sensitive ion channels [372]. Among these secondarily active channels are N-Methyl-D-aspartic Acid Receptors (NMDARs) and Voltage-gated Calcium Channels (VGCCs), which are responsible for the majority of Ca^{2+} influx into the cell following an event [334]. Although many ions and channels are involved in creating the voltage potential Ca^{2+} influx alone is sufficient to induce modulation of the signalling response strength, size and permanency of synapse and post synaptic density [247].

As a result of Ca^{2+} activity spines change in size, number and location to encode long-term memories and behavioural programmes [435]. The process of spine turnover and shape modulation in response to neural activity throughout an organism's lifespan is the defining marker of synaptic plasticity [181]. These changes in spine size are what allow synaptic plasticity to be investigated by optical microscopy methods [120, 341].

Rat hippocampal pyramidal neurons are fairly typical for excitatory cortical neurons with synapse counts on the order of 10,000 synapses per neuron [32]. Some of these may be what are called silent synapses, these are synapses that lack functional AMPAR on their surface. These synapses usually have low to no responsivity to input (hence silent) but may still be contributors to the neuronal Ca^{2+} background through the activity of NMDAR receptors. Silent synapses are known to occur prevalently in the pyramidal neurons of younger rats (12 days or younger) [73].

Controversy Regarding Pre and Postsynaptic Plasticity

It is well established that simultaneous activity between the pre- and postsynaptic neurons is necessary for synaptic plasticity [253]. However, it is a matter of some controversy as to which of these have a greater functional role to play in the plasticity of the strength and reliability of the synaptic response and often times, a purely postsynaptic explanation is proffered [416], neglecting the contribution of the presynaptic boutton in plasticity [440]. In addition, both pre and postsynaptic morphology can change in order to balance the excitatory and inhibitory activity that is needed to maintain electrochemical equilibrium and neuronal health [96]. This

type of synaptic change to maintain proper cellular and signal integrity is known as homeostatic plasticity [397]. It's function may be similar to that of batch normalisation in deep neural networks, which is used to ensure fidelity of the backpropagating signal between layers [349].

Despite their similarities, presynaptic plasticity can be induced independently of postsynaptic plasticity and may play an important role in maintaining inhibitory-excitatory balance in the brain through changes in neurotransmitter release probability [268]. The downplaying of the presynaptic role in plasticity is at least in part due to the difficulties in imaging presynaptic activity as opposed to postsynaptic activity, leading to a bias in the availability of data and thus apparent importance.

6.1.3 The Role of Non-Excitatory Cells

Some neurons lack spines altogether ($\sim 20\%$). These neuron are mostly inhibitory [324]. Inhibitory neurons dampen the signals generated by excitatory cells. They do not send excitatory impulses and are modulated by different and less well-understood mechanisms that may also involve intracellular Ca^{2+} release [282]. Inhibitory cells are very important for maintaining the excitatory-inhibitory balance, preventing destabilising positive excitatory feedback loops. Since inhibitory synapses have no spines and their activation is not directly connected to Ca^{2+} release, we cannot include them in our analysis. A range of other cells, called glial cells, also serve important regulatory and structural roles in the brain but do not send or receive excitatory signals that could be detected by Ca^{2+} imaging [24].

6.1.4 The Dynamic Effects of Ca^{2+}

The voltage-dependent influx of Ca^{2+} through spines and into the dendrite is confined spatially and temporally by the small diameter of the adjoining spine neck and by the action of Ca^{2+} buffers. Ca^{2+} buffering in the dendrite proximal to an event occurs on very fast time scales, at a half decay time of roughly 15ms [322]. Propagation of Ca^{2+} signal towards the soma is thus mainly due to shared voltage conditions and the action of other ions (namely Na^{+}) that spread VGCC activity and increase the chance of NMDAR activity [75]. This in turn can increase de-polarisation and, when activity is strong enough, cause the formation of a spike or transient Ca^{2+} event. A Ca^{2+} event is defined by a local 'peak' in dendritic Ca^{2+} levels that can last on the order of a few hundreds of milliseconds to a few seconds seconds. This is due in part

to dendritic signal propagation, but also the much slower dynamics of the fluorophores used to image Ca^{2+} activity in practice (see Section 6.1.5) [297].

Ca^{2+} Induced Synaptic Plasticity

Repeated Ca^{2+} influx into a spine is the catalyst for changes in spine morphology. These changes are frequency dependent. When Ca^{2+} activity at a spine is consistently accompanied by simultaneous AP firing at that spine's neuron this can lead to spinal enlargement and an increase in signalling strength and stability of the spine. This increase in synaptic response strength and spine size is termed Long Term Potentiation (LTP) [114]. When Ca^{2+} influx into a spine is infrequent or is not accompanied by a bAP, the response strength of a synapse may weaken and prolonged mismatch in activity may cause the spine to be reabsorbed into the dendrite, through a process known as Long Term Depression (LTD) [255]. LTP and LTD as they occur in the hippocampus are the most well-understood models of synaptic plasticity [246].

This simple description of synaptic plasticity is complicated by so-called heterosynaptic plasticity wherein modulation at one spine can have an opposite effect at neighbouring spines [280]. The role of heterosynaptic plasticity is poorly understood but it may be employed to increase input-output specificity of neurons or contribute to network stability (as a form of homeostatic plasticity). This research is aimed at first finding ways to identify methods for analysis of whole cell Ca^{2+} dynamics, so that experimenters can determine how heterosynaptic plasticity may affect neuronal integration and information storage in further studies.

6.1.5 Imaging of Dendritic Ca^{2+} Activity

Historically, initial experiments to detect the presence of dynamic patterns of Ca^{2+} activity within dendrites were performed using a charge-coupled camera [95]. Current methods make use of synthetic fluorescent Ca^{2+} binding dyes or fluorescent protein complex which bind to Ca^{2+} and respond to light of a specific wavelength. The two most widely used methods for imaging have been confocal microscopy [13] and, most recently, two-photon microscopy experiments [18, 256]. Two-photon imaging has a higher penetrance than confocal microscopy, with a small focal volume which can allow for highly site-specific, deep-tissue imaging of a small number of spines [372]. In addition, two-photon miscroscopy can be more easily coupled with methods

to excite specific spines [105] (see Section 6.1.6).

Like confocal microscopy, LSM requires the development of specific equipment to work with site-specific excitation (excitation of a specific synapse). The advantage of LSM is that it can be employed to rapidly image large dendritic volume in a short period of time, using a specialised set-up called *fast volume-scanning by light sheet microscope* [167]. The high image capture speed to volume ratio in fast volume-scanning LSM is due to the mode of dye fluorescence. In ordinary LSM, a whole cross-sectional plane of the slice is illuminated at once using a highly focused sheet of laser light (produced by a specialised lens) which passes through the entire focal plane. In the fast volume-scanning LSM set-up, rapid repositioning of the sheet or brain slice to capture planar images at different depths allows for full slice volumes to be captured in as little as 20ms (see Appendix F.1).

Recent advances in transgenic rodents, especially the development of transgenic Cre-GFP fluorescent complexes, which utilise a fluoroscent protein expressing virus, has made neuronal microscopy more reliable, is able to target whole populations of cells and is less likely to damage the neuron during imaging than synthetic fluorescent compounds [87]. LSM can be coupled with this approach to target single cells if the viral concentration is low enough to only infect a small number of cells within a tissue sample. Such a set-up has the advantage of being less toxic, allowing for longer in vitro imaging and enabling behavioural imaging experiments with live animals [197]. We differentiate three different types of image stacks (collections of 2D images): 2D videos or XYT image stacks, these are conventional videos of activity, where X and Y are the two axes within the focal plane of the laser and T is time; XYZT image stacks, where Z is the axial direction (perpendicular to the imaging plane in XY images); and 3D structural XYZ images, these are used for topological reconstruction (tracing) of the neuron.

6.1.6 Ex Vivo Ca^{2+} Brain Slice Imaging

Ex vivo brain slice imaging involves the extraction of a thin cross-section of the brain for imaging. It is ideal for study of individual neurons since the cell can be directly manipulated and the flat slice can be imaged more easily than in a live animal (e.g. through an intracranial window). We focus here on experiments with organotypic brain slices as distinct from acute slices. Organotypics are slices that have been allowed to culture and rejuvenate following extraction, making them less volatile [103]. However, the reformation of synaptic connections that occurs

during culture makes the cells morphologically distinct from acute (uncultured) slices, which are imaged directly after extraction. Rewiring in organotypic slices can also lead to pathological behaviours such as electrochemical over excitation and seizure-like activity [10, 97].

After culture or extraction the brain slice is placed on the microscope stage and submerged in Artificial Cerebral Spinal Fluid (ACSF). ACSF is a fluid manufactured to match the essential ions, oxygen and pH levels required to maintain the capacity for neuronal activity [212]. A brain slice can remain submerged in ACSF and continue to function and respond to stimuli for several hours [74].

6.1.7 The Free Energy Principle, Integration and Synaptic Plasticity

The free energy principle as introduced in Section 1.3 states that any self-organising system that is in equilibrium with its environment must act so as to minimise its free energy [140, 141]. This free energy is the same as was encountered in Chapter 4 and functions as an upper bound on the surprise of an observation about an organism's environment given a model of that environment.

The free energy principle has been applied fruitfully as a framework to understand the optimisation objectives of neuronal integration and synaptic plasticity [429], where synaptic plasticity tends to produce neuronal circuits that minimise the difference in timing between pre and post-synaptic APs. This minimises the surprise between the two cells and therefore the free energy [81].

6.2 Light Sheet Pilot Data Analysis

We perform an exploratory analysis of the spatiotemporal autocorrelations in a dynamic Ca^{2+} image stack produced in an LSM proof of concept study by our collaborators [167]. Our first approach was to emulate the pre-processing and analysis steps taken by our collaborators using the Regions of Interest (ROIs) that they identified in the study. This helps us to understand what elements of the method need to be improved in future steps. Our second approach was to perform analysis of local spatial correlations present in the Ca^{2+} fluorescence intensity time series. Lastly, we discuss the implications of these correlations for any automated detection methods.

6.2.1 Original Imaging Setup

The original LSM setup used to produce the data set in this exploratory analysis was constructed from a composite image in which a single exposure image was taken while the fluorescing laser moved vertically through the slice, creating a blurred XYZT image stack. The images were of the CA1 pyramidal hippocampal neurons of a single 7 day old rat. Activity was averaged in the Z direction to produce a single XYT composite stack (see Appendix F.1). This was the only set of 3D images produced by the experiment and the low axial resolution and large artefacts precluded the use of these image stacks to study 3D Ca^{2+} events.

In this experiment, OGB1, a Ca^{2+} sensitive dye was loaded onto the slice. This dye increases in intensity when fluoresced, but only if bound to Ca^{2+}. In this way, a snapshot of Ca^{2+} activity over a significant proportion of the cell is captured. In addition, synthetic dyes have the disadvantage of being more phototoxic than protein-based fluorophores contributing to the necessity of low temporal resolution (to limit exposure) and large signal deterioration and increase camera exposure.

Each image in the stack is taken over a 20ms exposure, imaging a volume with breadth (X), length (Y) and depth (Z) of $(X, Y, Z) = (300\mu\text{m}, Y = 300\mu\text{m}, Z = 160\mu\text{m})$, creating a single $300\mu\text{m} \times 300\mu\text{m}$ XY planar image per time point with pixel dimensions 2048×2048 and a depth of field of 160μm. The experiment was run over a period of 10 seconds, generating a total of 501 images per stack for a single slice. This data comes from a proof of concept paper [167]. In Sections 6.2.2 to 6.2.5 we describe some of the analysis methods employed in this study that we also employed to motivate our CUSUM-based method CUSTD (CUSUM with Standard Deviation) applied to later pilot data and our simulation approach.

6.2.2 Initial Excititatory Experiment

Our collaborators performed global excitatory experiments. This was done by changing the concentration of K^+ ions in solution in order to induce high numbers of subthreshold Ca^{2+} events for imaging. They used organotypic slices extracted from the CA1 area of the hippocampus [167]. Organotypic slices were chosen despite their morphological abnormalities due to their robustness to imaging and manipulation. Global experiments require less experimental equipment to set up and are therefore the first to be investigated under the new set-up once control (non-excitatory)

images can be produced. All that is required for global excitation is for the ACSF preparation to be prepared with an increased K^+ concentration. The result of the global depolarisation induced by the excess K^+ in the slide solution is to make the opening of VGCCs and NMDAR channels more likely. This in turn increases the probability of observing Ca^{2+} events within the FOV [48].

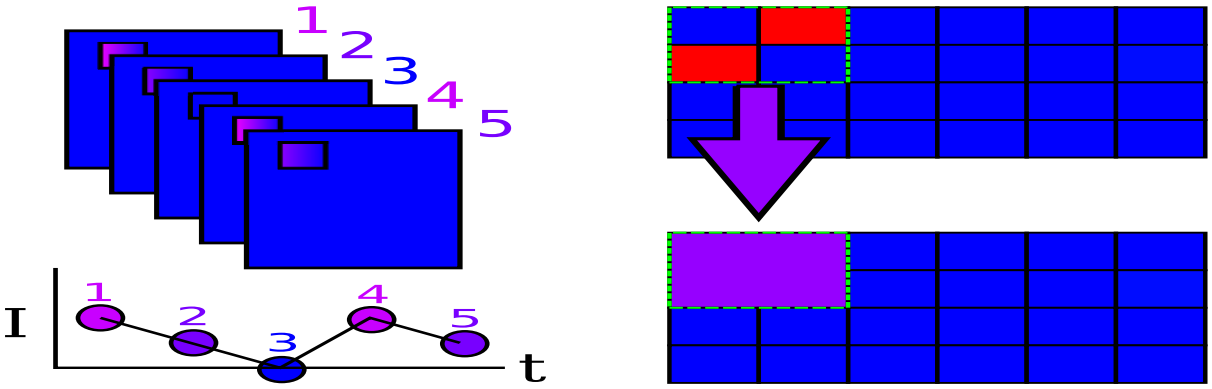
6.2.3 Specifying Regions of Interest

Regions of Interest (ROIs) are rectangular XY subregions of an XYT image. In their analysis, our collaborators selected ROI area and location based on observation of some phenomena such as a transient Ca^{2+} event, or as representative of a larger image region. We consider two types of ROIs, dendritic ROIs that span a section of dendrite, and extracellular or background ROIs that do not overlap with any of the visible dendrite.

The pixels enclosed by an ROI have variable intensity depending on their position in the stack. By taking the sample mean across the ROI pixel intensities in each image we generated a single aggregated intensity time series for each ROI (see Figure 6.3a). The size of ROIs is a compromise between the resolution of the data, the size of the object being imaged and the level of noise in the intensity time series.

In the original proof of concept paper [167], ROI area was selected on an ad hoc basis. In order to regularise analysis we chose the size of dendritic ROIs to enclose an area of 12×12 pixels ($1.76^2\mu m^2 = 3.09\mu m^2$); where a rat CA1 dendrite has a diameter of roughly $1-3\mu$m [131]. The area of extracellular ROIs on the other hand was larger in accordance with [167]. These ROIs had an area of 30×30 pixels ($4.39^2\mu m^2 = 19.3\mu m^2$).

A stack can also be divided into non-overlapping gridded subregions (see Figure 6.3b). Each subregion represents a volume with depth of field 160μm (the imaging depth) and specified XY dimensions. Dividing an image in this way is like creating a new set of pixels each capturing activity of a specific volume. As with an ordinary ROI, each grid ROI has a particular aggregated intensity profile over time. We gridded our Ca^{2+} fluorescence image stack by dividing it into an XY grid where each square is 9×9 pixels in area ($1.32^2\mu m^2 = 1.74\mu m^2$). Grid ROIs were chosen to be smaller than ROIs because noise at individual point in the grid is assumed to be less important than the fluorescence patterns across the grid.



(a) (b)

Figure 6.3: This diagram illustrates how ROIs and grids are aggregated and intensity (I) time series are produced. Once an ROI has been selected (small square in Figure 6.3a) the aggregated intensity is calculated from the pixels in the region (colour of the numerical time point of the image). Once all the aggregated intensities have been calculated at each time point, the intensity time series (I) is the series formed by the aggregated value at each time point. Figure 6.3b shows the result of aggregation of four pixels to produce a single intensity value associated with a voxel (green dotted line). The dotted line defines a volume given by the pixel area (2×2) and the depth of field of the image. As with an ROI, by considering the whole stack, each grid ROI is associated with an aggregated intensity time series as in Figure 6.3a. The resulted griding are new images with reduced resolution. After calculation of all the grid ROIs the resolution in the figure will be reduced from 6×4 pixels to 3×2 voxels.

6.2.4 De-bleaching, Normalisation and Phototoxicity

De-bleaching is the correction of an image for the effects of continued light exposure on fluorescence. Bleaching occurs because fluorescent dyes undergo an irreversible reaction when exposed to light that degrades the image quality over time and releases free radicals that gradually kill the cell. This leads to a decay in image intensity and quality that can be modelled by a simple exponential decay function,

$$f(t|h) = a \exp(-bt) + c, \tag{6.1}$$

where t is the time since light exposure began, h is an intensity time series to be fitted. The constants $a, b > 0$, and c are fitted directly from $h(t)$ for $t \in [0, 1, ..., T]$. In [167] the authors subtracted the estimated bleaching curve fitted to a background ROI to de-bleach the dentritic ROIs. The curve was fit to the background ROI aggregated time series, followed by non-linear least squares fitting of Eq. (6.1).

The rate and degree of bleaching can depend on the medium, for example, whether the dye or fluorescent protein is present intra- or extracellularly. Therefore, they performed an additional bleaching curve fit to the dendritic ROI after initial background subtraction. The resultant

de-bleached ROI time series $i(t)$ was

$$i(t) = \mathrm{I}^*(t) - f(t|\mathrm{I}^*) \qquad \mathrm{I}^*(t) = \mathrm{I}(t) - f(t|\mathrm{B}), \tag{6.2}$$

where $\mathrm{I}(t)$ is the raw mean intensity time series of the ROI, $\mathrm{I}^*(t)$ is the time series after background bleaching subtraction, $\mathrm{B}(t)$ is the intensity time series of the background ROI and i is the de-bleached intensity time series for the ROI. Note that the two curves modelling the effects of bleaching are fitted iteratively using non-linear least squares by first fitting $f(t|\mathrm{B})$ and then $f(t|\mathrm{I}^*)$. This is a simple but computationally expensive technique when considering a large number of ROIs such as in a grid. In particular, the fast grid-ROI de-bleaching procedure, when fitting the explicit bleaching (Eq. (6.1)), failed due to poor convergence (convergence failure may be due to regions with high levels of noise or low levels of bleaching).

As a computationally scalable alternative for an image composed of many ROIs (such as grids) we fit a negative linear trend l to each ROI in the image stack to approximate image-wide de-bleaching in the case of grid-ROI analysis,

$$l(t|\mathrm{I}_{x,y}) = -at + b,$$

where $\mathrm{I}_{x,y}$ is the mean intensity at the (x, y) location in the 2D, XYT image stack and $a, b > 0$ to be fitted to the the ROI intensity time series $\mathrm{I}_{x,y}$. This is a common procedure to produce trend stationarity in neuronal microscopy time series [52]. We do not normalise the intensities to allow for spatial differences in ROI intensity. The resultant detrended ROI intensity is

$$\mathrm{i}_{x,y}(t) = \mathrm{I}^*_{x,y}(t) - l(t|\mathrm{I}^*_{x,y}). \tag{6.3}$$

An another simple method used in [167] for normalising ROI intensity time series is to subtract and then divide by the baseline (mean) intensity $\bar{i}$. This gives the $\Delta F/F$ normalised fluorescence time series,

$$\frac{\Delta F}{F}(t) = \frac{i(t) - \bar{i}}{h(i)}, \tag{6.4}$$

where $h(i)$ is the mean of the first one second of the ROI mean time series i. Different bleaching corrections are used, as appropriate to different types of analyses. Eq. (6.4) was used for all sections that reference $\Delta F/F$ (primarily in analysing specific regions in the FOV), while linear debleaching is used in grid-ROI analysis. Background exponential subtraction alone is used in the proposed pipeline ($I^*(t)$ in Eq. (6.2)) since the double exponential fit does not always converge and the proposed detection method, while it does not require second order stationarity, does benefit from removing confounding background activity.

More difficult to control for than photobleaching are the effects of phototoxicity, the damage or death of cells caused by persistent light exposure. Phototoxicity causes a loss of cell-membrane and functional integrity that can lead to a loss of function or a pathological ion flux (including Ca^{2+} ions into or out of the cell), leading to over or under excitability. Phototoxicity can be partially mitigated by limiting effective light illumination exposure, either in intensity or in exposure time, as was done in these experiments where continuous exposure was limited to just 10 seconds or less [172].

6.2.5 Ca^{2+} Event Identification and Selection of Regions of Interest

Following de-bleaching, Ca^{2+} events in a particular ROI were defined by a local, transient increase in normalised Ca^{2+} fluorescence. This simple threshold approach formed the basis of the later detection approach in Section 6.4.5. In [167] Ca^{2+} events are defined as a 1 second time period in an ROI time series in which a fitted rise-decay curve had a 'peak' sufficiently above the noise level in a 1-2 second base-line period at the start of the experiment [167]. The formula for the rise-decay function r is given by

$$r(t|i_s) = at\exp(-bt) + c, \tag{6.5}$$

where $a, b > 0$ and c are all functions to be fit to i_s (by non-linear least squares) which is a one second (50 time point) snippet of an intensity time series starting at s. The function r has a single global maximum $\hat{r}$ that defines the peak of a potential event after fitting [84].

An event peak was deemed significant if it was above the mean of the intensity over the baseline

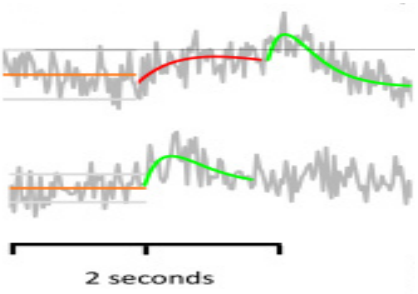


Figure 6.4: Event fitting and significance testing for time series from two separate ROIs from Figures generated in [167]. The bottom scale bar covers the baseline and first event periods. The baseline mean μ_b is given by the orange line. The grey lines define the $\mu_b+\sigma_b$ significance threshold and the $\mu_b-\sigma_b$ line. In the top plot, the first fitted event curve (red) is just below the significance threshold. Green curves show the rise-decay curve of two significant events in the top and bottom plots. The red curve shows a non-significant event and the orange line shows the baseline in the absence of events.

period μ_b plus the sample standard deviation of the baseline signal

$$\hat{r} > \mu_b + \sigma_b. \tag{6.6}$$

Figure 6.4 shows the results of event fitting for two ROIs. In the top plot, the first fitted event (red) is just slightly too low to be significant. This Figure is based on [167]. In this study ROIs were manually chosen and dispersed across two dendrites in the Field of View adding additional ROIs for areas with significant fitted events. Background ROIs were selected to be nearest the dendritic ROIs without touching any of the dendrite (determined by eye).

6.2.6 Time Series Models and Testing

We considered two types of models fitted to the last 498 time points of the normalised intensity time series of a ROI (the first 3 time points were discarded as their values were unstable). Autoregressive (AR) and Vector Autoregressive (VAR) models are related time series models which can be fitted to univariate and multivariate time series respectively. In both cases it is important that the time series is second-order stationary, i.e. its mean and autocovariance do not vary in time, which we shall refer to simply as stationary. For both types of autoregressive model a linear model is fitted to the values x_t of a time series $\mathbf{x}$ using the value of the last p time points, where p is often termed the lag order of the model. The autoregressive, AR(p), model is

$$x_t = \sum_{k=1}^{p} B_k x_{t-k} + \epsilon_t, \tag{6.7}$$

where $\epsilon_t \sim N(0, \Sigma)$ is itself a stationary time series with $\Sigma = \sigma\mathbb{I}$ (for $\mathbb{I}$ the identity matrix and $\sigma > 0$), and B_k is a coefficient fitted by maximum likelihood estimation or an approximate method. Stationarity of the original time series (and thus applicability of the autoregressive model) was determined via a unit root test (see below). In the case of a VAR(p) model, B_k is a matrix rather than a single value and Σ is positive definite [355].

The Autocorrelation Function, ACF(τ), is the Pearson correlation of a time series with itself delayed a set number of time points $\tau \in \mathbb{N}$. Similarly, the Partial Autocorrelation Function, PACF(τ) is the lagged correlation of a time series with itself conditioned on the correlation at all lags less than τ [355]. Generally, the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) give an indication of the order of an AR process, where the order is that of the highest significant lag in the PACF. Also relevant is the Cross-Correlation Function (CCF) which gives an indication of lagged cross-correlation between variables in a multivariate time series. We considered AR and VAR models with lags up to order 8 (due to computational limitations), which we fitted by minimising the Aikake Information Criterion or AIC (see Appendix F.2).

Testing for Stationarity

An autoregressive model is appropriate when there is significant autocorrelation, there is no moving average component and there does not appear to be integration or co-integration, which are related and defined as follows. An integrated time series (differentiated from Ca^{2+} integration that occurs in neurons) is one in which AR error terms are not a stationary process and instead to achieve stationarity the time series needs to be differenced with the difference operator Δ defined by

$$\Delta y_t = y_t - y_{t-1}.$$

The result of applying Δ is a new time series with one fewer time point. If repeated application of the difference operator up to m times is required for the time series to become stationary then the model is said to be integrated of order m. Two time series of integration order m_1 and m_2 respectively are said to be co-integrated if there exists a non-trivial linear combination of the

two time series that has lower order of integration than $\max(m_1, m_2)$.

A unit root test is a test designed to determine whether a given AR process is integrated. Formally, a unit root test such as the Augmented Dickey-Fuller Test, tests whether the characteristic equation of an AR process,

$$m^p + m^{p-1}B_1 + m^{p-2}B_2 + \ldots + mB_{p-1} + B_p = 0 \tag{6.8}$$

admits a root $m = 1$. If this is the case then the AR process is said to be integrated of order equal to the multiplicity of $m = 1$. An AR process that does not admit a unit root is stationary or is stationary with deterministic trend (which we have fit by de-bleaching, see Section 6.2.4). Importantly, the Augmented Dickey-Fuller test does not directly fit an AR model of order p to test for integration, but rather fits a related model based on Δy_t and successive difference terms [325].

The order of lagged differences to include for the Augmented Dickey-Fuller test unit is difficult to know a priori and is usually selected as a trade-off between the power of the test (which decreases with lags included) and the bias produced by serial correlations in a model with too few lags. One recommended choice is to use a lag order for unit root tests based on the time series length of $12(n/100)^{1/4}$ [339], where n is the length of the time series. We chose this value as it allowed us to perform the test uniformly across ROIs. The recommended lag order for our ROI time series with length $T = 498$ is 17.

Testing for Significant Autocorrelation and Local Causality

If an AR process is stationary, it is possible to test whether significant lagged autocorrelations exist by performing a test known as the Box-Ljung test. The Box-Ljung test checks whether a given set of lagged autocorrelations are significantly different from zero using a normalised sum of the individual lagged correlations. The resultant sum is distributed χ^2_m under the null hypothesis that the first m lagged autocorrelations are not significantly different from zero [355].

Lastly, the Granger 'causality' test can be used to determine whether one time series (such as from a neighbouring ROI) is a significant predictor of another. The test involves fitting of a two dimensional VAR(p) model with special attention paid to y, the proposed 'dependent' time variable, and x, the 'causal' variable. In order to derive a test statistic the VAR(p) model is

compared to an AR(p) model that includes only lagged terms from the y time series, using standard ANOVA significance test [355].

$$y_t = \sum_{k=1}^{p} B_k x_{t-k} + C_k y_{t-k} + \epsilon_t \qquad \text{vs.} \qquad \sum_{k=1}^{p} D_k y_{t-k} + \epsilon_t.$$

The test compares the two models under the null hypothesis that none of the p lags of x contribute significant explanatory power in predicting y.

6.3 Results of Exploratory Analysis

We produced two sets of results using standard time series analysis techniques. The first is based on direct Region of Interest (ROI) selection, based on the methods and data presented in [167]. The second uses the same data to investigate Ca^{2+} activity across the whole section of dendrite in-field, using a simple gridded image approach (see Section 6.2.3).

6.3.1 Analysis of Preselected Regions of Interest

An ROI is specified by a rectangular XY subregion of an XYT image stack. The size of each ROI was varied depending on whether the ROI was a control area of off-dendrite (extracellular) background meant for background subtraction or an dendritic ROI, meant for analysis (see Section 6.2.3). The dendritic ROIs we analysed corresponded roughly to the ROIs preselected by Haslehurst [167] (see Figure 6.5).

Analysis of the Off-Dendrite Background

Before investigating any of the on-dendrite ROIs we first assess whether any structure in the form of temporal autocorrelations are present in a large extracellular region after controlling for the effects of bleaching (Section 6.2.4). Figure 6.6a shows the first image in a stack with an off-dendrite background ROI selected in yellow. Following subtraction of the exponential trend which we fit (see Eq. (6.1)) using the background ROI, there is little evidence of significant temporal autocorrelation with only a few marginally significant lags. The Box-Ljung test, gave a p-value of 0.775, across 25 lags tested. The results of the Augmented Dickey-Fuller test (see Section 6.2.6) is a unit root test for autoregressive time series with an alternative hypothesis of stationarity. The results of the test suggest that there is no unit root for an AR process with up

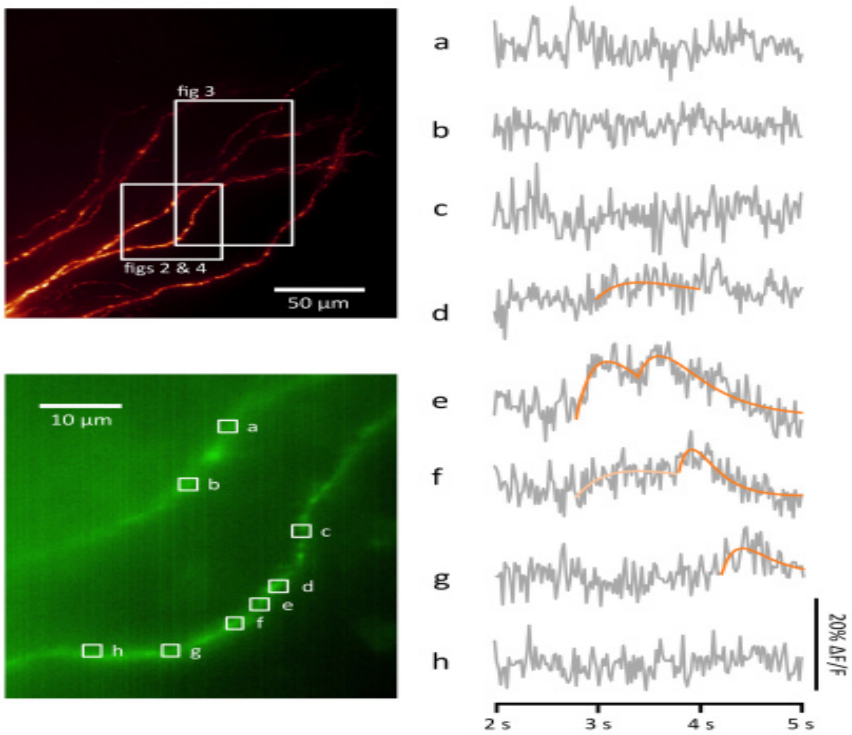
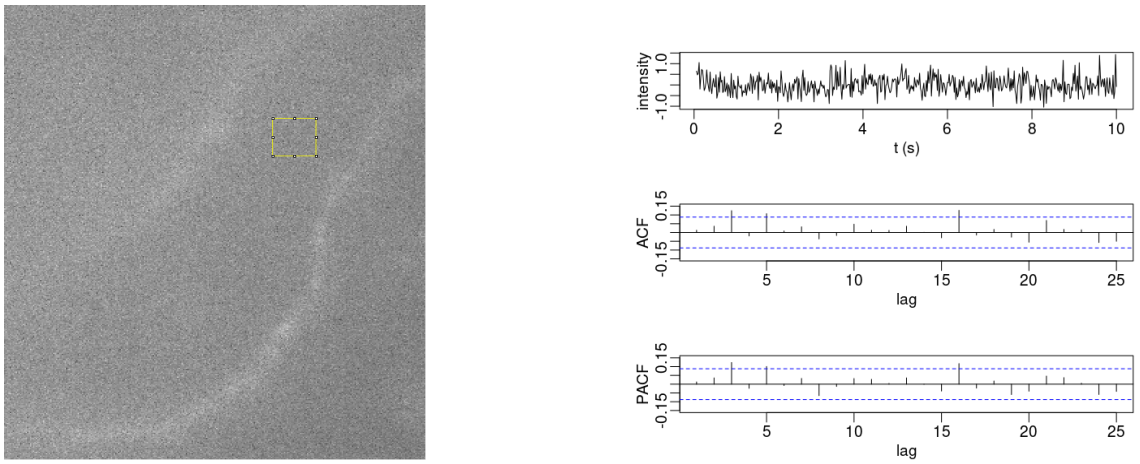


Figure 6.5: These images, produced for [167], give an overview of Ca^{2+} activity captured over a ten second span during a elevated K^+ experiment using the volume-scanning LSM method (see Section 6.1.6). The top left image is a maximum pixel intensity projection of an in field region of a single arbour with scale bar in white. The bottom left image is a mean pixel intensity projection showing two separate branches cropped from the same stack (cropped to 292×362 pixels or $43\mu\text{m} \times 53\mu\text{m}$).The white boxes from a to h show on-dendrite ROIs selected in [167]. The graphs on the right are of background subtracted ROI intensities. Fitted curves in orange indicate a possible Ca^{2+} events (light orange indicates a speculative, sub-threshold event), fitted according to a rise-decay curve (see Eq. (6.5)). These graphs suggest that events can be spread over multiple ROIs but attenuate, making them undetectable in regions sufficiently far away from the initiation site. In this case, events seem to be isolated to the centre of the bottom-most branch (ROIs d to g) and not detectable at all in the branch above (ROIs a and b).

to 17 lags ($p < 0.01$), which is consistent with the hypothesis that the process is approximately stationary.



(a) (b)

Figure 6.6: Background Region of Interest (ROI), selected in accordance with background ROI selection in [167]. Figure 6.6a generated using Fiji [333], shows the first image in the stack of an XY cropped version of a $300\mu m \times 300\mu m$ XYT image stack (cropped to 292×362 pixels or 43μm$\times 53\mu$m) produced from the fast volume-scanning LSM set-up, the same XYT stack used in Figure 6.5. Figure 6.6b shows the detrended intensity of the background ROI over the last 498 time points. The lagged autocorrelation function (ACF) and partial autocorrelation (PACF) plots suggests that the background intensity time series has individually significant autocorrelations that are significant (above the blue, dotted line). The results of a Box-Ljung test for significant overall significance of ACF lags returned a non-significant value of 0.775, suggesting that the correlations that do exist result from randomness in the sampling process.

Multivariate Analysis of Dendritic ROIs on Separate Branches

We examined dendritic ROIs, coincident with ROIs b and d in Figure 6.5, with control region as in Figure 6.6a. Following normalisation and de-bleaching (see Section 6.2.4), the Augmented Dickey-Fuller results suggest that the individual time series associated with ROIs d and b have no unit roots up to order 17, again suggestive of stationarity of the fitted AR process ($p < 0.01$). The Ljung-Box test suggests that both ROIs d ($p < 10^{-5}$) and b ($p < 0.05$) exhibit significant temporal autocorrelations. The marginal significance of temporal autocorrelation in b could be due to the lack of Ca^{2+} events present (see Figure 6.5).

The lagged correlation of signals at d and b is significant as indicated by the CCF in Figure 6.7. There are a number of highly significant lags, clustered about lag 20, or around 400ms in time. This could potentially represent the time taken for the signal to travel between these two branch-separated regions, or it could be related to the difference in arrival times between signals originating at the branch point.

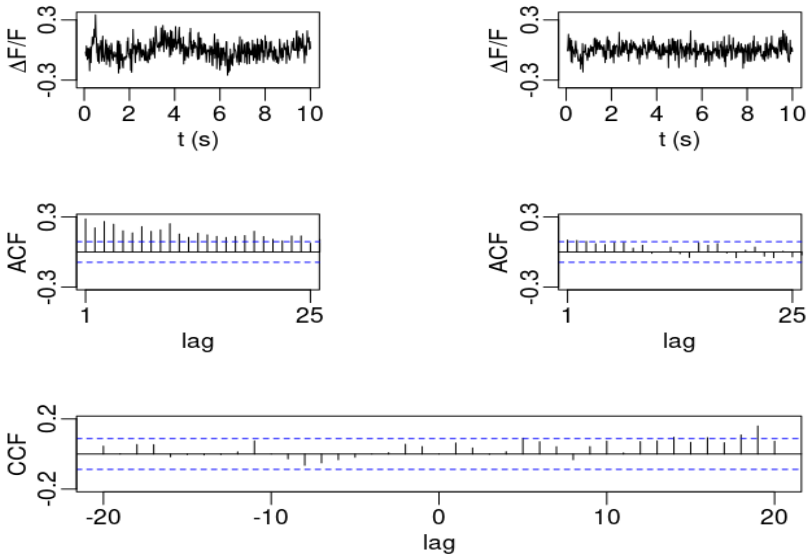


Figure 6.7: Comparison of de-bleached fluorescence intensity between dendritic ROIs b and d (see Figure 6.5), following bleaching controls and normalisation. The top two plots show the de-bleached time series for these two regions while the two centre plots show the ACF evaluated between lags 1 and 25 (d left, b right). The bottom plot is the lagged CCF of the b and d fluorescence time series between lags -20 and 20. The ACF for ROI d is suggestive of periodic effects, while ROI b has multiple marginally significant lags, at the 5% level (blue, dotted line), but no clear structure from the ACF. The CCF has a number of significant lags around lag 20. This could perhaps be indicative of the time taken for signal in ROI d to reach b in the other branch (a time delay of roughly 400ms).

Multivariate Analysis of Neighbouring Dendritic ROIs

Diagnostic plots of shared Ca^{2+} fluorescence dynamics for neighbouring ROIs c and d (after de-bleaching and normalisation) are shown in Figure 6.8. Both mean fluorescence time series are individually non-integrated up to order 17, according to the Augmented Dickey-Fuller test with $p < 0.01$. The AR processes are therefore not co-integrated. While some individual lags are significant in the ACF of ROI c, the Box-Ljung test indicates no significant overall autocorrelation structure ($p = 0.241$), consistent with their being no shifts in Ca^{2+} dynamics (i.e. an event) at this point. On the other hand, cross-correlation lags -4 and 5 are marginally significant ($p < 0.05$) with a number of significant lags around 10 to 15. Although the marginal lags may be indicative of some sharing of voltage conditions across the branch leading to correlated VGCC activity, this could also be due to chance, as we can expect 2 significant lags out of 41 at the 5% level under the null hypothesis of white noise.

The cut-off in significant PACF lags in Figure 6.9 combined with the large number of significant ACF lags in Figure 6.8, are suggestive of such an AR model. Assuming again that there are

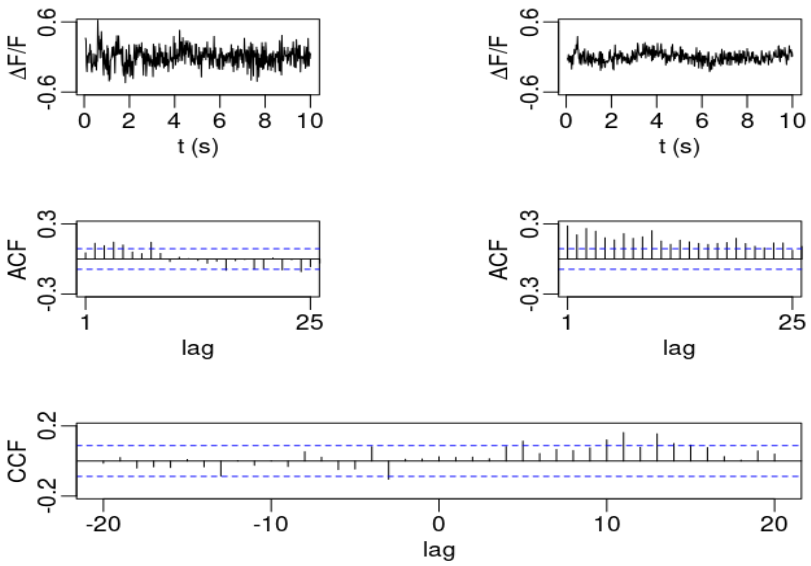
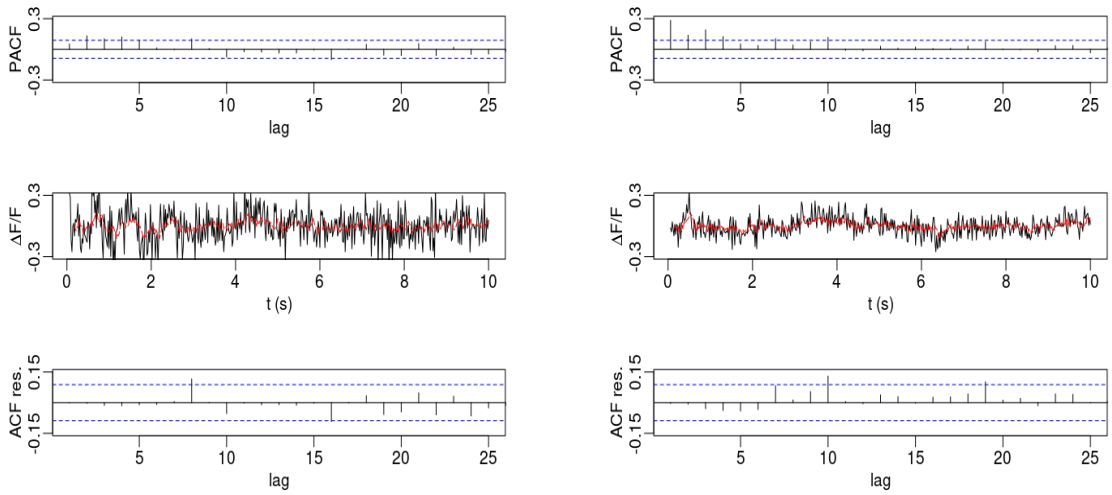


Figure 6.8: Comparison of mean $\Delta F/F$ fluorescence between dendritic ROIs c and d (see Figure 6.5), following bleaching controls and normalisation. The top two plots show the $\Delta F/F$ time series for these two regions while the two centre centre plots show the ACF evaluated between lags 1 and 25 (c left, d right). The bottom plot is the CCF of the b and d fluorescence time series between lags -20 and 20. The ACF for ROI c has multiple significant lags, at the 5% level (blue, dotted line), between 1 and 5. The CCF has a number of significant lags at around lag 11, but also at lower lags -4 and 5. This could be indicative of the sharing of voltage conditions within the branch (leading to the opening of VGCCs) even when no Ca^{2+} event is detectable at c.

no moving average components, in part because of computational constraints, we fitted a VAR model, selecting the number of lags by minimising the AIC. The resulting ACF of the residuals indicate that there is still some temporal correlation left at higher lags (e.g. lag 10 in Figure 6.9a) but not to a significant extent (Box-Ljung, $p > 0.3$). Consistent with the local branch depolarisation hypothesis, we found that the Granger causality test (see Section 6.2.6) used to measure whether the fluorescence at ROI d is a significant predictor of ROI c, is significant with $p \leq 0.01$. Tests of whether signal at c 'Granger causes' d were not significant ($p = 0.186$).

6.3.2 Grid-ROI Analysis

Gridded image normalisation was performed using a slightly different protocol to individual ROIs. First, pixels were binned into non-overlapping 9×9 ($2.7^2\mu m^2$=$24\mu m^2$) grid ROIs and a ROI wide average intensity for each image in the stack was calculated to create a new grid XYT image stack. A gridded version based on the cropped image stack from Figure 6.5 is shown in Figure 6.10a which was generated using a maximum intensity projection (the image formed by



(a) ROI c (b) ROI d

Figure 6.9: This figure shows the results of fitting a Vector Autoregressive (VAR) model to the aggregated, normalised fluorescence time series of neighbouring ROIs c and d. The top two plots are diagnostic plots to determine the number of lags to be fitted. The PACF of ROI c in Figure 6.9a indicate a drop off in significant PACF lags at the 5% level (blue, dotted line) at lag 5 and similarly for ROI d at lag 4 in Figure 6.9b. The AIC was used to select the number of joint lag terms to include and a model with 5 total lags was selected. The middle two plots show the result of fitting the VAR model to the $\Delta F/F$ fluorescence values (black) with fitted values in red. The bottom two plots show the ACF of the residuals for the fitted model for ROIs c and d. The ACF indicates that there is little autocorrelation left in the residuals at specific lags, where we would expect at least one significant residual occurring by chance at the 5% level.

taking the maximum ROI intensity across image stacks for each grid ROI).

A linear detrend of the ROI intensity (Eq. (6.3)) was used to mitigate bleaching effects instead of two exponential fits to background and ROI. This was done because of problems of convergence in the exponential fit (Eq. (6.1)) to certain ROIs on or near the original ROIs in the lower dendritic branch in Figure 6.5. The poor fit could be due to high levels of Ca^{2+} flux in this area, where there are known Ca^{2+} events that could interfere with the fit.

Figure 6.10b shows the blurring effect of linear trend removal on the maximum of the Ca^{2+} fluorescence intensity in each ROI. By examining the result of an Augmented Dickey-Fuller (ADF) test applied to each ROI, it is possible to distinguish some of the ROIs associated with Ca^{2+} events from the rest of the in-field image. This is expected since the arrival of transient Ca^{2+} events likely causes a shift in the mean and variance of the time series.

Comparison of Figure 6.10b and 6.10d suggests that it is possible to recover an outline of the dendrite from passive dendritic dynamics using the magnitude of the neighbour correlation of each ROI. The neighbour correlation is the average (mean) of the absolute Pearson correlation

between the ROI and its 8 nearest neighbours. This plot indicates high levels of local spatial correlation within the dendrite in Ca^{2+} intensity. This is likely due to more than just confinement of Ca^{2+} within the dendrite as this would not account for the spread (see Section 6.1.4). It could also depend on NMDAR, VGCC and internal Ca^{2+} storage activity. The given contrast may provide a way to distinguish dendritic from intracellular regions for further analysis. Note that the local spatial correlation is highest at the site of known Ca^{2+} events and is still high closer to the soma. This could be an example of directional Ca^{2+} dynamics towards the point of voltage signal integration (the soma), which has been noted for some cell types [124].

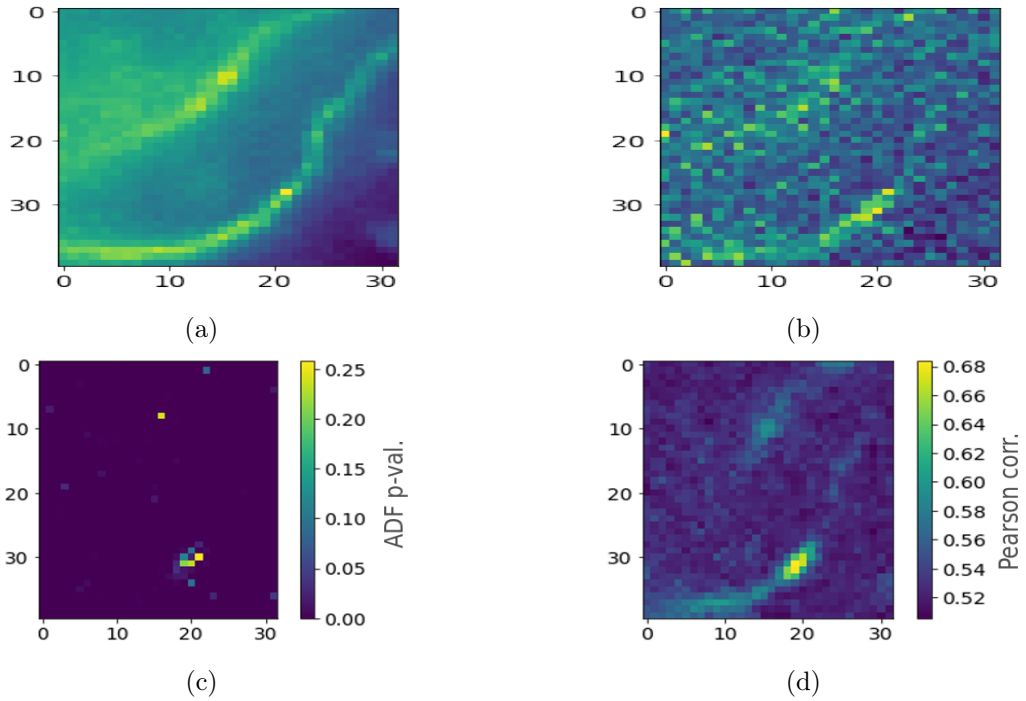


Figure 6.10: The results of gridded image analysis of the cropped XYT stack used in Section 6.3.1 are shown here, where each pixel in the gridded plot represents a volume enclosed by 9×9 pixels over an area of $1.73\mu m^2$ or a projected volume of $278\mu m$. The coordinates of each of the 32×40 ROIs are shown on the X and Y axes. Each point in Figure 6.10a is the temporal maximum of aggregated intensities of pixels in that ROI. The result is a blurred image resembling the cropped image in Figure 6.5 from Haslehurst, 2018 [167]. Figure 6.10b is the maximum intensity in each ROI following linear detrending of intensity time series associated with the ROI. While some structure is present, it is difficult to distinguish active from inactive dendrite or even on-dendrite from extracellular image regions. Figure 6.10c shows the p-values of the Augmented Dickey-Fuller unit root test to each detrended ROI time series. The result suggests that linear detrending is sufficient to produce stationarity, mitigating the majority of bleaching effects. The only areas where p is not significantly low, are areas identified with Ca^{2+} events (i.e. ROIs d-g) and a single ROI on the opposite branch. Pixels in Figure 6.10d give the mean absolute Pearson correlation of the 8 nearest neighbours of each ROI with itself. These spatial correlations are consistent with coordinated Ca^{2+} activity at and below the site of the Ca^{2+} event, moving towards the soma (off image).

6.3.3 Discussion of Preliminary Results

We were limited in the generality of our conclusions by the availability of dynamic Ca^{2+} image stacks. However, the dendritic dynamics analysed exhibits broad spatial and temporal patterns such as local coordination of Ca^{2+} activity at the site of transient Ca^{2+} events. More generally we now have possible guiding principles for distinguishing dendritic from extracellular and active from inactive image regions based on the properties of their Ca^{2+} dynamics. Namely, We found that background, extracellular regions resemble white noise following de-bleaching, while dendritic regions have significant temporal autocorrelation and higher local spatial correlation. In addition, changes in Ca^{2+} fluorescence are consistent with both local branch depolarisation and widespread bidirectional propagation of VGCC activity restricted to this branch.

Our results suggest a possible link between non-stationary unit roots in ROI time series as an indicator of Ca^{2+} events, which could be used to filter ROIs for events to semi-automate manual event detection. In the next section, we propose a simple exploratory method for automated Ca^{2+} event detection (see Section 6.4.5) as well as a novel pipeline for creating morphologically and biophysically realistic neuron simulations from image data. These will be useful in determining what possible recommendations in terms of detection methods or experimental protocols might be useful to future experimenters.

6.4 Experimentation with Automated Event Detection

This section details follow-up experiments that were used to refine a pipeline for generating realistic simulations from XYZ images covered in Section 6.4.2. First, the experimental overview is presented with reference to changes from the first experiment. Notably, these experiments will include 3D structural images of the whole neuron using new transgenic, and thus less phototoxic, Ca^{2+}-sensitive fluorophores that allow for a longer exposure. This analysis pipeline requires the extraction of neuron topology from filtered XYZ images using iterated automated tracing tracing of the neuron.

6.4.1 New Imaging Set-up

In order to capture high quality images with a broad enough FOV to span the dendritic arbour of an entire pyramidal neuron our collaborators have modified the initial lens to emit a curved

beam that is able to fluoresce over a larger area of the slice, at a lower intensity, lowering phototoxicity. The beam produced from this lens, termed an Airy beam lens, needs less light to produce a similar fluoresce intensity [405]. However, additional processing steps must be taken to correct for the curvature of the beam (see Appendix F.4). Even after correction, artefacts persist (see Figure 6.11) that must be corrected in order to accurately locate the neuronal cell body in the image. In addition, this set up may be poorly suited to XYT imaging given the 3D curvature of the beam produces inhomogeneous fluorescence in a single XY plane [405].

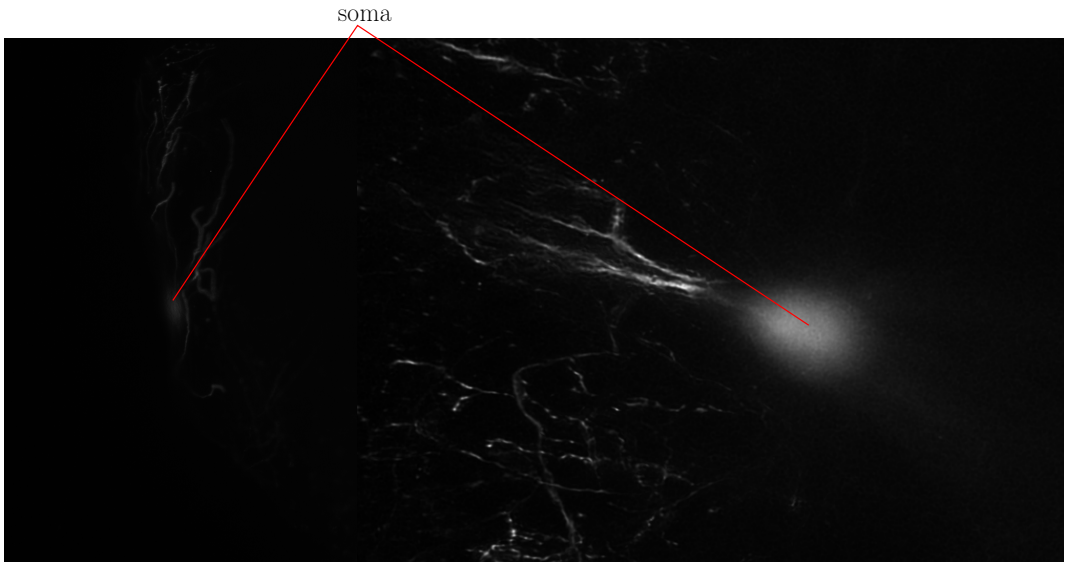


Figure 6.11: This image shows a maximum intensity projection of the soma of two 3D structural images of neurons from early testing data produced by a Light Sheet Microscope (LSM) by taking the maximum pixel intensity from each image in a 3D stack to produce a composite. The imaging technique leads to blur near the soma (or cell body).

Traditional Ca^{2+} indicators must be patch-pipetted into the slice which is time consuming and risks damage to the cell. The second round of pilot images were produced using transgenic Ca^{2+} fluorescent indicators. The intrinsic expression of these indicators makes them more resistant to bleaching, less phototoxic, and less time consuming to prepare and image [297]. Development of this imaging set-up was delayed by COVID-19 and experimental slices require an additional two week incubation period to allow for the sample to become infected with the GCAMP7 transgenic adenovirus [87]. Results utilising the GCAMP7s transgenic Ca^{2+}-binding protein are shown in Section 6.4.5.

Structural XYZ images were produced by pipetting synthetic Alexa Fluor dye directly into the cell after XYT acquisition. We use this single pair of XYT and XYZ images as further pilot

data on which to base our simulation and Ca^{2+} detection methods.

As stated, this set-up has a broader FOV than the original experimental set-up (see Section 6.2.1) with X×Y dimensions of 600μm × 600μm (four times the area of the original), but with the same pixel dimensions of 2048×2048. This broader FOV is what allows for the imaging of whole neurons and thus extraction of the complete neuronal topology (see Section 6.4.3). Exposure was reduced to 10ms and limited to a single 2D image plane (as opposed to a 3D scanning exposure in Section 6.2.1), significantly increasing dampling rate. The lower exposure time allows the initial sharp rise in Ca^{2+} to be resolved more readily, however, it also may reduce effective detection by introducing artefacts.

Structural images were captured using a FOV with the same XY dimensions but a mirror was moved vertically through the slice to fluoresce at depth increments of $0.5\mu m$, passing through the same FOV as the XY image. Individual, long exposure (200ms) images over a $160\mu m$ depth from the surface of the slice were combined to create a single X×Y×Z image stack with dimensions 600μm × Y = 600μm × 160μm) or a pixel resolution of 2048 × 2048 × 320. This resolution is not isotropic (having a lower resolution in Z) which is a problem for most 3D image processing methods, including neuron tracing [432]. This caveat was corrected by standard cubic interpolation which is part of the ImageJ package, Fiji [333]. The resulting cubic voxels (3D pixels) have a volume of $0.29^3\mu m^3$.

Limitations and Advantages of GCAMPXs

GCAMP7s is a highly stable transgenic fluorophore with a high peak fluorescence, increasing the potential signal to noise ratio [145]. It is one of a variety of Ca^{2+} fluorophores belonging to the GCAMP family including GCAMP6s, GCAMP7f and GCAMP6f. Of these two variants GCAMP6s and GCAMP7s are the slower variants with a significantly higher half decay time (the time taken for brightness to reduce by half), for instance the decay time of GCAMP6s is roughly 550ms, compared to 142ms for GCAMP6f [205]. This could make it difficult to effectively isolate single events occurring at a frequency above 1.81Hz, which may occur even in the subthreshold regime, where inputs are more sparse. However, when local input frequencies are low, increased smoothing by GCAMP6/7s may help to make peaks more robust and easy to detect. These slower variants also tend to be brighter in fluorescence.

6.4.2 Second Pilot Imaging Experiments

In the second round of pilot experiments, images were produced in order to perform proof of concept testing of a simple, automated event detection method, as well as neuron tracing and segmentation pipeline (for simulation and analysis). The experimental set-up was similar to the first pilot experiment but used a new Airy beam. After slice extraction and incubation of the GCAMP7s viral vector the slide was prepared and a single GCAMP7s expressing transgenic neuron was located. The imaging plane was selected by scanning through the axial Z direction for an XY plane that had high cross-sectional overlap with the soma and dendrites. No additional K^+ was added to the solution when the XYT image was produced and the neuronal XY plane was imaged for a total of 5 seconds at a rate of 100Hz. After imaging, the XYZ image was deconvolved using a proprietary algorithm designed by M Squared Lasers Ltd. The algorithm is unpublished but Appendix F.4 gives a simple description of the basic principles of deconvolution for light microscopy.

6.4.3 Neuronal Segmentation

Event-like activity can occur in the image background as a result of activity in neighbouring neurons not in the 2D image plane or other sources of random fluctuations in background fluorescence. It is therefore useful to refine event search to only those pixels belonging to the neuron being imaged. One way to do this is to manually trace the neuron using a 2D maximum intensity projection of the 2D Ca^{2+}-dye dynamic image stack. This method is prone to human error due to the level of detail required in tracing and the noisiness of the 2D image but is easy to do with standard imaging software [333]. Alternatively, neurons can be traced from the 3D, higher contrast, XYZ images.

Automatic Neuron Tracing

We used the APP2 automatic neuron tracing algorithm to generate neuronal traces from a 3D image (see Appendix F.3) [432]. In brief, APP2 is a fast-marching, cellular automata-based algorithm which is the flagship algorithm of the vaa3d Open Source suite of 3D image processing tools for neuronal microscopy [292]. The vaa3d package also provides a specialised method for segmenting neuronal spines from background pixels adjacent to dendrites using a similar cellular automata-based algorithm [187]. The resulting neuronal trace is stored as a tree with XYZ

position and radius for each node emanating from the brightest point in the image (usually the soma). The resulting trace includes dendrites and soma, as well as estimated volumes of each of these substructures (important for compartmental simulation of neuronal activity), but excludes the axon which is generally too thin in the image to reliably appear in the trace. We alternate between APP2 tracing and a novel filtering method to identify the soma and remove somatic artefacts from the XYZ image, thereby improving the final trace with APP2.

Segmentation of the Soma and Dendrites

A distinction between the functionally distinct units of the soma and neurites (namely the dendrites) is useful, both to refine the search for Ca^{2+} events (the soma does not have EPSPs but has a high variance in Ca^{2+} activity) and to distinguish biologically distinct cellular components for accurate simulation (see Section 6.5). We adapted a simple segmentation technique based on the Jenk's Natural Breaks method to create this separation of components in a biologically plausible way [194].

In order to segment the neuron tree it is useful to understand how the data is structured. The tree generated by APP2 is stored as a rooted tree list with an entry for each node and a column each for X, Y and Z coordinates, numeric node ID, node type, parent ID and the node radius. For a node, n, in the table with ID k_n, the radius $r_i > 0$ defines the spherical volume of the node. The node ID is a non-negative integer index given by a depth first ordering on the nodes of the tree. The depth first ordering follows the path from the root node to the most distal leaf node (the longest branch). The process is repeated starting from the branch point of the most distal branch off this path. Node IDs are reordered at the end of the process so that the first branch is always the longest branch starting from the root node by depth first search. The type is an integer identifier that is the same for all nodes in the same branch. The root of the tree is the node with the lowest Grey Scale Distance (see Appendix F.3 for details). This node is usually near the geometric centre of the soma which was validated by visual inpection of the XYZ image stack in ImageJ [432].

The root node n_0 has at least two child nodes (see Figure 6.12 below) each associated with different distinct branches. The nodes labels are depth first ordered with n_0 always lying on the longest branch, A (typically containing the central trunk of the apical dendrite and a half or more of the soma). The second longest branch, B, typically contains the basal dendrite and

some somatic nodes (see Figure 6.12). Since A and B are the longest branches of the tree they must necessarily include most if not all of the soma, which is a broader than average component from which all dedritic branches originate.

Moving proximally to distally out from these branches the exact change point from soma to dendrite is not clear cut and a transitional region exists at both ends (the basal portions of A and B). However, for model building and detection methodological simplicity it is useful to define a hard boundary between somatic and dendritic nodes. We achieve this by clustering along the subtree, P, induced by $A \cup B$. Note that P forms a path ordering inherited from the neuron tree and numeric node labels.

Jenk's natural breaks is a simple method for partitioning of an ordered set into N contiguous classes in one dimension [194]. It is related to k-means clustering used in multiple dimsensions. We used the Jenk's framework to partition the nodes of A and B each into two distinct classes using the nodes radii. The two interior classes, with the largest radii then define the soma.

In the Jenk's framework a boundary node C_* is chosen for an ordered list of nodes $C = \{c_1, c_2, ..., c_n\}$, each with radius $w(c) > 0$ for each $c \in C$. C_* is chosen by selecting i which minimises the sum of within class variances

$$var(\{w(c_k)\}_{k=1}^{i}) + var(\{w(c_k)\}_{k=i+1}^{n}), \tag{6.9}$$

so that $C_* = c_i$. We make the modification from the ordinary Jenk's clustering framework by using the unscaled sum of variances in the optimisation step so as to weight somatic and dendritic components equally.

Solving separately for the boundary nodes C_A and C_B for both A and B, the soma is thus defined by the set of nodes $S = \{a_1, ..., C_A, b_1, .., C_B\}$. The segmentation is completed by classifying all other nodes not in S as dendritic.

6.4.4 Artefact Correction for Neuron Tracing

The new LSM setup uses an Airy beam lens. This type of lens is well suited to capturing broad FOV, 3D images. However, bending of the beam through the lens causes artefacts that are not completely removed by the proprietary deconvolution method used to correct such arte-

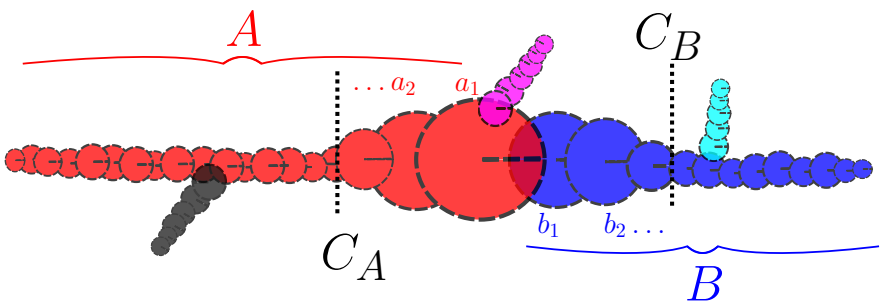


Figure 6.12: 2D representation of the two largest branches connected to the root node of a neuron traced with APP2. The longest branch, A, is shown in red while the shorter, B is in blue. The tips of other minor branches are also shown. Nodes a_i and b_i are shown radiating outwards from the root node (a_1 is the root). Black dotted lines demarcate a plausible Jenk's clustering (based on Jenk's Natural Breaks), dividing the neuron length into 3 classes, with class boundaries C_A and C_B. The soma is the central class $[C_A, C_B]$ while the other two classes are dendritic branches.

facts [405]. For a brief overview of deconvolution see Appendix F.4. Artefacts can lead to errors in tracing where artefactual voxels are misclassified as part of the dendritic tree. Crucially this artefactual blurring occurs near the soma but at a lower intensity allowing the blurred non-somatic pixels to be removed from around the soma by appropriate thresholding.

We propose an augmented tracing procedure which breaks neuronal tracing into an iterative three-step process. First the neuron is traced using an established method e.g. APP2 (see Section 6.4.3). Then the neuron is anatomically segmented into the soma and dendritic components (see Section 6.4.3). After this a cube containing and centred on the soma, but three times its radius is defined. This box contains somatic, artefactual and background voxels. Otsu thresholding, the 3D equivalent of the Jenk's Natural Breaks or k-means clustering methods from Section 6.4.3 is then used to separate somatic voxels from the lower intensity background and artefactual voxels [433]. Non-somatic voxels are then set to zero intensity and the tracing algorithm is rerun on the entire image.

6.4.5 Cummulative Summation Method for Automated Event Detection

Methods are yet to be developed for automatic detection of Ca^{2+} events within cells, either in 2D or 3D Ca^{2+} imaging. A simple approach is to treat the onset of a synaptic Ca^{2+} event as a transient change point, a sudden spike followed by a gradual decline in local neuronal Ca^{2+} concentration relative to recent neuronal activity in a region around the spine and to follow a

CUSUM-based approach to detect the largest Ca^{2+} events.

In Section 6.3, we see that Ca^{2+} events led to a locally coordinated spike in Ca^{2+} activity. We utilise this observation and the 3D neuronal tree extraction pipeline already discussed (see Section 6.4.3) to refine the event search and confine it to only dendritic cellular regions. Confining the search is done by producing a mask from the extracted neuron structure, using the volume of each node to define a neuronal search area. However, since only one viable single XYT Ca^{2+} image stack has been produced for this data set with no paired structural images, we chose to manually trace the neurites in the XY image plane using ImageJ's simple neurite tracer [333], a semi-automated tool for neurite tracing. Only pixels belonging to the neurites in the image are searched using a moving window approach. The window size, $W = 8$, was chosen so that moving windows were 8×8 pixels ($5.84\mu m^2$) in size, as suggested by our collaborators, due to the effective resolution of the imaging set-up. Pixel intensities across the window were spatially averaged to smooth out spatial noise, concentrate signal from correlated areas of activity and produce a time series, p_t for each XY pixel on the neuron at each t time point.

Following neuron segmentation (in accordance with Section 6.4.3) and anatomical labelling, simplified bleaching correction was performed by least squares fitting of Eq. (6.1) to background activity, the pixel-wise mean of temporal activity across all non-neuronal pixels in the XYT image. The event detection algorithm is applied independently for each pixel that is on the neuron (except somatic regions which are excluded).

We then use a CUSUM-based method to determine spatiotemporal event locations. The CUSUM method calculates a change point score s_t at each time point based on previous time points (usually with a moving average or exponentially weighted average offset). CUSUM methods classify as a change point each time where $s_t > b\sigma$ for σ the time series standard deviation and $b = 5$.

Our method, CUSTD is a slight variation on these that includes a local noise penalty that further takes into account any local non-stationarity in the signal (e.g. due to LRD). The event score is calculated recursively from the previous score, with initial score $s_0 = 0$, and subsequent scores

$$s_t = s_{t-1} + \max(0, p_t - m_t - 2\sigma_t) \tag{6.10}$$

where σ_t and m_t are the moving mean and standard deviation respectively, based on the previous $k = 20$ time steps. For $t \leq k$, $m_t = \sigma_t = \infty$ for $t \leq k$ so that events before $t = k$ are effectively undetectable (a shortcoming of the method for large k). Here, b and k were chosen so as to obtain a good fit to the data. In Section 6.6 we will propose a method to fit b and k to the data based on simulation and the free energy.

6.4.6 Calculation of Signal to Noise Ratio

We used a random sample of $N = 5$ events identified by the experimenter to determine a crude baseline signal to noise ratio by subtracting the mean intensity baseline over a 100ms window before the event, z_e, from the peak intensity at that event's location in time and space, m_e. The approximate signal to noise ratio, $S\hat{N}R$, is then

$$\hat{\text{SNR}} = \frac{1}{N}\sum_{e=1}^{N}\frac{m_e}{z_e}. \tag{6.11}$$

Exploratory Results

Figure 6.13a shows the spatial location of 17 events (some overlapping) which were expertly labelled by the experimenter (Peter Haslehurst) as well as a predicted probability density of locating an event according to the CUSTD method. Most events are in high density locations in the image except one that may have been excluded due its proximity to the soma. These events are up to a few tens of microns in extent along the dendrite and so spread in the estimated density is likely due to error. Figure 6.13b shows the scoring of events for a single windowed ROI in the XYT stack. In this example ROI there are two labelled events, at frame 160 and 275. Both of these events were near to two separate actual events in time at the same location (possibly at the same synapse).

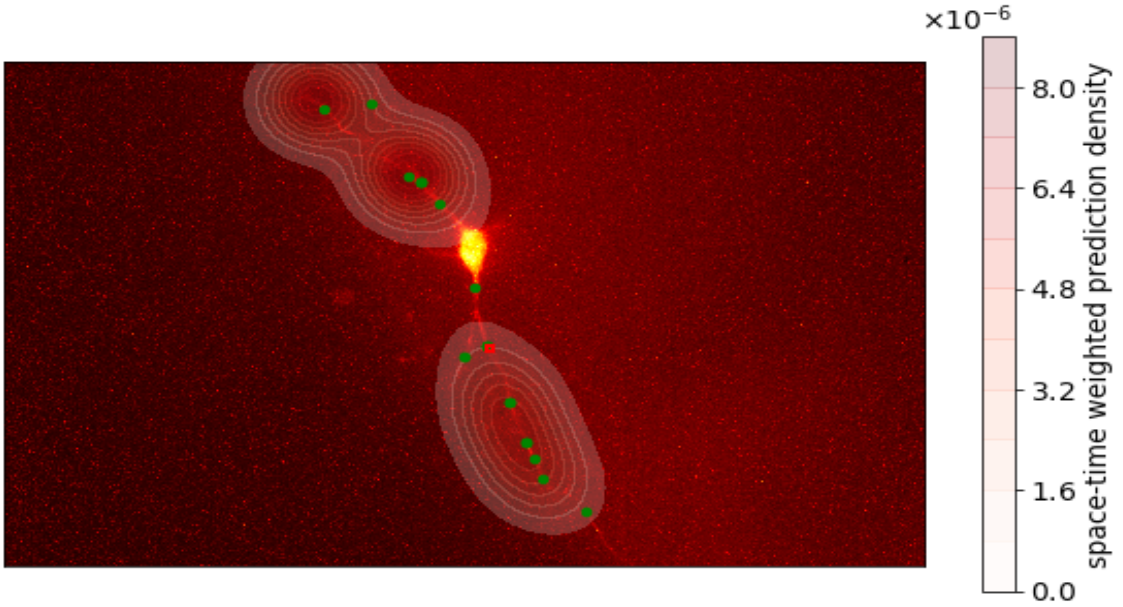
From Eq. (6.11) the approximate signal to noise ratio is 2.5. This is likely to an over estimate of the true signal to noise ratio as it takes into account only events that were sufficiently above the noise level as to be readily identifiable by the experimenter who identified events by eye without validation by the manual detection method in Section 6.2.5.

It is well established that intraneuronal Ca^{2+} activity power scales roughly in a log-log fashion with frequency with scaling exponent α approximately in the range $-3 < \alpha < -0.5$ [47, 297], depending on modality, subcellular location and sub or supratheshold activity regime. We found

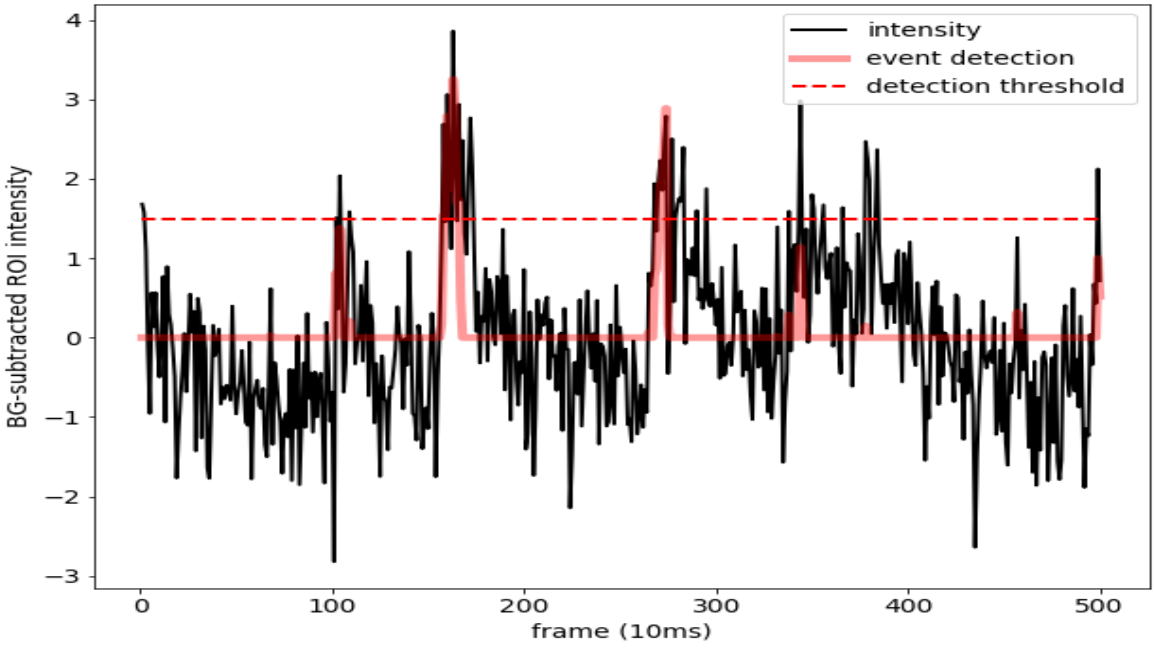
that the Power Spectral Density (PSD) of mean neuronal Ca^{2+} activity followed a strict log-log power law in time over the frequency range of 0.1Hz to 10Hz, with an estimated scaling parameter of $\hat{\alpha} = -2.35$. This result suggests that multiscale temporal LRD is present in the neuronal Ca^{2+} signal, with a large proportion of the signal's power lying in lower frequencies.

As seen in Figure 6.10 Ca^{2+} events in the subthreshold regime can propagate locally in a wave-like pattern. In Figure 6.13a we account for this spread using a simple Gaussian kernel smoothing function in time and space. Ca^{2+} waves may account for some of the spread in Figure 6.13a along the dendrites, however a large proportion of the event density ends up in the extracellular medium which is unrealistic (as events only occur intracellularly).

The problem of estimating the location of events on neurons is non-trivial due to the non-Euclidean, graph-like topology of the neuron and the spreading of the events in time and space along the neuron branches through neuronal integration. In Section 6.6.4 we improve upon this crude estimate using a spatiotemporal graph heat kernel to capture the spread of simulated event probability density in a semi-realistic way. We also investigate the suitability of using a faster fluorophore, such as GCAMP6/7f, in any hypothetical future experiments.



(a)



(b)

Figure 6.13: This figure shows the results of spatiotemporal Ca^{2+} event detection. (a.) Kernel density plot of Ca^{2+} events (red shading), overlayed on a maximum intensity projection of an XYT Ca^{2+} dynamic image stack produced from transgenic GCAMP7s. Event hotspots from 17 total events, labelled by LSM developer and neuroscientist Peter Haslehurst, are shown in green. In order to construct the estimate detected events are weighted by their space-time proximity (Euclidean distance in XYT space) to labelled events. The red box shows the ROI used to produce the plot in (b.). (b.) The plot shows the event detection method in action with the event score at each time point shown in red. When the score crosses a pre-specified threshold, the time point and ROI location are registered as containing an event. Events in neighbouring ROIs at a similar time contribute to the probability of events being present locally, as does the time spent above threshold.

6.5 Simulation

This section details the simulation of Ca^{2+} events based on the previous sections and data. Our simulations of neuronal activity are based on a previous simulation model for pyramidal neurons in mice and rats within the NEURON simulation environment [61,361]. NEURON is a platform optimised for simulating neurons as collections of coupled electrochemical compartments using ordinary differential equations [78]. NEURON allows for realistic simulation of both passive and active neuronal properties including Ca^{2+} events and APs. For full details of the model framework we employ to simulate Ca^{2+} activity see Appendix F.4.1. We use this framework with slight modification in Section 6.7 to model events and then test which of a representative sample of Ca^{2+} event detection methods provides a probability distribution over space and time that best reflects the true event distribution as measured by free energy.

The simulation framework for pyramidal excitatory neurons was still being updated as of February 2021 [251], but requires a number of changes to be used in studying Ca^{2+} dynamics as would be realistically observed using light microscopy. We model these augmentations on the results of the previous sections and add, silent synapses, technical noise caused by the interaction between fluorophore activity and LSM equipment, Ca^{2+} fluorophore signal decay and additional mechanisms for realistic Ca^{2+} exchange (such as Ca^{2+} influx via NMDAR channels). Properties such as bleaching and phototoxicity were not implemented and are outside the scope of this simulation. We utilise a neuron morphology generated from the XYZ image using our pipeline (see Figure 6.14 and Section 6.4.3).

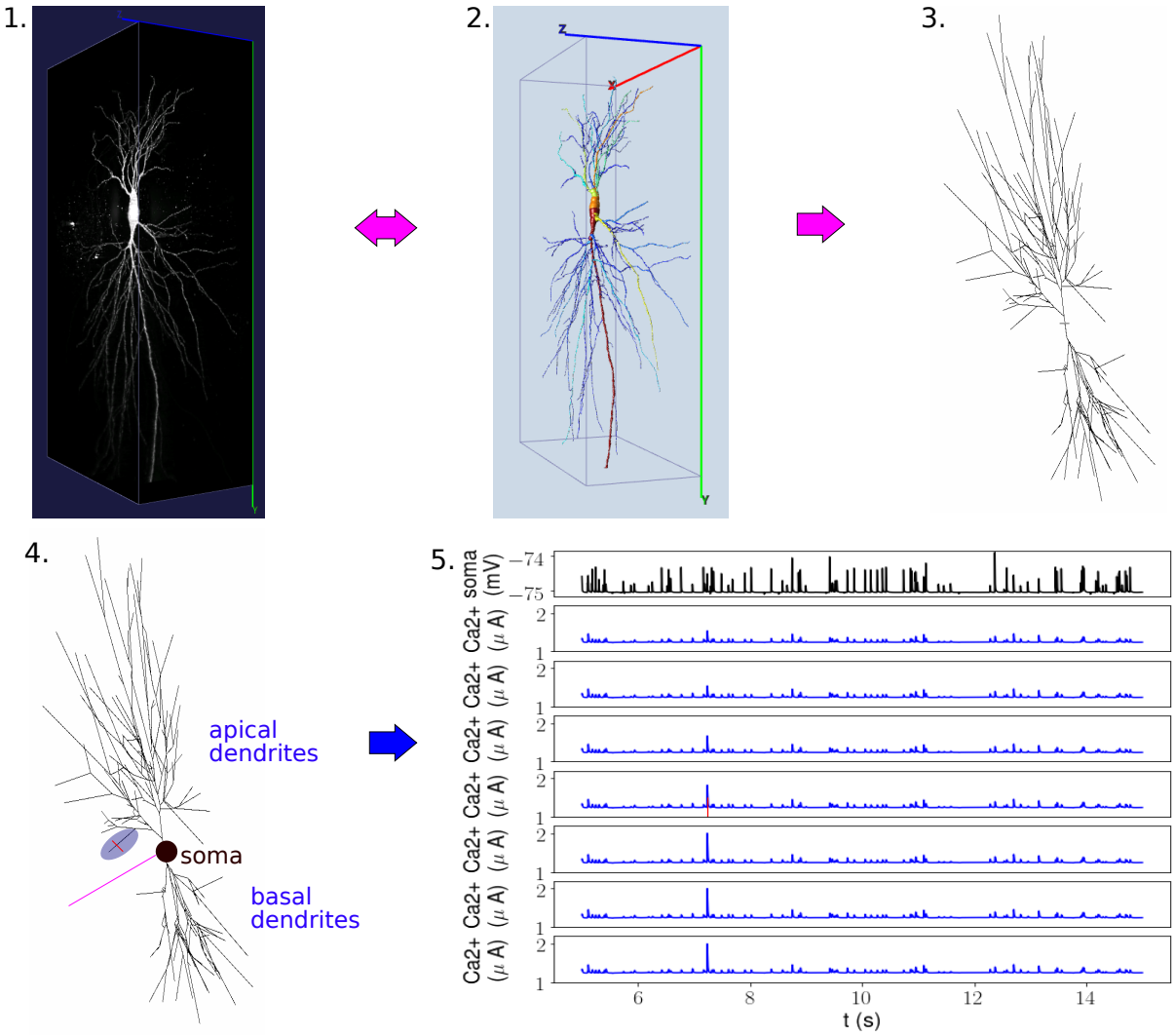


Figure 6.14: Pipeline diagram summarising the entire process for extracting neuronal trees for simulation. 1. Neurons are extracted from background and artefact corrected using the procedure shown in Section 6.4.4. 2. This step requires the neuron to be segmented into anatomical components and then the tracing step is repeated. 3. Next, an abstraction of the neuron topology, as a tree, is produced for simulation. 4. This abstract representation is embellished upon by adding biophysical properties of the soma and apical, and basal dendrites. A simple axon is added for simulation (long straight neurite in purple). 5. Lastly, the neuron is simulated as a collection of functionally distinct compartments (see Appendix F.4.1). The first time series shows the somatic voltage (in mV), remaining subthreshold for the duration of the 10 second simulation. The following figures show Ca^{2+} current flowing into the branch highlighted in blue, with compartments arranged proximal-distally. The simulation has a single Ca^{2+} event at 7 seconds in the fourth branch compartment (shown in red on the neuron and the time series). Note however that the current peaks towards the distal tip, a directional propagation effect of neuronal impedance.

6.5.1 Neuronal Geometry and the Neuronal Tree

The topology and morphology of a single hippocampal pyramidal neuron was extracted following the method in Section 6.4.4. The resulting tree-like topology was then discretised into $N = 320$ cylindrical cellular compartments with a uniform length of 21μm and median width of 2μm (estimated by APP2), including one spherical somatic compartment with a diameter of 15μm. The total number of dendritic branches is 37. A single unbranched axon of 300μm and width 1μm was connected to the soma. The axon remained largely inactive throughout simulation and was excluded from event detection due to its lack of synapses and propagating Ca^{2+} channels. The dendrites and soma can be considered nodes of a rooted tree G (excluding the axon) with the soma as root node with index $1 \in V$, where $V = \mathbb{N}_N$ (see Section 6.4.3).

Biophysical Properties of the Simulation

Ca^{2+} events can be delivered to an individual compartment through pre-programmed synaptic events. Due to computational constraints, we consider the case of only $S = 1000$ active excitatory synapses. In addition we added $S_{inh} = 400$ inhibitory synapses. Lastly we added $S_{sil} = 1000$ silent synapses with no AMPAR channels and NMDAR conductance set to one tenth of normal to account for the lack of AMPAR facilitation. The total of 2000 excitatory synapses is much lower than the roughly 10,000 excitatory synapses present on real neurons, however, previous studies have shown that as few as 100 active spines are sufficient to generate APs in simulation [361], with the distribution and strength of synapses mattering more than their total number.

Each synapse is modelled as a temporal Poisson point process generating inputs at a rate of $r_s = 8 \times 10^{-3} Hz$ events per second for each $s \in \mathbb{N}_S$. This implies an effective event rate of $r = 8Hz$ across the neuron. This is just under the stimulation frequency required to induce LTD in real neurons and below that required to induce sustained AP generation [117].

We assume that synaptic EPSPs and inhibitory signals arrive according to a Poisson process with fixed rate independently across all synapses along the dendrites as in [61]. Despite its established shortcomings [3], the Poisson process is a common simplified model for events arriving at the synapse [337]. The Poisson process assumes a uniform expected arrival time given by r_s where the intervening time of the event preceding the arrival of the i^{th} event, $t_{s,i}$, arriving at any one

synapse s, are independently and exponentially distributed across synapses and in time as,

$$P(t_{s,i}|r_s) = \lambda e^{-r_s t} \quad t_{s,i} \geq 0 \qquad P(t_{s,i}|r_s) = 0 \quad t_{s,i} < 0.$$

On the other hand synapse locations are determined by assuming a similar number of synapses to be uniformly distributed across the length of the dendritic arbour (as in Figure 6.14). Such a distribution is known as a spatial point process, defined over the length of the neuron, while the addition of Poison distributed synaptic events produces a spatio-temporal point process in which events arrive at a uniformly random rate across space and time given the synapse locations. We hold the position of synapses constant after initialisation of the model.

Biological and Experimental Noise

The raw outputs of the NEURON simulation are the Va^{2+} current, $x_{n,t}$, and voltage, $v_{n,t}$, time series for each $n \in V$ and time t. V includes dendritic segment as well as the soma. Given the much slower rate of transgenic fluorophore intensity decay (see Section 6.1.4), we assume that the time constant of intracellular Ca^{2+} depletion by buffering is instantaneous from the perspective of the model so that radial and axial diffusion of Ca^{2+} is negligible. This implies that for the purposes of the simulation,

$$x_{n,t} \propto c_{n,t},$$

where $c_{n,t}$ is the axial intracellular concentration accessible to imaging. We thus use $x_{n,t}$ as a proxy for the Ca^{2+} signal.

Ca^{2+} events are not the only source of variation in the Ca^{2+} signal in the subthreshold regime as fluctuations also occur due to a combination of systematic, biological and fluorophore specific noise. In addition, decay of the brightness of the fluorophore is an important part of modelling the observed fluorescence intensity.

The fluorescence decay rate has a substantial effect on the shape of Ca^{2+} events observed with fluorescence microscopy. In this simulation, we compare the two transgenic fluorophores GCAMP6s and GCAMP6f for which rely decay time estimates are available (see Section 6.4.1). The half decay time for GCAMP6s is 550ms, while for GCAMP6f it is 142ms. Assuming an exponential

decay function, the lifetime d is given by

$$e^{-\frac{\tau_{1/2}}{d}} = \frac{1}{2}, \tag{6.12}$$

where $\tau_{1/2}$ is the half decay time. The resulting values of d are $d_s \approx 793ms$ for GCAMP6s, and $d_f = 204ms$ for GCAMP6f. The decay of the fluorescence intensity has a smoothing effect on $x_{n,t}$ which we account for by smoothing the true Ca^{2+} signal using an AR process of order $p = 1$, giving us the noise free Ca^{2+} fluorescence, $I_{n,t}^{true}$ as

$$I_{n,t}^{true} = x_{n,t} + \beta I_{n,t-1}, \tag{6.13}$$

where $\beta = 1 - \Delta t/d$ for $\Delta t = 10$ms, the sampling frequency.

In order to factor in technical noise into the simulated Ca^{2+} it is important to first define the signal strength, which is related to the SNR as defined in Eq. (6.11). From Section 6.4.6 the estimated $\hat{\text{NSR}} = 0.4$, the inverse of the $\hat{\text{SNR}}$ and δ_n is the average signal strength,

$$\delta_n = \frac{1}{C_n} \sum_{e=1}^{C_n} (m_{n,e} - z_{n,e})$$

where C_n is the total number of events at compartment n. Baseline $z_{n,e}$ and peak intensity $m_{n,e}$ are as in Eq. (6.11). To reflect variability of the experimental equipment and fluorescence over and above the effects of decay we include additive white noise to the true Ca^{2+} fluorescence. This white noise $\epsilon_t \sim N(0, a_n^2)$ has a standard deviation of $a_n = \text{NSR}\delta_n$ (where $\text{NSR} = \text{SNR}^{-1}$). The modelled observed intensity $I_{n,t}$ is then

$$I_{n,t} = I_{n,t}^{true} + \epsilon_t$$

where $\epsilon_t \sim N(0, a_n)$.

6.6 Event Detection and Method Comparison

Ca^{2+} event detection requires methods that can distinguish local transient spikes in Ca^{2+} from the noisy inhomogeneous microscopy signal. We explore two different models for event detection based on the simple CUSTD method from Section 6.4.5 and a deconvolution based method

called Online Active Set method to Infer Spikes (OASIS) usually used for AP detection [139]. These methods generate sparse discrete distributions over event space, the space of possible Ca^{2+} events in space and time. We propose a model evaluation method based on the free energy principle discussed in Section 6.1.7.

Ca^{2+} events can occur in one of N neuronal compartments across T discrete time points in a data set. We call this the event space, $Z = V \times \mathbb{N}_T$, that is defined in space and time. The problem of comparing event detection to data can then be cast as a distributional comparison comparing the true distribution q to the candidate distribution p obtained from the event detector. Typically both p and q are sparse distributions over Z.

Naive comparison of p and q, $|p(z) - q(z)|$ for $z \in Z$, is problematic as when $q(z) > 0$ and $p(z) = 0$ or vice-versa, $|p(z) - q(z)|$ will tend to be high even if $q(z^*) = p(z)$ for some $z^* \in Z$, near to z. Due to the local spread of Ca^{2+} signal, the inaccuracy of detection methods and the sparsity of events in the sub-threshold regime. This is likely and so a possible cause of inaccuracy for model comparison. It is thus useful to consider a smoothing function on space and time that can take into account the non-Euclidean structure of the neuron. We propose such a smoothing kernel and show that it is simple to compute, even for relatively large datasets with high spatial and temporal resolutions (large Z).

6.6.1 Simulation of Ca^{2+} and Voltage Data Sets

We generated a total of 15 ten second simulated outputs, three sets of five each. Each set uses the same randomly generated synapse locations. The first is a training set of five simulations of ten seconds. In addition to the training set, we generate a validation set of 5 simulations each ten seconds in length, assuming GCAMP6s decay dynamics for both training and validation. We also generate 5 simulations using GCAMP6f indicator decay dynamics for comparison of the two fluorophores in terms of their capacity for event detection. The total of 15 simulations were chosen to be in line with the number of Ca^{2+} images that were to be expected from the third phase of the experiments (but which were not produced by our collaborators). The Ca^{2+} and voltage time series across all compartments were simulated at a sampling rate of 10kHz for stability of the integrator but were downsampled to 100Hz to be in line with the imaging rate of the LSM.

6.6.2 Event Detection Methods

We propose two candidate event detection procedures. The first is the CUSTD method we introduce in Section 6.4.5. For this method we use the same initial parameters as in Section 6.4.6, with $k = 20$ and $b = 5$. The second, OASIS, is a more complex method which assumes that events arrive according to an exponential distribution. We will briefly explain the second of these methods as was developed in [139]. For each model we consider the probability mass function over Z to be the probability of an event occurring at a space and time $z \in Z$, given the data, $\{I_z\}_{z \in Z}$.

The OASIS Algorithm

As we have assumed in Section 6.5.1, the OASIS algorithm assumes that in the absence of noise Ca^{2+} fluorescence $\hat{I}_n = (\hat{I}_{n,1}, ..., \hat{I}_{n,T})$, follows an AR process of order $p = 1$. The estimated fluorescence depends on the estimated Ca^{2+} current $x_n = (\hat{x}_{n,1}, ..., x_{N,T})$ x_n can be related by an autoregressive matrix

$$\hat{I}_n = G\hat{x}_n,$$

where G is the lower triangular matrix given by

$$G_{i,j} = \begin{cases} 1 & \text{if } i = j \\ -\gamma & \text{if } j = i - 1 \\ 0 & \text{if otherwise.} \end{cases} \tag{6.14}$$

and γ is the signal decay coefficient which is assumed to be in $\gamma \in (0, 1)$ for stationarity. Since the dominant dynamics are in large part due to decay in the fluorescence intensity [119]. The goal then is to obtain the estimated Ca^{2+} signal $\hat{x}_n$ via Eq. (6.13). OASIS does this by minimising the squared error loss

$$\mathcal{L} = ||\hat{I}_n - I_n||^2,$$

for fixed n and true fluorescence intensity $I_n = (I(n,1), ..., I_{n,T}$, under the constraint that the estimated fluorescence is non-negative,

$$\hat{I}_{n,t} = [G\hat{x}_n]_t > 0,$$

for all t.

The optimisation is performed using the pooled adjacent violators algorithm, which is an active set optimisation method for sampling with incomplete information which is a reasonably efficient algorithm for such sparse optimisation problems [28]. First, we assume that γ can be approximated by β (which is a known property of the fluorophore), giving $\hat{I}_{n,t}$ exactly the same form as the observed fluorescence $I_{n,t}$ according to the simulation model. Later (Section 6.7.2), we test the consistency of the model selection process by arriving at $\hat{\gamma}_{free} = \beta$ through free energy based model selection.

6.6.3 Model Selection by Free Energy Minimisation

Given a distribution q, modelling the uncertainty in event location $(n,t) \in Z = V \times \mathbb{N}_T$, with true distribution p, the free energy is

$$F(q||p) = \sum_{z \in Z} q(z) \log\left(\frac{q(z)}{p(z)}\right) = \sum_{z \in Z} q(z) \log(p(z)) - H(q), \tag{6.15}$$

where $H(q)$ is the entropy of q. Note that Eq. (6.15) is maximised only when $p(z) = q(z)\ \forall z \in Z$. Since p is known from simulation (or potentially from expert labelling), it is theoretically possible to calculate $F(q||p)$ from the detection method. Since we have full knowledge of p (at least in the case of simulation) and q (since we calculate this from the detection results), then

$$FE(q||p) = D_{KL}(q||p), \tag{6.16}$$

where $D_{KL}(q||p)$ is the Kullback-Liebler divergence of q from p, the relative entropy or the information in bits required to compress q using a code generated from p.

Note that if $p(z) = 0$ while $q(z) > 0$ then Eq. (6.15) is undefined while otherwise $q(z)\log(q(z)/p(z)) =$

0 by convention, we therefore require that the support of q contains that of q,

$$p(z) > 0 \implies q(z) > 0 \qquad \forall z \in Z. \tag{6.17}$$

Baseline performance was given by a uniform distribution, u of event occurrence over space and time. This simple prior over Z is described by $u(z) = (NT)^{-1}\ \forall z \in Z$. As a more informative indicator we therefore use

$$FE^*(q||p) = F(q||p) - F(u||p), \tag{6.18}$$

to indicate model performance. This model selection criterion is less than zero when q is at least as informative as u and greater than zero otherwise.

Given that both event detections in Section 6.6.2 detect discrete events and thus are potentially sparse in Z, it is difficult to enforce the condition in Eq. (6.17). An appropriate spatial smoothing function, one that improves on that from Figure 6.13a, by respecting the topology of the neuron tree graph $G = (V, E)$ is needed.

6.6.4 The Multiplex Spatio-temporal Graph Heat Kernel

Let $G = (V, E)$ be the simple graph representing the neuron topology over V and edges E. The heat kernel is the solution to the heat diffusion equations on graphs. From Section 2.4.3, the heat kernel on G, K_G can be written as

$$K_G(\lambda) = e^{-\lambda L_G}, \tag{6.19}$$

where $L_G(\lambda)$ is the combinatorial graph Laplacian of G (see Section 2.1.7), and $\lambda \geq 0$. We term λ the spatial coupling parameter. $K_G(\lambda)$ can be interpreted as an operator on a function $f : V \to \mathbb{R}$ defined over nodes V.

In order to see one way in which G can be extended to the spatiotemporal domain, consider a new graph G_Z, with nodes $z \in Z$ and edge set E_Z. For fixed t^*, $(n, t^*), (n', t^*) \in E_Z$ if and only if $\{n, n'\} \in E$ and each node is connected to at the next time point $\{(n, t-1), (n, t)\} \in E$ for all $t \leq T - 1$. This is exactly the graph Cartesian product of G with the line graph P_T of length

T. Thus, G_Z is an undirected multiplex network of copies of G (see Section 2.3), $G_Z = G \times P_T$. The Laplacian L_Z of G_Z graph has dimension $|Z| = NT$.

By definition, the Laplacian $L_Z = D_Z - A_Z$ (for A_Z the adjacency and D_Z the degree matrices of G_Z) is the spatiotemporal differential operator on G_Z which assumes that information can travel only within layers or between consecutive copies of n and behaves symmetrically in time. The Laplacian L_Z can be alternatively written as the Kronecker sum, $\oplus$ (see Appendix F.5), of L_T (the Laplacian of P_T) and L_G which is given by the sum of Kronecker products, denoted $\otimes$,

$$L_Z = L_T \oplus L_G.$$

As we can control the spatial coupling in G_Z, by scaling the intralayer edges by a factor $\lambda > 0$, we can also scale the temporal coupling in G_Z the temporal coupling parameter between interlayer edges of G_Z, $\tau > 0$. The factor τ is the temporal coupling parameter. The fact that L_Z has on the order of 10^{11} entries for $T = 1001$ time points and $N = 320$ nodes makes the computation of

$$K_Z(\lambda, \tau) = \exp(-(\tau L_T \oplus \lambda L_G)) \tag{6.20}$$

numerically unstable and computationally intractable. We now show that the resulting kernel can be efficiently computed by vectorisation of Eq. (6.20).

Efficient Computation of K_Z

Theorem 1: In the above setting,

$$K_Z(\lambda, \tau)\mathrm{vec}(F) = \mathrm{vec}(K_G(\lambda) F K_T(\tau)),$$

where $K_T(\tau)$ is the heat kernel of P_T, $F : Z \to \mathbb{R}$ is a real valued function (expressed as a matrix) over Z and $\mathrm{vec}(\Psi)$ is the vectorisation of Ψ to a column vector.

Proof. We have

$$\begin{aligned}\mathrm{vec}(K_G(\lambda)FK_T(\tau)) &= K_T(\tau)^T \otimes K_G(\lambda)\mathrm{vec}(F) && (6.21)\\ &= K_T(\tau) \otimes K_G(\lambda)\mathrm{vec}(F) && (6.22)\\ &= \exp(-(\tau L_T \oplus \lambda L_G))\mathrm{vec}(F) && (6.23)\\ &= K_Z(\lambda,\tau)\mathrm{vec}(F). && (6.24)\end{aligned}$$

Eq. (6.21) is a standard property of vectorisation, $\mathrm{vec}(ABC) = (C^T \otimes A)B$ for appropriate matrices A, B and C [358]. Since L_T is symmetric, Eq. (6.22) follows from the symmetry of any constant multiple τ of L_T and its exponential $K_T(\tau)$. Eq. (6.23) follows from another property of the matrix exponential, namely that the Kronecker product of the exponentials is the exponential of the Kronecker sum [327]. Eq. (6.24) follows from the definition of K_Z as the heat kernel of G_Z. □

Theorem 1 implies that kernel K_Z acting on F can be simply rewritten using the much more computationally tractable formulation

$$K_Z(\lambda,\tau)F = K_G(\lambda)FK_T(\tau). \tag{6.25}$$

To see that this represents a major reduction in complexity, it is important to note that the matrix exponential algorithms have at best a worst-case time complexity of $O(n^3)$ for a square matrix of size n [267]. For the naive calculation of $K_Z(\lambda,\tau)$ this presents a complexity of $O(N^3T^3)$. Given that matrix multiplication of the matrices $K_G(\lambda)FK_T(\tau)$ has complexity $O(N^2T+T^2N)$, the worst-case complexity for Eq. (6.25) is $\max\{O(N^2T), O(T^2N), O(T^3), O(N^3)\}$. In most cases we expect $T \gg N$, reflecting the high temporal resolution and imaging time needed to collect enough Ca^{2+} events. The most likely worst-case complexity is thus only limited by time improving its performance over the naive algorithm, $O(T^3) \ll O(N^3T^3)$.

Long Range Behaviour of K_Z

By increasing the coupling constants $\lambda, \tau > 0$ when applied to a probability mass function p over Z we obtain independent smoothing of the probability mass in space and time respectively.

Note that as for L_G, for some F, with all $F_z \geq 0$ and finite sum m over Z, we obtain the limit

$$K_G(\lambda)FK_T(\tau) \xrightarrow[\infty]{\tau} K_G(\lambda)F_T$$

where F_T is the temporal mean of F across all $(n^*, t) \in Z$ for fixed $n^* \in V$. This follows simply from the fact that $K_Z(\lambda, \tau)$ has the same eigenvectors as L_Z, where L_Z has $\mathbf{1}$ as its zero eigenvalues [93]. Lastly,

$$K_G(\lambda)F_T \xrightarrow[\infty]{\lambda} M$$

where $M_{n,t} = \frac{1}{m}$, for the same reason.

6.6.5 Kernel Smoothing of Event Probabilities for Model Selection

As mentioned in Section 6.6.3, kernel smoothing can help to avoid spurious mismatch between two sparse mass functions p^* and q^*. We contend that the multiplex spatiotemporal heat kernel is a reasonable smoothing function on the space of event locations Z when we remain unsure about directional bias in p^* and q^*. The kernel approximates local diffusion of information on the neuron both spatially and temporally. For these reason we apply $K_Z(1, 1)$ to both p^* and q^* prior to calculating the normalised free energy in Eq. (6.15). The final smoothed probability mass functions are

$$p(z) = [K_Z(1, 1)p]_z \qquad \text{and} \qquad q(z) = [K_Z(1, 1)q]_z.$$

6.7 Simulation and Event Detection Results

In this section we explore the results of applying Ca^{2+} detection methods to simulated data. In order to do this we first examine the properties of the model.

6.7.1 Long Range Dependence in Simulated Ca^{2+} Intensity

Figure 6.15 shows the results of a single simulation. In Figure 6.15a, Ca^{2+} intensity appears highly non-stationary with two spiking Ca^{2+} events clearly visible from the signal alone. We include the voltage time series of the same dendritic region, Figure 6.15b, in order to understand

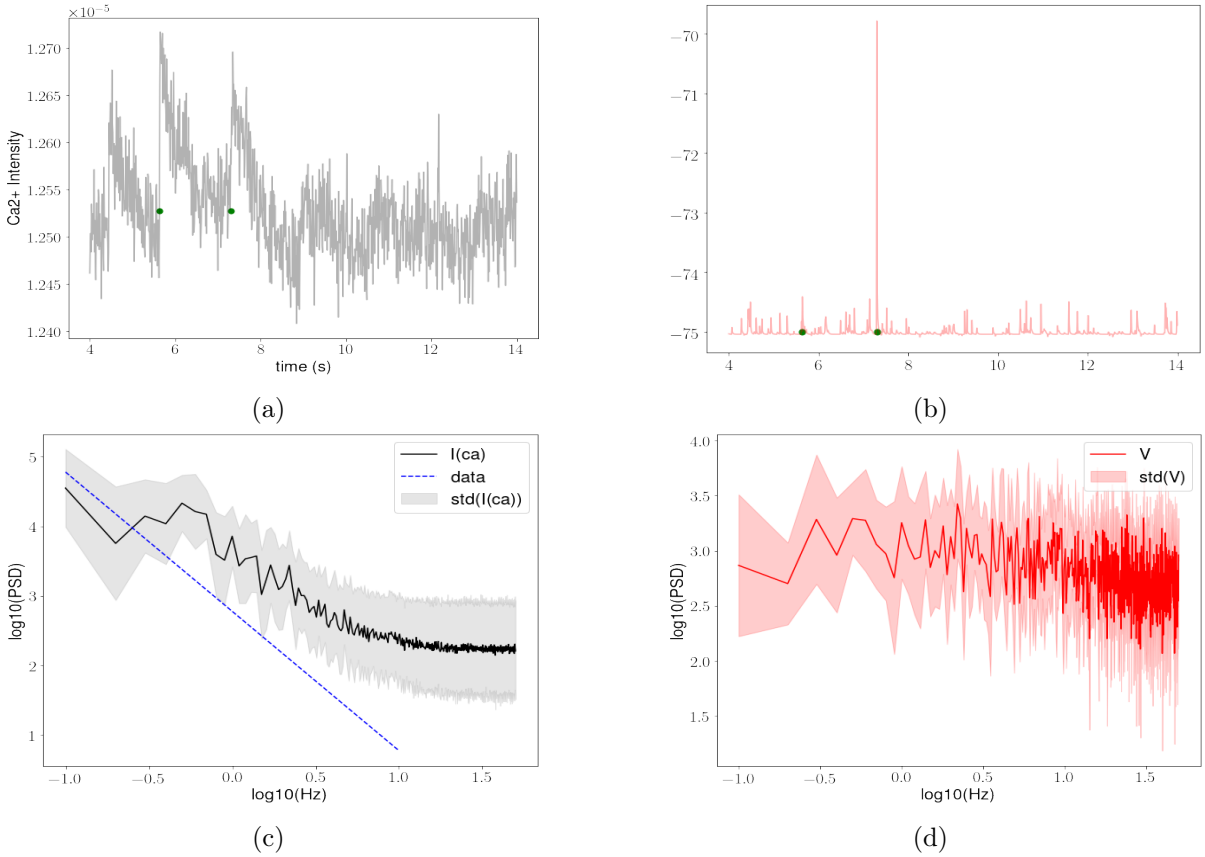
how this relates to Ca^{2+} intensity. It is notable that two spikes are clearly visible in the voltage at the location of the two events but that, even withstanding the noise in Figure 6.15a, the voltage and Ca^{2+} are noticeably different. For instance the largest of the spikes in voltage events does not correspond to that of the Ca^{2+} events and vice versa, suggesting some additional dependencies of the Ca^{2+} signal on the delay between voltage change and Ca^{2+} influx through VGCCs and NMDAR channels.

In Figure 6.15c the log-log plot of the PSD versus frequency clearly shows a degree of power law-like scaling in the Ca^{2+} intensity. Considering the same region as in Section 6.4.6 of $[0.1, 10]$ the resulting scaling exponent is estimated as $\hat{\alpha} = -1.68$. This is lower than that of the true data but still provides evidence of significant power in the lower frequencies of the signal and thus potential for LRD. This dependence may very well be due to the Ca^{2+} indicator itself, as without the exponential smoothing of the indicator, $\alpha \approx 0$. On the other hand, the voltage signal exhibits a much less extreme log-log relationship with $\alpha = -0.14$ over the same interval, suggesting some degree of LRD in the simulated intracellular voltage.

6.7.2 Model Selection for Detection Methods

Direct comparison of the OASIS and CUSTD methods, with event density $o(z)$ and $q(z)$, on the initial set of 5 training simulations with GCAMP6s fluorescence dynamics suggests that both methods are at least as informative as u, the uniform prior on event space Z, with mean normalised free energy of $FE^*(q) = -21.33$ and $FE^*(o) = -27.51$. The discrepancy in performance may in part be due to CUSTD detecting zero events over one entire simulation, with between 0 and 7 events identified per simulation, making it conservative in comparison to OASIS, which identifies a consistent number of events across simulations with mean of 4197 events and standard deviation 281.

Figure 6.16 shows the performance of the two event detectors normalised by the free energy of the uniform distribution. The normalised free energy of CUSTD in the third simulation is zero, meaning that it performs as well as the uniform prior according to the free energy. This is due to zero event detections for CUSTD in simulation 3, as the thresholding in the CUSTD procedure can lead to all or nothing detection of events. In Figure 6.17a, we see that refinement of the CUSTD threshold b on the training set by minimising the free energy can lead to an improvement over the initial model in validation simulations. Overall, both automated models



(a) (b) (c) (d)

Figure 6.15: This figure is summarises the output of a single dynamic Ca^{2+} simulation lasting 10 seconds. Figure 6.15a shows the simulated Ca^{2+} intensity at a single dendritic compartment with two synaptic Va^{2+} events (green dots), which coincides to a transient influx of intracellular Ca^{2+} ions. Figure 6.15b show voltage increases at the same compartment which facilitate the Ca^{2+} uptake which is picked up in the Ca^{2+} signal. Figure 6.15c shows the Power Spectral Density (PSD) of the mean Ca^{2+} intensity across densritic compartments (I(ca)) as a log-log plot (standard deviation in grey). Also shown is the actual log-log linear fit (blue, dotted line) to experimental data of intracellular Ca^{2+}. Figure 6.15d shows a much milder form of long range dependent scaling of power bands in voltage as a function of frequency.

are a better fit to the data than the uniform prior according to the free energy criterion since $FE^* \leq 0$ for both methods except when the event threshold b is very large for CUSTD. In this case $FE^*(q) \approx 0$, as very few events pass the event threshold. The refined choice for the event threshold b, $\hat{b}_{free} = 2.5$ has mean $FE^*(q) = -33.00$ over the validation data set, with mean number of events detected 208 and standard deviation 40. As an aside, we also compared both methods against a simple CUSUM-based method with no noise penalty (σ_t term in Eq. (6.10)) and found that both proposed methods outperformed simple CUSUM which had a significantly larger free energy and a very large number of events per simulation (with more than 14000 events detected per simulation), diffusely distributed in space and time.

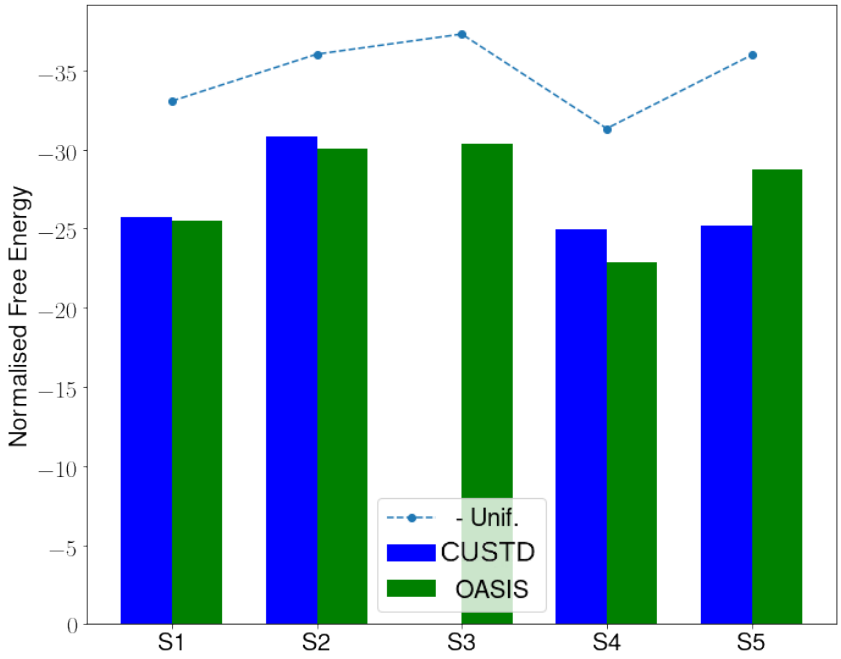
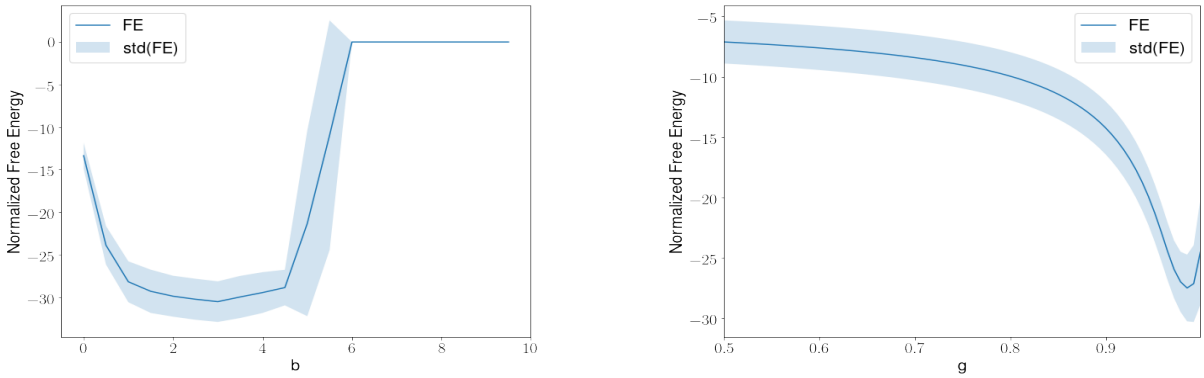


Figure 6.16: This bar graph shows a comparison of the Cumulative Summation with Standard Deviation (CUSTD) and deconvolutional (OASIS) methods across the 5 training simulations. Free energy (a measure of event encoding efficiency), is normalised by the free energy of the uniform distribution on event space. The bars show the normalised free energy for CUSTD (blue) and OASIS (green). The results show that while CUSTD marginally outperforms OASIS in three out of five simulations, its conservative scoring of events leads to one simulation in which no events are detected (S3). The dotted blue line is of the baseline negative free energy of the uniform distribution over event space which depends on the location of events.

Figure 6.17b shows that the normalised free energy minimisation used to estimate the correct AR coefficient in the absence of any information on the decay dynamics of GCAMP6s passed to the model. The model that minimises the normalised free energy has parameter $\hat{\gamma}_{free} = \beta = 0.987$, which is consistent with the true GCAMP6s decay time. The CUSTD model outperforms OASIS after model selection in validation data, the OASIS model had mean $FE^*(o) = -31.72$ so

$FE^*(q) < FE^*(o)$ on average. The CUSTD model with $b = \hat{b}_{free}$ is thus a better choice for detecting Ca^{2+} events than the OASIS model with $\gamma = \hat{\gamma}_{free}$.

Using the free energy optimal parametrisation $b = 2.5$, we examine the distribution of event detection error in neuron space in Figure 6.18. Looking at the absolute difference between the modelled and true event probability mass functions, smoothed using the heat kernel. From the figure it appears that high error regions exist mostly towards the most distal and proximal regions of some dendrites. However, error is low at the branch points themselves. This is likely due to non-linear effects of the impedance and conductance on Ca^{2+} ion flux at the extreme ends of the dendrites, which can contribute to non-linearity in neuronal integration [202].



(a) (b)

Figure 6.17: This figure shows normalised free energy for different parameter values of the CUSTD and OASIS methods based on the data from the 5 training simulations. (a.) The event threshold for CUSTD is varied in the range $b \in [0, 10]$ with the minimum normalised free energy at $\hat{b}_{free} = 2.5$. (b.) For the OASIS method, the AR1 coefficient was varied in the range $\gamma \in [0.1, 1)$. Both parameters show clear minima in the normalised free energy. In the case of OASIS, the model minimising the normalised free energy is the one that takes the true decay time for GCAMP6s ($\hat{\gamma}_{free} = 0.987$).

6.7.3 Comparison of Transgenic Ca^{2+} Indicators

We use normalised free energy with kernel density smoothing to make an important experimental choice, the transgenic Ca^{2+} indicator. Using a model with Ca^{2+} decay dynamics based on GCAMP6f ($d = 142$ms), we train a model using the optimised CUSTD parameters (mean $FE^*(q) = -27.65$). Surprisingly, the preferred choice, based on the recoverability of the Ca^{2+} event distribution is given by GCAMP6s. This could reflect the effect of smoothing out noise that a longer decay time has on the Ca^{2+} signal.

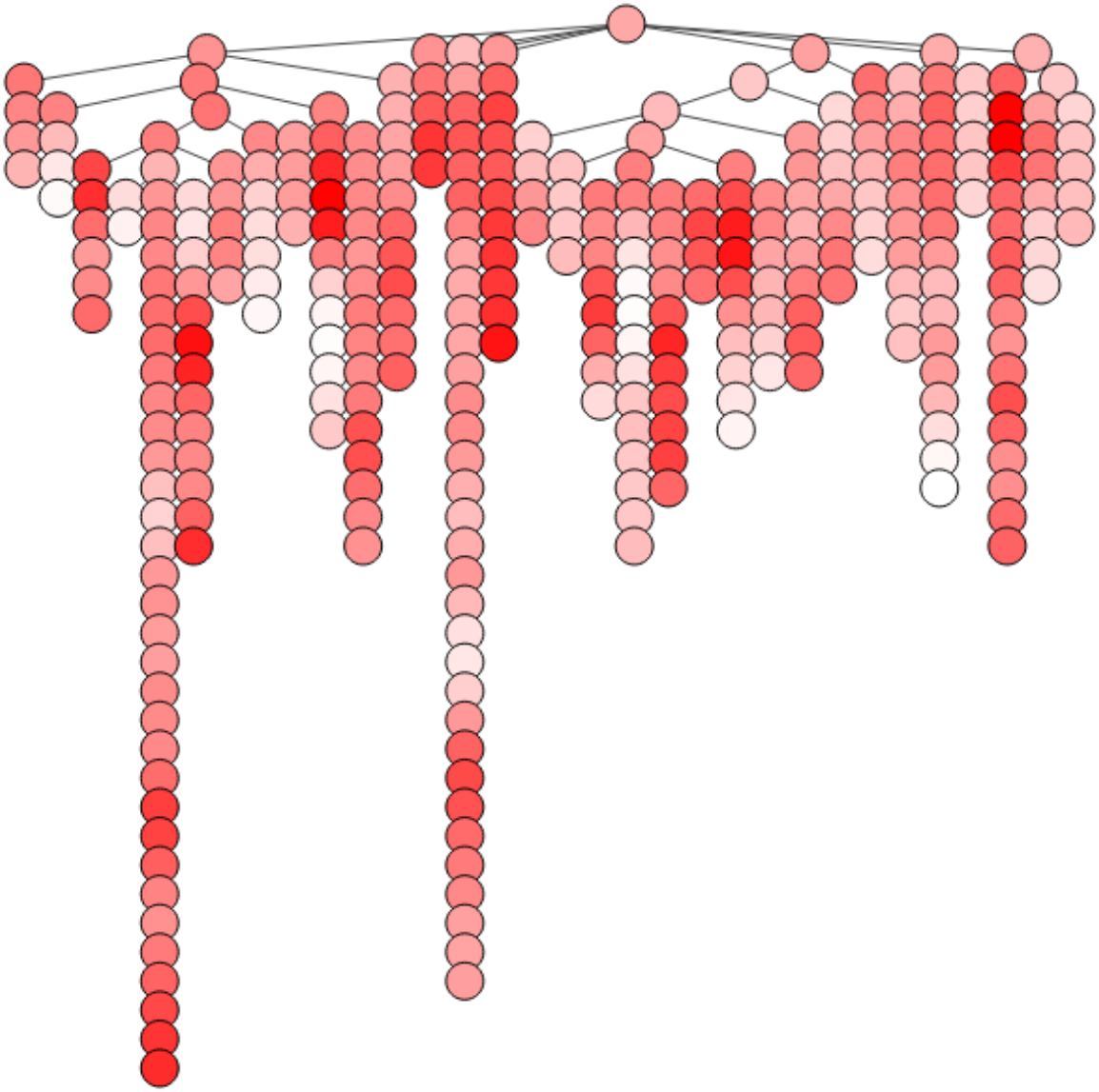


Figure 6.18: This graph is a tree representation of the neuronal topology with nodes for each computational compartment. The red shading is more intense in neuronal compartments where error in event detection is high using the 5 validation simulations. The mean absolute difference between the modelled and true event probability mass functions after smoothing is relatively small with mean 2.110×10^{-4} and standard deviation 9.756×10^{-5}.

6.8 Discussion

Our collaborators were most motivated to understand the complex process of neuronal integration and plasticity. Before AP-induced plasticity can be studied it is important to be able to accurately identify subthreshold integration. However, there is a dearth of modelling and analysis tools for subthreshold single neuronal Ca^{2+} dynamics. We present a novel pipeline for

modelling and analysis of both real and simulated subthreshold transient, neuronal Ca^{2+} events.

We show that LRD is a feature of subcellular neuronal dynamics as observed in light microscopy and simulated voltage signals, connecting to similar patterns seen in macroscale brain activity (see Chapter 5). These effects are well established in simulation [264, 294], however few studies exist that demonstrate LRD in the subthreshold regime in both simulated and real Ca^{2+} data [381]. However, it is important to note that our simulation results suggest that at least some of this power law-like scaling is due to the interaction between the intrinsic Ca^{2+} activity and decay of transgenic Ca^{2+} indicators.

Our novel, graph-based kernelised smoothing approach is appropriate to the neuronal topology and computationally efficient. This kernel could be adapted to examine multiple scales of information diffusion on neurons by varying the spatial and temporal coupling parameters. Consensus information could even be gained in a manner similar to the consensus community clustering in Chapter 5 (Section 5.1.6). This kernel, using ideas from neuroscience and free energy theory, allows us to determine a detection model that best matches the data.

Both OASIS deconvolution and CUSTD automated detection methods far outperformed the baseline uniform prior on events. The method selected by the free energy criterion, CUSTD, may be more tolerant to the long range dependencies and inhomogeneities in neuronal Ca^{2+} intensity than the more sophisticated method, OASIS, that assumes strong background stationarity. Indeed, the brain's own pathways may partially compensate for apparent spatiotemporal non-stationarity in neuron activity by varying the size, density and number of synapses along the dendrite to partially average out response strength along the dendrite [202, 278].

It was noted that none of the methods explored correctly predicted the number of events. This type of overcounting (or undercounting) of events may become an issue when too many (or too few) detected events make it hard to distinguish areas of high or low density. This may be a problem in the case of the basic CUSUM method for which the event counts were very high. However, since the intended purpose of event detection here is to determine regions in time and space with a high probability of events, the number of events is less important than the distributional fit and some over counting is expected in order to increase certainty around a high probability (easily detectable) event or hotspot (area of high event activity).

The model selection procedure is founded in both neuroscientific and statistical theory [140], and

is able to recover the true value (known from biophysical considerations) of model parameters in a constrained parameter search. Based on data from these simulations, we make recommendations for future experiments using transgenic Ca^{2+} indicators to study transient Ca^{2+} events. CUSUM-based methods have been recommended for neuronal event detection in other contexts due to its computational efficiency, relative accuracy, parametric simplicity and potential to be used for online event detection [155, 326]. However, the uniform distribution baseline on events used in our selection procedure fails to take account of the natural clustering of events that tends to occur in real event data. It represents a first pass approach to model selection on neurons; borrowing from other fields, such as neuroimaging may provide future avenues for improvement. In particular, combining both graph-based kernel smoothing and established random field theory-based multiple testing methods could provide firmer frameworks against which to test future models [204].

Although the set of further experiments that this pipeline was meant to be applied to were not completed (step 9 in Figure 6.1), these methods provides a means to conduct rigorous testing of experimental assumptions and methodological choices prior to and in conjunction with costly experimentation. It is therefore a potential aid to experimenters looking for a cost-effective means of hypothesis generation and testing. Light microscopy and in particular light sheet microscopy will require new methods to analyse non-linear subcellular integration for the study of synaptic plasticity and learning. As the quality and ease of acquisition of XYZT images continues to improve, statistics that apply to the unique 3D topology of the neuron will only become more relevant.

7 | Conclusion

Through a general framework of simulating local network exploration and diffusion of information on biological networks we have shown that Long Range Dependent activity is dominant across neurobiological networks and network dynamics. In seeking to apply a broad methodological framework to neurobiological data from a range of systems, protein and living neural networks, and scales, neurons to brains, we have arrived at a number of observations.

7.1 Inspection Bias is an Epistemological Challenge

The state of knowledge in the biomedical sciences is shaped around what researchers find interesting, experimentally accessible and can obtain funding for [67,128]. In this thesis we presented results suggesting that using a metric for PIN functional module selection, that accounts for inspection bias and metaplexity, also performs well when validated on a (mostly) orthogonal transcriptomics data set (see Section 3.3.4). Molecular biology and neuroscience are young and rapidly growing fields with wide-ranging applications in industry and medicine leading to similar research incentives and approaches. It is plausible that metadata in neuroscience and molecular biology would have similar biases [169].

Figure 7.1 shows a log-log plot of neurosynth term frequencies (based on an 11,000 fMRI rest and task study corpus). Possible power-law like scaling is evident ($\hat{\alpha} = -1.91$). The top five terms by frequency involve just a few regions in the DMN and the descriptor 'visual' (which may refer to the visual cortex or visual tasks). This is understandable as many studies focus on vision or visual processing tasks [186], while differences in the activity of the DMN and prefrontal areas are prevalent in behavioural and language neuroimaging studies among many others [214].

While these top brain regions and tasks are likely prominent in the literature due to their importance [243], the possibility that researchers may shape their analyses around these areas while

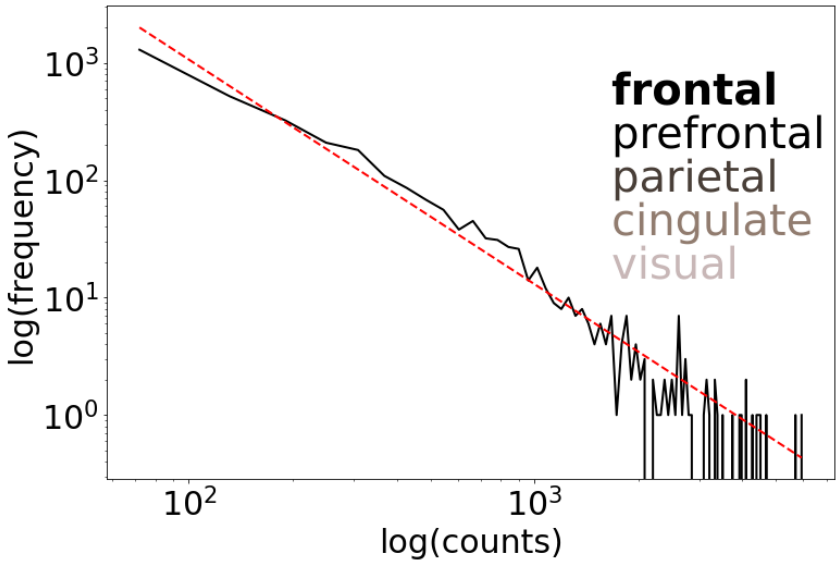


Figure 7.1: This figure shows a log-log plot of the frequency of neurosynth term occurrence (counts) accross the 2017 database of 11,000 rest and task related fMRI studies (excluding terms related to fMRI which occurred in most studies). The dotted red line shows the linear fit to the log-log plot. The scaling parameter is $\hat{\alpha} = -1.91$. Power law-like scaling of neurosynth term frequency shows that across studies certain brain region-related terms tend to dominate studies. The most frequent region-related terms (in the top 20 terms) are frontal, prefrontal, parietal, cingulate and visual, though visual may refer to task type or region (visual cortex), the other regions are all key regions of the DMN.

neglecting others should be considered. For instance, one study which looked at the degree of "corticocentrism" in the literature roughly estimated that between 50% and 80% of studies that reported significant subcortical foci did not comment on them in their analyses [289]. On the other hand, a meta-analysis investigating publication bias (bias towards positive results in the literature) found no significant differences in bias towards any specific frontal areas. However, significant publication bias was found in frontal lobe studies as a whole [195]. If these biases continues to be an issue in the field, meta-analysis based on the literature (of the kind addressed in Chapter 4) may be skewed towards over-represented terms with well defined activity maps (such as cortical and vision terms), in comparison to high uncertainty under-represented terms even when the brain map being queried would be more accurately characterised by the latter.

Ultimately, inspection bias could have far reaching consequences on the state of knowledge in science and poses an epistemological challenge to researchers. Metadata analysis tools are relatively new in neuroscience and an area of active development [137, 436]. As tools that address metaplexity in neuroscience, like neurosynth (under new development in the NiMARE package), grow in maturity, they will need to take positive steps to address potential biases. Our attempt at this in the neuroscientific context has been to adjust for background expectations by sampling

from the local community background with functional activity and connectivity data, using this as an importance filter (Chapter 5) prior or in conjunction with meta-analysis tools (Chapter 4). In the proteomics context, publication data was used directly to score the networks. In the future it could be interesting to assess the underlying biases more directly through a careful analysis of both publication content and citation history for molecular biology as well as neuroscience, however such an analysis is beyond the scope of this thesis [121].

7.2 Discrete Brain State Models are Underdetermined

In Chapter 1 and 4 we defined brain states as configurations of brain activity evoked in response to a stimuli or to facilitate more complex responses. On a practical level this can result in fragmentation of the familiar static network models. However, another issue with this definition is that these responses are often hierarchical with nested levels of subresponses and subprocesses involving interactions between multiple regions [31, 200]. How then can a general concept of brain state be defined?

The solution we have presented to this underspecification problem is to use entropy maximisation to select states and multiplex networks approaches to combine the potentially fragmented functional networks across time. Another solution is to tie brain state dynamics to the dynamics of behaviour. This could be done by training semi-supervised state models that start on a small dataset with timestamped behavioural labels. This may help restrict the initial space of possible states once training is continued on unlabelled data (but limits applicability to cases in which behavioural recordings are possible).

Functional fragmentation may be less of an issue for continuous state space models. Supervised, continuous state space models have recently been developed which use neural network-based frameworks [417]. In general neural network models are difficult to interpret, however these models are an important first step, providing a characterisation of what task-related brain states may look like. This could allow researchers to move away from static network characterisations (such as comparing states to the RSNs). Once more interpretable, continuous and/or semi-supervised brain state models are available, these will be important tools in validating the results of discrete models such as HMGMs and HMMs.

7.3 Non-Stationarity Methods Are Important in Microscopy

In our analysis of intraneuronal synaptic Ca^{2+} spike-like events we find that methods that accounted for the inhomogeneity and Long Range Dependence (LRD) of background activity in active dendrites perform best. This is evident in the superior performance of CUSTD (a method that assumes weak local stationarity) compared to OASIS (a deconvolutional approach that assumes a stationary background). Not only this but LRD is evident in our LSM pilot data and simulations in both the biological and non-biological sources of noise. However, researchers continue to use methods that assume a stationary or trend-stationary level of background activity [84]. This is especially problematic in Ca^{2+} imaging of subcellular activity where indicators have an outsized effect on the decay of background fluctuations [205].

7.4 Long Range Dependence Can Be Beneficial

In this thesis, a common theme has been the ways in which LRD can make it difficult to separate signal from noise. In many cases the definition of signal even becomes hard to parse. However, emerging evidence suggests that the processes of the brain depend on LRD to compute efficiently and construct functional models of the external environment [5]. Researchers should thus be optimistic about the effects of LRD and the prospect of harnessing it to improve our understanding of living systems.

References

[1] Nimare. `hhttps://nimare.readthedocs.io/en/latest/about.html`, November 2019. (Accessed on 10/09/2021).

[2] The gene ontology resource: enriching a gold mine. *Nucleic Acids Research*, 49(D1):D325–D334, 2021.

[3] A. Abenavoli, L. Forti, and A. Malgaroli. Mechanisms of spontaneous miniature activity at ca3-ca1 synapses: evidence for a divergence from a random poisson process. *The Biological Bulletin*, 199(2):184–186, 2000.

[4] L. Acciai, P. Soda, and G. Iannello. Automated neuron tracing methods: an updated account. *Neuroinformatics*, 14(4):353–367, 2016.

[5] K. Aghababaiyan. Improving performance of neurons by adding colour noise. *IET Nanobiotechnology*, 14(5):433–439, 2020.

[6] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann. Link communities reveal multiscale complexity in networks. *Nature*, 466(7307):761, 2010.

[7] E. M. Airoldi, D. M. Blei, S. E. Fienberg, E. P. Xing, and T. Jaakkola. Mixed membership stochastic block models for relational data with application to protein-protein interactions. In *Proceedings of the International Biometrics Society Annual Meeting*, volume 15, 2006.

[8] T. Aittokallio and B. Schwikowski. Graph-based methods for analysing networks in cell biology. *Briefings in Bioinformatics*, 7(3):243–255, 2006.

[9] R. Albert. Scale-free networks in cell biology. *Journal of Cell Science*, 118(21):4947–4957, 2005.

[10] K. Albus, A. Wahab, and U. Heinemann. Standard antiepileptic drugs fail to block epileptiform activity in rat organotypic hippocampal slice cultures. *British Journal of Pharmacology*, 154(3):709–724, 2008.

[11] A. Alexa and J. Rahnenführer. topGO: Enrichment analysis for gene ontology. *R Package*, 2(0), 2010.

[12] A. Alexa, J. Rahnenführer, and T. Lengauer. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13):1600–1607, 2006.

[13] S. Alford, B. Frenguelli, J. Schofield, and G. Collingridge. Characterization of Ca^{2+} signals induced in hippocampal CA1 neurones by the synaptic activation of NMDA receptors. *Journal of Physiology*, 469(1):693–716, 1993.

[14] W. Ali, C. Deane, and G. Reinert. Protein interaction networks and their statistical analysis. *Handbook of Statistical Systems Biology*, pages 200–234, 2011.

[15] H. J. Alitto and W. M. Usrey. Corticothalamic feedback and sensory processing. *Current Opinion in Neurobiology*, 13(4):440–445, 2003.

[16] P. Allegrini, D. Menicucci, R. Bedini, L. Fronzoni, A. Gemignani, P. Grigolini, B. J. West, and P. Paradisi. Spontaneous brain activity as a source of ideal 1/f noise. *Physical Review E*, 80(6):061914, 2009.

[17] M. Altaf-Ul-Amin, F. M. Afendi, S. K. Kiboi, and S. Kanaya. Systems biology in the context of big data and networks. *BioMed Research International*, 2014, 2014.

[18] J. M. Amatrudo, J. P. Olson, H. K. Agarwal, and G. C. Ellis-Davies. Caged compounds for multichromic optical interrogation of neural systems. *European Journal of Neuroscience*, 41(1):5–16, 2015.

[19] M. Andreasen and J. D. Lambert. Factors determining the efficacy of distal excitatory synapses in rat hippocampal CA1 pyramidal neurones. *Journal of Physiology*, 507(2):441–462, 1998.

[20] J. R. Andrews-Hanna. The brain's default network and its adaptive role in internal mentation. *The Neuroscientist*, 18(3):251–270, 2012.

[21] J. R. Andrews-Hanna, J. Smallwood, and R. N. Spreng. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, 1316(1):29, 2014.

[22] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan. An introduction to mcmc for machine learning. *Machine Learning*, 50(1):5–43, 2003.

[23] S. J. Ansley, J. L. Badano, O. E. Blacque, J. Hill, B. E. Hoskins, C. C. Leitch, J. C. Kim, A. J. Ross, E. R. Eichers, T. M. Teslovich, et al. Basal body dysfunction is a likely cause of pleiotropic Bardet–Biedl syndrome. *Nature*, 425(6958):628, 2003.

[24] A. Araque, V. Parpura, R. P. Sanzgiri, and P. G. Haydon. Tripartite synapses: glia, the unacknowledged partner. *Trends in Neurosciences*, 22(5):208–215, 1999.

[25] C. N. Areshenkoff, J. Y. Nashed, R. M. Hutchison, M. Hutchison, R. Levy, D. J. Cook, R. S. Menon, S. Everling, and J. P. Gallivan. Muting, not fragmentation, of functional brain networks under general anesthesia. *NeuroImage*, 231:117830, 2021.

[26] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, et al. Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25, 2000.

[27] K. Avrachenkov, N. Litvak, M. Sokol, and D. Towsley. Quick detection of nodes with large degrees. *Internet Mathematics*, 10(1-2):1–19, 2014.

[28] M. Ayer, H. D. Brunk, G. M. Ewing, W. T. Reid, and E. Silverman. An empirical distribution function for sampling with incomplete information. *The Annals of Mathematical Statistics*, pages 641–647, 1955.

[29] T. Aynaud, V. D. Blondel, J.-L. Guillaume, and R. Lambiotte. *Graph Partitioning*. John Wiley & Sons, Inc., Feb 2013.

[30] Y. Bai, J. He, X. Xia, Y. Wang, Y. Yang, H. Di, X. Li, and U. Ziemann. Spontaneous transient brain states in EEG source space in disorders of consciousness. *NeuroImage*, 240:118407, 2021.

[31] D. H. Ballard. *Brain computation as hierarchical abstraction*. MIT Press, 2015.

[32] N. Bannister and A. Larkman. Dendritic morphology of CA1 pyramidal neurones from the rat hippocampus: Ii. spine distributions. *Journal of Comparative Neurology*, 360(1):161–171, 1995.

[33] A.-L. Barabási. Scale-free networks: a decade and beyond. *Science*, 325(5939):412–413, 2009.

[34] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[35] A.-L. Barabási and E. Bonabeau. Scale-free networks. *Scientific American*, 288(5):60–69, 2003.

[36] A.-L. Barabasi and Z. N. Oltvai. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 5(2):101–113, 2004.

[37] G. Bardella, A. Bifone, A. Gabrielli, A. Gozzi, and T. Squartini. Hierarchical organization of functional connectivity in the mouse brain: a complex network approach. *Scientific Reports*, 6(1):1–11, 2016.

[38] G. P. Basharin, A. N. Langville, and V. A. Naumov. The life and work of A.A. markov. *Linear algebra and its applications*, 386:3–26, 2004.

[39] D. S. Bassett and E. T. Bullmore. Small-world brain networks revisited. *The Neuroscientist*, 23(5):499–516, 2017.

[40] D. S. Bassett and O. Sporns. Network neuroscience. *Nature Neuroscience*, 20(3):353–364, 2017.

[41] V. Batagelj and A. Mrvar. Pajek data: Yeast. `http://vlado.fmf.uni-lj.si/pub/networks/data/bio/yeast/yeast.htm`, 2008. (Accessed on 01/07/2022).

[42] M. J. Beal. *Variational algorithms for approximate Bayesian inference*. University College London, 2003.

[43] M. J. Beal, Z. Ghahramani, and C. E. Rasmussen. The infinite hidden markov model. In *Advances in neural information processing systems*, pages 577–584, 2002.

[44] J. M. Bekkers. Pyramidal neurons. *Current Biology*, 21(24):R975, 2011.

[45] P. T. Bell and J. M. Shine. Subcortical contributions to large-scale network communication. *Neuroscience & Biobehavioral Reviews*, 71:313–322, 2016.

[46] R. Bell, A. Hubbard, R. Chettier, D. Chen, J. P. Miller, P. Kapahi, M. Tarnopolsky, S. Sahasrabuhde, S. Melov, and R. E. Hughes. A human protein interaction network shows conservation of aging processes between human and invertebrate species. *PLoS genetics*, 5(3):e1000414, 2009.

[47] T. Bellay, A. Klaus, S. Seshadri, and D. Plenz. Irregular spiking of pyramidal neurons organizes as scale-invariant neuronal avalanches in the awake state. *eLife*, 4:e07224, 2015.

[48] Y. Ben-Ari and M. Gho. Long-lasting modification of the synaptic properties of rat ca3 hippocampal neurones induced by kainic acid. *Journal of Physiology*, 404(1):365–384, 1988.

[49] J. Beran. Statistical methods for data with long-range dependence. *Statistical Science*, pages 404–416, 1992.

[50] T. Berggård, S. Linse, and P. James. Methods for the detection and analysis of protein–protein interactions. *Proteomics*, 7(16):2833–2842, 2007.

[51] N. Berline, E. Getzler, and M. Vergne. *Heat kernels and Dirac operators*. Springer Science & Business Media, 2003.

[52] L. K. Bicks, K. Yamamuro, M. E. Flanigan, J. M. Kim, D. Kato, E. K. Lucas, H. Koike, M. S. Peng, D. M. Brady, S. Chandrasekaran, et al. Prefrontal parvalbumin interneurons require juvenile social experience to establish adult social behavior. *Nature Communications*, 11(1):1–15, 2020.

[53] D. Binns, E. Dimmer, R. Huntley, D. Barrell, C. O'donovan, and R. Apweiler. Quickgo: a web-based tool for gene ontology searching. *Bioinformatics*, 25(22):3045–3046, 2009.

[54] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*, 34(4):537–541, 1995.

[55] Å. K. Björklund, S. Light, L. Hedin, and A. Elofsson. Quantitative assessment of the structural bias in protein–protein interaction assays. *Proteomics*, 8(22):4657–4667, 2008.

[56] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.

[57] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308, 2006.

[58] M. Boly, S. Sasai, O. Gosseries, M. Oizumi, A. Casali, M. Massimini, and G. Tononi. Stimulus set meaningfulness and neurophysiological differentiation: a functional magnetic resonance imaging study. *PloS One*, 10(5):e0125337, 2015.

[59] J. A. Bondy, U. S. R. Murty, et al. *Graph theory with applications*, volume 290. Macmillan London, 1976.

[60] V. Bonhomme, C. Staquet, J. Montupil, A. Defresne, M. Kirsch, C. Martial, A. Vanhaudenhuyse, C. Chatelle, S. K. Larroque, F. Raimondo, et al. General anesthesia: a probe to explore consciousness. *Frontiers in Systems Neuroscience*, 13:36, 2019.

[61] T. Branco, B. A. Clark, and M. Häusser. Dendritic discrimination of temporal input sequences in cortical neurons. *Science*, 329(5999):1671–1675, 2010.

[62] U. Brandes, D. Delling, M. Gaertler, R. Gorke, M. Hoefer, Z. Nikoloski, and D. Wagner. On modularity clustering. *IEEE Transactions on Knowledge and Data Engineering*, 20(2):172–188, 2008.

[63] T. S. Breusch and A. R. Pagan. A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, pages 1287–1294, 1979.

[64] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1-7):107–117, 1998.

[65] A. Brockwell. Likelihood-based analysis of a class of generalized long-memory time series models. *Journal of Time Series Analysis*, 28(3):386–407, 2007.

[66] A. D. Broido and A. Clauset. Scale-free networks are rare. *Nature Communications*, 10(1):1–10, 2019.

[67] L. Bromham, R. Dinnage, and X. Hua. Interdisciplinary research has consistently lower funding success. *Nature*, 534(7609):684–687, 2016.

[68] R. Brown. What is a brain state? *Philosophical Psychology*, 19(6):729–742, 2006.

[69] A. Brudno. Entropy and the complexity of the trajectories of a dynamic system. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 44:124–149, 1982.

[70] D. Bu, Y. Zhao, L. Cai, H. Xue, X. Zhu, H. Lu, J. Zhang, S. Sun, L. Ling, N. Zhang, et al. Topological structure analysis of the protein–protein interaction network in budding yeast. *Nucleic Acids Research*, 31(9):2443–2450, 2003.

[71] Y. Burak and I. R. Fiete. Fundamental limits on persistent activity in networks of noisy neurons. *Proceedings of the National Academy of Sciences*, 109(43):17645–17650, 2012.

[72] Z. Burda, J. Duda, J.-M. Luck, and B. Waclaw. Localization of the maximal entropy random walk. *Physical review letters*, 102(16):160602, 2009.

[73] G. Busetto, M. J. Higley, and B. L. Sabatini. Developmental presence and disappearance of postsynaptically silent synapses on dendritic spines of rat layer 2/3 pyramidal neurons. *Journal of Physiology*, 586(6):1519–1527, 2008.

[74] Y. Buskila, P. P. Breen, J. Tapson, A. Van Schaik, M. Barton, and J. W. Morley. Extending the viability of acute brain slices. *Scientific Reports*, 4:5309, 2014.

[75] G. Buzsaki, M. Penttonen, Z. Nadasdy, and A. Bragin. Pattern and inhibition-dependent invasion of pyramidal cell dendrites by fast spikes in the hippocampus in vivo. *Proceedings of the National Academy of Sciences*, 93(18):9921–9925, 1996.

[76] V. D. Calhoun and T. Adali. Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE Reviews in Biomedical Engineering*, 5:60–73, 2012.

[77] V. D. Calhoun, T. Adalı, and J. J. Pekar. A method for comparing group fMRI data using independent component analysis: application to visual, motor and visuomotor tasks. *Magnetic Resonance Imaging*, 22(9):1181–1191, 2004.

[78] N. T. Carnevale and M. L. Hines. *The NEURON book*. Cambridge University Press, 2006.

[79] V. S. Caviness, J. Meyer, N. Makris, and D. N. Kennedy. MRI-based topographic parcellation of human neocortex: an anatomically specified method with estimate of reliability. *Journal of Cognitive Neuroscience*, 8(6):566–587, 1996.

[80] D. J. Chalmers. Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3):200–219, 1995.

[81] C. L. Chapman, J. J. Wright, and P. D. Bourke. Spatial eigenmodes and synchronous oscillation: co-incidence detection in simulated cerebral cortex. *Journal of Mathematical Biology*, 45(1):57–78, 2002.

[82] C. J. Charpentier, P. Faulkner, E. R. Pool, V. Ly, M. S. Tollenaar, L. M. Kluen, A. Fransen, Y. Yamamori, N. Lally, A. Mkrtchian, et al. How representative are neuroimaging samples? large-scale evidence for trait anxiety differences between fMRI and behaviour-only research participants. *Social Cognitive and Affective Neuroscience*, 2021.

[83] A. Chatr-Aryamontri, B.-J. Breitkreutz, R. Oughtred, L. Boucher, S. Heinicke, D. Chen, C. Stark, A. Breitkreutz, N. Kolas, L. O'donnell, et al. The BioGRID interaction database: 2015 update. *Nucleic Acids Research*, 43(D1):D470–D478, 2014.

[84] C. E. Cheetham, M. S. Hammond, C. E. Edwards, and G. T. Finnerty. Sensory experience alters cortical connectivity and synaptic function site specifically. *Journal of Neuroscience*, 27(13):3456–3465, 2007.

[85] J. Chen and B. Yuan. Detecting functional modules in the yeast protein–protein interaction network. *Bioinformatics*, 22(18):2283–2290, 2006.

[86] M. Chen, K. Kuzmin, and B. K. Szymanski. Community detection via maximization of modularity and its variants. *IEEE Transactions on Computational Social Systems*, 1(1):46–65, 2014.

[87] T.-W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, 499(7458):295–300, 2013.

[88] C. Cherniak, M. Changizi, et al. Large-scale optimization of neuron arbors. *Physical Review E*, 59(5):6001, 1999.

[89] S.-Y. Cho, S.-G. Park, D.-H. Lee, and B.-C. Park. Protein-protein interaction networks: from interactions to networks. *BMB Reports*, 37(1):45–52, 2004.

[90] Y.-R. Cho, W. Hwang, M. Ramanathan, and A. Zhang. Semantic integration to identify overlapping functional modules in protein interaction networks. *BMC Bioinformatics*, 8(1):265, 2007.

[91] A. W. Chung, E. Pesce, R. P. Monti, and G. Montana. Classifying HCP task-fMRI networks using heat kernels. In *2016 International Workshop on Pattern Recognition in NeuroImaging (PRNI)*, pages 1–4. IEEE, 2016.

[92] F. Chung. The heat kernel as the pagerank of a graph. *Proceedings of the National Academy of Sciences*, 104(50):19735–19740, 2007.

[93] K. D. Cole, M. R. Yavari, and P. K. Rao. Computational heat transfer with spectral graph theory: Quantitative verification. *International Journal of Thermal Sciences*, 153:106383, 2020.

[94] G. Collell and J. Fauquet. Brain activity and cognition: a connection from thermodynamics and information theory. *Frontiers in Psychology*, 6:818, 2015.

[95] J. A. Connor. Digital imaging of free calcium changes and of spatial gradients in growing processes in single, mammalian central nervous system cells. *Proceedings of the National Academy of Sciences*, 83(16):6179–6183, 1986.

[96] R. P. Costa, B. E. Mizusaki, P. J. Sjöström, and M. C. van Rossum. Functional consequences of pre-and postsynaptic expression of synaptic plasticity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1715):20160153, 2017.

[97] C. Croft, H. Futch, B. Moore, and T. Golde. Organotypic brain slice cultures to model neurodegenerative proteinopathies. *Molecular Neurodegeneration*, 14(1):1–11, 2019.

[98] Y. Cui, M. Cai, and H. E. Stanley. Discovering disease-associated genes in weighted protein–protein interaction networks. *Physica A: Statistical Mechanics and its Applications*, 496:53–61, 2018.

[99] M. E. Cusick, H. Yu, A. Smolyar, K. Venkatesan, A.-R. Carvunis, N. Simonis, J.-F. Rual, H. Borick, P. Braun, M. Dreze, et al. Literature-curated protein interaction datasets. *Nature Methods*, 6(1):39, 2009.

[100] J. Das and H. Yu. HINT: High-quality protein interactomes and their applications in understanding human disease. *BMC Systems Biology*, 6(1):92, 2012.

[101] M. Dasilva, A. Camassa, A. Navarro-Guzman, A. Pazienti, L. Perez-Mendez, G. Zamora-López, M. Mattia, and M. V. Sanchez-Vives. Modulation of cortical slow oscillations and complexity across anesthesia levels. *NeuroImage*, 224:117415, 2021.

[102] O. David, D. Cosmelli, and K. J. Friston. Evaluation of different measures of functional connectivity using a neural mass model. *NeuroImage*, 21(2):659–673, 2004.

[103] A. De Simoni, C. B. Griesinger, and F. A. Edwards. Development of rat CA1 neurones in acute versus organotypic slices: role of experience in synaptic morphology and activity. *The Journal of Physiology*, 550(1):135–147, 2003.

[104] F. de Vos, M. Koini, T. M. Schouten, S. Seiler, J. van der Grond, A. Lechner, R. Schmidt, M. de Rooij, and S. A. Rombouts. A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer's disease. *NeuroImage*, 167:62–72, 2018.

[105] W. Denk, J. H. Strickler, and W. W. Webb. Two-photon laser scanning fluorescence microscopy. *Science*, 248(4951):73–76, 1990.

[106] R. Devooght, A. Mantrach, I. Kivimäki, H. Bersini, A. Jaimes, and M. Saerens. Random walks based modularity: application to semi-supervised learning. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 213–224. ACM, 2014.

[107] G. Didier, C. Brun, and A. Baudot. Identifying communities from multiplex biological networks. *PeerJ*, 3:e1525, 2015.

[108] C. Doerr and N. Blenn. Metric convergence in social network sampling. In *Proceedings of the 5th ACM Workshop on HotPlanet*, pages 45–50, 2013.

[109] L. Donetti and M. A. Munoz. Detecting network communities: a new systematic and efficient algorithm. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(10):P10012, 2004.

[110] P. Doukhan, G. Oppenheim, and M. Taqqu. *Theory and applications of long-range dependence*. Springer Science & Business Media, 2002.

[111] M. Dueck, F. Petzke, H. Gerbershagen, M. Paul, V. Hesselmann, R. Girnus, B. Krug, B. Sorger, R. Goebel, R. Lehrke, et al. Propofol attenuates responses of the auditory cortex to acoustic stimulation in a dose-dependent manner: A fMRI study. *Acta Anaesthesiologica Scandinavica*, 49(6):784–791, 2005.

[112] R. I. Dunbar. Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, 16(4):681–694, 1993.

[113] R. Dunn, F. Dudbridge, and C. M. Sanderson. The use of edge-betweenness clustering to investigate biological function in protein interaction networks. *BMC Bioinformatics*, 6(1):39, 2005.

[114] T. Dunwiddie and G. Lynch. Long-term potentiation and depression of synaptic responses in the rat hippocampus: localization and frequency dependency. *Journal of Physiology*, 276(1):353–367, 1978.

[115] P. Dupont, J. Callut, G. Dooms, J.-N. Monette, Y. Deville, and B. Sainte. Relevant subgraph extraction from random walks in a graph. *Universite Catholique de Louvain, UCL/INGI, Number RR*, 7, 2006.

[116] P. Ebbes, Z. Huang, and A. Rangaswamy. Sampling designs for recovering local and global characteristics of social networks. *International Journal of Research in Marketing*, 33(3):578–599, 2016.

[117] G. C. Ellis-Davies. Two-photon uncaging of glutamate. *Frontiers in Synaptic Neuroscience*, 10:48, 2019.

[118] V. Embar, A. Handen, and M. K. Ganapathiraju. Is the average shortest path length of gene set a reflection of their biological relatedness? *Journal of bioinformatics and computational biology*, 14(06):1660002, 2016.

[119] N. Emptage, T. V. Bliss, and A. Fine. Single synaptic events evoke NMDA receptor–mediated release of calcium from internal stores in hippocampal dendritic spines. *Neuron*, 22(1):115–124, 1999.

[120] F. Engert and T. Bonhoeffer. Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature*, 399(6731):66, 1999.

[121] Y.-H. Eom and S. Fortunato. Characterizing and modeling citation dynamics. *PloS One*, 6(9):e24926, 2011.

[122] C. Errando, J. Sigl, M. Robles, E. Calabuig, J. Garcia, F. Arocas, R. Higueras, E. Del Rosario, D. López, C. Peiró, et al. Awareness with recall during general anaesthesia:

a prospective observational evaluation of 4001 patients. *British Journal of Anaesthesia*, 101(2):178–185, 2008.

[123] E. Estrada. When local and global clustering of networks diverge. *Linear Algebra and its Applications*, 488:249–263, 2016.

[124] T. Euler, P. B. Detwiler, and W. Denk. Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature*, 418(6900):845, 2002.

[125] S.-Z. Fan, J.-R. Yeh, B.-C. Chen, J.-S. Shieh, et al. Comparison of EEG approximate entropy and complexity measures of depth of anaesthesia during inhalational general anaesthesia. *Journal of Medical and Biological Engineering*, 31(5):359–366, 2011.

[126] S. J. Farlow. *Partial differential equations for scientists and engineers*. Courier Corporation, 1993.

[127] M. Fellenberg, K. Albermann, A. Zollner, H.-W. Mewes, and J. Hani. Integrative analysis of protein interaction data. In *ISMB*, volume 8, pages 152–161, 2000.

[128] L. K. Fellows, A. S. Heberlein, D. A. Morales, G. Shivde, S. Waller, and D. H. Wu. Method matters: an empirical study of impact in cognitive neuroscience. *Journal of Cognitive Neuroscience*, 17(6):850–858, 2005.

[129] D. Fernández-Espejo, A. Soddu, D. Cruse, E. M. Palacios, C. Junque, A. Vanhaudenhuyse, E. Rivas, V. Newcombe, D. K. Menon, J. D. Pickard, et al. A role for the default mode network in the bases of disorders of consciousness. *Annals of Neurology*, 72(3):335–343, 2012.

[130] R. P. Feynman. Statistical mechanics: A set of lectures. 2018.

[131] J. C. Fiala and K. M. Harris. Dendrite structure. *Dendrites*, 2:1–11, 1999.

[132] V. Fonov, A. C. Evans, K. Botteron, C. R. Almli, R. C. McKinstry, D. L. Collins, B. D. C. Group, et al. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54(1):313–327, 2011.

[133] F. Forrest, M. Tooley, P. Saunders, and C. Prys-Roberts. Propofol infusion and the suppression of consciousness: the EEG and dose requirements. *BJA: British Journal of Anaesthesia*, 72(1):35–41, 1994.

[134] E. Forsythe and P. L. Beales. Bardet–Biedl syndrome. *European Journal of Human Genetics*, 21(1):8, 2013.

[135] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5):75–174, 2010.

[136] S. Fortunato and M. Barthelemy. Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007.

[137] P. T. Fox and J. L. Lancaster. Mapping context and content: the brainmap model. *Nature Reviews Neuroscience*, 3(4):319–321, 2002.

[138] D. Fricker, J. A. Verheugen, and R. Miles. Cell-attached measurements of the firing threshold of rat hippocampal neurones. *The Journal of physiology*, 517(3):791–804, 1999.

[139] J. Friedrich and L. Paninski. Fast active set methods for online spike inference from calcium imaging. In *Advances In Neural Information Processing Systems*, pages 1984–1992, 2016.

[140] K. Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

[141] K. Friston, J. Kilner, and L. Harrison. A free energy principle for the brain. *Journal of Physiology - Paris*, 100(1-3):70–87, 2006.

[142] B. A. Fritz, P. L. Kalarickal, H. R. Maybrier, M. R. Muench, D. Dearth, Y. Chen, K. E. Escallier, A. B. Abdallah, N. Lin, and M. S. Avidan. Intraoperative electroencephalogram suppression predicts postoperative delirium. *Anesthesia and Analgesia*, 122(1):234, 2016.

[143] F.-F. Fu and H.-W. Xue. Co-expression analysis identifies rice starch regulator1 (rsr1), a rice ap2/erebp family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiology*, pages pp–110, 2010.

[144] S. Galatolo, M. Hoyrup, and C. Rojas. Effective symbolic dynamics, random points, statistical behavior, complexity and entropy. *Information and Computation*, 208(1):23–41, 2010.

[145] M. I. Garcia, J. J. Chen, and D. Boehning. Genetically encoded calcium indicators for studying long-term calcium dynamics during apoptosis. *Cell Calcium*, 61:44–49, 2017.

[146] A. Gaskell and J. Sleigh. The quagmire of postoperative delirium: does dose matter? *British Journal of Anaesthesia*, 127(5):664–666, 2021.

[147] A.-C. Gavin, K. Maeda, and S. Kühner. Recent advances in charting protein–protein interaction: mass spectrometry-based approaches. *Current Opinion in Biotechnology*, 22(1):42–49, 2011.

[148] A. Ghazalpour, S. Doss, B. Zhang, S. Wang, C. Plaisier, R. Castellanos, A. Brozell, E. E. Schadt, T. A. Drake, A. J. Lusis, et al. Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genetics*, 2(8):e130, 2006.

[149] J. Gillis, S. Ballouz, and P. Pavlidis. Bias tradeoffs in the creation and analysis of protein–protein interaction networks. *Journal of Proteomics*, 100:44–54, 2014.

[150] A. Giovannucci, J. Friedrich, P. Gunn, J. Kalfon, B. L. Brown, S. A. Koay, J. Taxidis, F. Najafi, J. L. Gauthier, P. Zhou, B. S. Khakh, D. W. Tank, D. B. Chklovskii, and E. A. Pnevmatikakis. Caiman: An open source tool for scalable calcium imaging data analysis. *eLife*, 8:e38173, 2019.

[151] M. Girvan and M. E. Newman. Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12):7821–7826, 2002.

[152] K.-I. Goh, E. Oh, H. Jeong, B. Kahng, and D. Kim. Classification of scale-free networks. *Proceedings of the National Academy of Sciences*, 99(20):12583–12588, 2002.

[153] J. Gómez-Gardenes and V. Latora. Entropy rate of diffusion processes on complex networks. *Physical Review E*, 78(6):065102, 2008.

[154] L. A. Goodman. Snowball sampling. *The annals of mathematical statistics*, pages 148–170, 1961.

[155] B. B. Gore, M. E. Soden, and L. S. Zweifel. Visualization of plasticity in fear-evoked calcium signals in midbrain dopamine neurons. *Learning & Memory*, 21(11):575–579, 2014.

[156] S. Gottwald and D. A. Braun. The two kinds of free energy and the bayesian revolution. *PLoS Computational Biology*, 16(12):e1008420, 2020.

[157] D. Graham, Y. Hao, et al. A selective diffusion model of brain network activity. In *Proceedings of the Conference on Cognitive Computational Neuroscience*, volume 1195, pages 2018–1195, 2018.

[158] A. Grigoriev. A relationship between gene expression and protein interactions on the proteome scale: analysis of the bacteriophage t7 and the yeast saccharomyces cerevisiae. *Nucleic Acids Research*, 29(17):3513–3519, 2001.

[159] A. Grover and J. Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 855–864, 2016.

[160] R. Guimera and L. A. N. Amaral. Functional cartography of complex metabolic networks. *Nature*, 433(7028):895–900, 2005.

[161] R. Haberman. *Elementary applied partial differential equations*, volume 987. Prentice Hall Englewood Cliffs, NJ, 1983.

[162] Z. Hajat, N. Ahmad, and J. Andrzejowski. The role and limitations of EEG-based depth of anaesthesia monitoring in theatres and intensive care. *Anaesthesia*, 72:38–47, 2017.

[163] L. Hakes, D. L. Robertson, S. G. Oliver, and S. C. Lovell. Protein interactions from complexes: a structural perspective. *Comparative and Functional Genomics*, 2007, 2006.

[164] Z. Hammoud and F. Kramer. Multilayer networks: aspects, implementations, and application in biomedicine. *Big Data Analytics*, 5(1):1–18, 2020.

[165] O. Hanteer and M. Magnani. Unspoken assumptions in multi-layer modularity maximization. *Scientific Reports*, 10(1):1–15, 2020.

[166] G. T. Hart, A. K. Ramani, and E. M. Marcotte. How complete are current yeast and human protein-interaction networks? *Genome Biology*, 7(11):120, 2006.

[167] P. Haslehurst, Z. Yang, K. Dholakia, and N. Emptage. Fast volume-scanning light sheet microscopy reveals transient neuronal events. *Biomedical Optics Express*, 9(5):2154–2167, 2018.

[168] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. 1970.

[169] W. A. Haynes, A. Tomczak, and P. Khatri. Gene annotation bias impedes biomedical research. *Scientific Reports*, 8(1):1362, 2018.

[170] L. Heine, A. Soddu, F. Gómez, A. Vanhaudenhuyse, L. Tshibanda, M. Thonnard, V. Charland-Verville, M. Kirsch, S. Laureys, and A. Demertzi. Resting state networks and consciousness. *Frontiers in Psychology*, 3:295, 2012.

[171] R. Hindriks, M. H. Adhikari, Y. Murayama, M. Ganzetti, D. Mantini, N. K. Logothetis, and G. Deco. Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *NeuroImage*, 127:242–256, 2016.

[172] R. Hoebe, C. Van Oven, T. Gadella Jr, P. Dhonukshe, C. Van Noorden, and E. Manders. Controlled light-exposure microscopy reduces photobleaching and phototoxicity in fluorescence live-cell imaging. *Nature Biotechnology*, 25(2):249, 2007.

[173] M. D. Hoffman, D. M. Blei, C. Wang, and J. Paisley. Stochastic variational inference. *Journal of Machine Learning Research*, 14(1):1303–1347, 2013.

[174] A. L. Hopkins. Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology*, 4(11):682–690, 2008.

[175] J. L. Horn. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185, 1965.

[176] A. C. Horton and M. D. Ehlers. Dual modes of endoplasmic reticulum-to-golgi transport in dendrites revealed by live-cell imaging. *Journal of Neuroscience*, 23(15):6188–6199, 2003.

[177] B. Horwitz and D. Poeppel. How can EEG/MEG and fMRI/PET data be combined? *Human Brain Mapping*, 17(1):1, 2002.

[178] J. Y. Hua and S. J. Smith. Neural activity and the dynamics of central nervous system development. *Nature Neuroscience*, 7(4):327, 2004.

[179] X. Huang, D. Chen, T. Ren, and D. Wang. A survey of community detection methods in multilayer networks. *Data Mining and Knowledge Discovery*, 35(1):1–45, 2021.

[180] L. Hubert and P. Arabie. Comparing partitions. *Journal of Classification*, 2(1):193–218, 1985.

[181] J. R. Hughes. Post-tetanic potentiation. *Physiological Reviews*, 38(1):91–113, 1958.

[182] J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt, and E. H. Stelzer. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science*, 305(5686):1007–1009, 2004.

[183] L. G. Humphreys and R. G. Montanelli Jr. An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research*, 10(2):193–205, 1975.

[184] H. E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil engineers*, 116(1):770–799, 1951.

[185] R. J. Huster, S. Debener, T. Eichele, and C. S. Herrmann. Methods for simultaneous EEG-fMRI: an introductory review. *Journal of Neuroscience*, 32(18):6053–6060, 2012.

[186] F. Hutmacher. Why is there so much more research on vision than on any other sensory modality? *Frontiers in Psychology*, 10:2246, 2019.

[187] D. M. Iascone, Y. Li, U. Sümbül, M. Doron, H. Chen, V. Andreu, F. Goudy, I. Segev, H. Peng, and F. Polleux. Whole-neuron synaptic mapping reveals local balance between excitatory and inhibitory synapse organization. *bioRxiv*, page 395384, 2018.

[188] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences*, 98(8):4569–4574, 2001.

[189] C. E. Jahr and C. F. Stevens. Calcium permeability of the N-methyl-D-aspartate receptor channel in hippocampal neurons in culture. *Proceedings of the National Academy of Sciences*, 90(24):11573–11577, 1993.

[190] M. Jalili. Functional brain networks: does the choice of dependency estimator and binarization method matter? *Scientific Reports*, 6(1):1–12, 2016.

[191] R. Jansen, D. Greenbaum, and M. Gerstein. Relating whole-genome expression data with protein-protein interactions. *Genome Research*, 12(1):37–46, 2002.

[192] E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620, 1957.

[193] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith. Fmrib software library (FSL). *NeuroImage*, 62(2):782–790, 2012.

[194] G. F. Jenks. The data model concept in statistical mapping. *International Yearbook of Cartography*, 7:186–190, 1967.

[195] R. G. Jennings and J. D. Van Horn. Publication bias in neuroimaging research: implications for meta-analyses. *Neuroinformatics*, 10(1):67–80, 2012.

[196] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41–42, 2001.

[197] P. Jercog, T. Rogerson, and M. J. Schnitzer. Large-scale fluorescence calcium-imaging methods for studies of long-term memory in behaving mammals. *Cold Spring Harbor Perspectives in Biology*, 8(5):a021824, 2016.

[198] J. Z. Ji, L. Jiao, C. C. Yang, J. W. Lv, and A. D. Zhang. MAE-FMD: Multi-agent evolutionary method for functional module detection in protein-protein interaction networks. *BMC Bioinformatics*, 15(1):325, 2014.

[199] A. M. Jurgens and J. P. Crutchfield. Shannon entropy rate of hidden markov processes. *Journal of Statistical Physics*, 183(2):1–18, 2021.

[200] R. Kanai, Y. Komura, S. Shipp, and K. Friston. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1668):20140169, 2015.

[201] F. I. Karahanoğlu and D. Van De Ville. Dynamics of large-scale fMRI networks: Deconstruct brain activity to build better models of brain function. *Current Opinion in Biomedical Engineering*, 3:28–36, 2017.

[202] Y. Katz, V. Menon, D. A. Nicholson, Y. Geinisman, W. L. Kath, and N. Spruston. Synapse distribution suggests a two-stage model of dendritic integration in CA1 pyramidal neurons. *Neuron*, 63(2):171–177, 2009.

[203] M. B. Kennedy. Signal-processing machines at the postsynaptic density. *Science*, 290(5492):750–754, 2000.

[204] J. M. Kilner, S. J. Kiebel, and K. J. Friston. Applications of random field theory to electrophysiology. *Neuroscience Letters*, 374(3):174–178, 2005.

[205] D. S. Kim, V. Jayaraman, L. L. Looger, and K. Svoboda. Engineering fluorescent calcium sensor proteins for imaging neural activity. *Society for Neuroscience*, pages 11–19, 2014.

[206] G. Kirchhoff. Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird. *Annalen der Physik*, 148(12):497–508, 1847.

[207] C. Koch, M. Massimini, M. Boly, and G. Tononi. Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, 17(5):307–321, 2016.

[208] S.-M. Kokkonen, J. Nikkinen, J. Remes, J. Kantola, T. Starck, M. Haapea, J. Tuominen, O. Tervonen, and V. Kiviniemi. Preoperative localization of the sensorimotor area using independent component analysis of resting-state fMRI. *Magnetic Resonance Imaging*, 27(6):733–740, 2009.

[209] B. Kolb, J. Cioe, and W. Comeau. Contrasting effects of motor and visual spatial learning tasks on dendritic arborization and spine density in rats. *Neurobiology of Learning and Memory*, 90(2):295–300, 2008.

[210] A. N. Kolmogorov. On tables of random numbers. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 369–376, 1963.

[211] R. I. Kondor and J. Lafferty. Diffusion kernels on graphs and other discrete structures. In *Proceedings of the 19th International Conference on Machine Learning*, volume 2002, pages 315–322, 2002.

[212] F. M. Kuenzi, S. M. Fitzjohn, R. A. Morton, G. L. Collingridge, and G. R. Seabrook. Reduced long-term potentiation in hippocampal slices prepared using sucrose-based artificial cerebrospinal fluid. *Journal of Neuroscience Methods*, 100(1-2):117–122, 2000.

[213] F. Kuhn and R. Oshman. Dynamic networks: models and algorithms. *ACM SIGACT News*, 42(1):82–96, 2011.

[214] A. R. Laird, S. B. Eickhoff, K. Li, D. A. Robin, D. C. Glahn, and P. T. Fox. Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. *Journal of Neuroscience*, 29(46):14496–14505, 2009.

[215] R. Lambiotte, J.-C. Delvenne, and M. Barahona. Laplacian dynamics and multiscale modular structure in networks. *arXiv*, 2008.

[216] A. Lancichinetti and S. Fortunato. Consensus clustering in complex networks. *Scientific Reports*, 2(1):1–7, 2012.

[217] P. Langfelder and S. Horvath. Wgcna: an r package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1):559, 2008.

[218] C.-H. Lee, X. Xu, and D. Y. Eun. Beyond random walk and metropolis-hastings samplers: why you should not backtrack for unbiased graph sampling. *ACM SIGMETRICS Performance evaluation review*, 40(1):319–330, 2012.

[219] J.-M. Lee, J. Hu, J. Gao, K. D. White, B. Crosson, C. E. Wierenga, K. McGregor, and K. K. Peck. Identification of brain activity by fractal scaling analysis of functional mri data. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 2, pages ii–137. IEEE, 2005.

[220] U. Lee, G. Oh, S. Kim, G. Noh, B. Choi, and G. A. Mashour. Brain networks maintain a scale-free organization across consciousness, anesthesia, and recovery: evidence for adaptive reconfiguration. *The Journal of the American Society of Anesthesiologists*, 113(5):1081–1091, 2010.

[221] X. Lei, Y. Wang, H. Yuan, and A. Chen. Brain scale-free properties in awake rest and nrem sleep: a simultaneous EEG/fMRI study. *Brain Topography*, 28(2):292–304, 2015.

[222] E. A. Leicht and M. E. Newman. Community structure in directed networks. *Physical Review Letters*, 100(11):118703, 2008.

[223] J. Leskovec and C. Faloutsos. Sampling from large graphs. In *Proceedings of the ACM SIGKDD*, volume 12, pages 631–636, 2006.

[224] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, pages 177–187, 2005.

[225] D. A. Levin and Y. Peres. *Markov chains and mixing times*, volume 107. American Mathematical Society, 2017.

[226] A. C. Lewis, N. S. Jones, M. A. Porter, and C. M. Deane. The function of communities in protein interaction networks at multiple scales. *BMC Systems Biology*, 4(1):100, 2010.

[227] L. D. Lewis, V. S. Weiner, E. A. Mukamel, J. A. Donoghue, E. N. Eskandar, J. R. Madsen, W. S. Anderson, L. R. Hochberg, S. S. Cash, Brown, and E. N. Rapid fragmentation of neuronal networks at the onset of propofol-induced unconsciousness. *Proceedings of the National Academy of Sciences*, 109(49):E3377–E3386, 2012.

[228] R.-H. Li, J. X. Yu, L. Qin, R. Mao, and T. Jin. On random walk based graph sampling. In *2015 IEEE 31st International Conference on Data Engineering*, pages 927–938. IEEE, 2015.

[229] W. Li, J. Xu, J. Huo, L. Wang, Y. Gao, and J. Luo. Distribution consistency based covariance metric networks for few-shot learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 8642–8649, 2019.

[230] L. Liang, V. Chen, K. Zhu, X. Fan, X. Lu, and S. Lu. Integrating data and knowledge to identify functional modules of genes: a multilayer approach. *BMC Bioinformatics*, 20(1):1–15, 2019.

[231] M. Liljeström, A. Hulten, L. Parkkonen, and R. Salmelin. Comparing meg and fmri views to naming actions and objects. *Human Brain Mapping*, 30(6):1845–1856, 2009.

[232] G. Liu, H. Wang, H. Chu, J. Yu, and X. Zhou. Functional diversity of topological modules in human protein-protein interaction networks. *Scientific Reports*, 7(1):1–13, 2017.

[233] Q.-H. Liu, X. Xiong, Q. Zhang, and N. Perra. Epidemic spreading on time-varying multiplex networks. *Physical Review E*, 98(6):062303, 2018.

[234] J. Lonsdale, J. Thomas, M. Salvatore, R. Phillips, E. Lo, S. Shad, R. Hasz, G. Walters, F. Garcia, N. Young, et al. The genotype-tissue expression (gtex) project. *Nature Genetics*, 45(6):580, 2013.

[235] H. Lu, B. Shi, G. Wu, Y. Zhang, X. Zhu, Z. Zhang, C. Liu, Y. Zhao, T. Wu, J. Wang, et al. Integrated analysis of multiple data sources reveals modular structure of biological networks. *Biochemical and Biophysical Research Communications*, 345(1):302–309, 2006.

[236] M. Luecken, M. Page, A. Crosby, S. Mason, G. Reinert, and C. Deane. Commwalker: correctly evaluating modules in molecular networks in light of annotation bias. *Bioinformatics*, 1:7, 2017.

[237] F. Luo, Y. Yang, C.-F. Chen, R. Chang, J. Zhou, and R. H. Scheuermann. Modular organization of protein interaction networks. *Bioinformatics*, 23(2):207–214, 2006.

[238] A. I. Luppi, D. Golkowski, A. Ranft, R. Ilg, D. Jordan, D. K. Menon, and E. A. Stamatakis. Brain network integration dynamics are associated with loss and recovery of consciousness induced by sevoflurane. *Human Brain Mapping*, 42(9):2802–2822, 2021.

[239] H. Ma, S. Leng, K. Aihara, W. Lin, and L. Chen. Randomly distributed embedding making short-term high-dimensional data predictable. *Proceedings of the National Academy of Sciences*, 115(43):E9994–E10002, 2018.

[240] A. S. Mahadevan, U. A. Tooley, M. A. Bertolero, A. P. Mackey, and D. S. Bassett. Evaluating the sensitivity of functional connectivity measures to motion artifact in resting-state fMRI data. *NeuroImage*, 241:118408, 2021.

[241] A. S. Maiya and T. Y. Berger-Wolf. Sampling community structure. In *Proceedings of the 19th International Conference on World Wide Web*, pages 701–710, 2010.

[242] A. S. Maiya and T. Y. Berger-Wolf. Benefits of bias: Towards better characterization of network sampling. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 105–113, 2011.

[243] L. E. Mak, L. Minuzzi, G. MacQueen, G. Hall, S. H. Kennedy, and R. Milev. The default mode network in healthy individuals: a systematic review and meta-analysis. *Brain Connectivity*, 7(1):25–33, 2017.

[244] S. Makeig, S. Debener, J. Onton, and A. Delorme. Mining event-related brain dynamics. *Trends in Cognitive Sciences*, 8(5):204–210, 2004.

[245] N. Makris, J. M. Goldstein, D. Kennedy, S. M. Hodge, V. S. Caviness, S. V. Faraone, M. T. Tsuang, and L. J. Seidman. Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophrenia Research*, 83(2-3):155–171, 2006.

[246] R. C. Malenka. Synaptic plasticity in the hippocampus: Ltp and ltd. *Cell*, 78(4):535–538, 1994.

[247] R. C. Malenka, J. A. Kauer, R. S. Zucker, and R. A. Nicoll. Postsynaptic calcium is sufficient for potentiation of hippocampal synaptic transmission. *Science*, 242(4875):81–84, 1988.

[248] M. Maletic-Savatic, R. Malinow, and K. Svoboda. Rapid dendritic morphogenesis in CA1 hippocampal dendrites induced by synaptic activity. *Science*, 283(5409):1923–1927, 1999.

[249] J. Malmivuo and R. Plonsey. Bioelectromagnetism. *Medical and Biological Engineering and Computing*, 34:9–12, 1996.

[250] B. B. Mandelbrot and J. R. Wallis. Noah, joseph, and operational hydrology. *Water Resources Research*, 4(5):909–918, 1968.

[251] B. Marin and P. Gleeson. Dendritic spikes enhance stimulus selectivity in cortical neurons in vivo, smith, smith, branco & häusser, nature 2013. `https://github.com/OpenSourceBrain/SmithEtAl2013-L23DendriticSpikes`, Feb 2021. (Accessed on 11/23/2021).

[252] W. Marshall, J. Gomez-Ramirez, and G. Tononi. Integrated information and state differentiation. *Frontiers in Psychology*, 7:926, 2016.

[253] S. J. Martin, P. D. Grimwood, and R. G. Morris. Synaptic plasticity and memory: an evaluation of the hypothesis. *Annual Review of Neuroscience*, 23(1):649–711, 2000.

[254] G. A. Mashour, A. Shanks, K. K. Tremper, S. Kheterpal, C. R. Turner, S. K. Ramachandran, P. Picton, C. Schueller, M. Morris, J. C. Vandervest, et al. Prevention of intraoperative awareness with explicit recall in an unselected surgical population: a randomized comparative effectiveness trial. *The Journal of the American Society of Anesthesiologists*, 117(4):717–725, 2012.

[255] P. V. Massey and Z. I. Bashir. Long-term depression: multiple forms and implications for brain function. *Trends in Neurosciences*, 30(4):176–184, 2007.

[256] M. Matsuzaki, G. C. Ellis-Davies, T. Nemoto, Y. Miyashita, M. Iino, and H. Kasai. Dendritic spine geometry is critical for AMPA receptor expression in hippocampal CA1 pyramidal neurons. *Nature Neuroscience*, 4(11):1086, 2001.

[257] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, et al. A four-dimensional probabilistic atlas of the human brain. *Journal of the American Medical Informatics Association*, 8(5):401–430, 2001.

[258] M. Meilă. Comparing clusterings—an information based distance. *Journal of Multivariate Analysis*, 98(5):873–895, 2007.

[259] B. W. Mel. Information processing in dendritic trees. *Neural Computation*, 6(6):1031–1085, 1994.

[260] V. Menon. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences*, 15(10):483–506, 2011.

[261] V. Menon. Salience network. In *Brain Mapping: An Encyclopedic Reference*, pages 597–611. Elsevier, 2015.

[262] R. N. Mhuircheartaigh, C. Warnaby, R. Rogers, S. Jbabdi, and I. Tracey. Slow-wave activity saturation and thalamocortical isolation during propofol anesthesia in humans. *Science Translational Medicine*, 5(208):208ra148–208ra148, 2013.

[263] A. M. Michael, M. Anderson, R. L. Miller, T. Adalı, and V. D. Calhoun. Preserving subject variability in group fMRI analysis: performance evaluation of gica vs. iva. *Frontiers in Systems Neuroscience*, 8:106, 2014.

[264] P. Miller and X.-J. Wang. Power-law neuronal fluctuations in a recurrent network model of parametric working memory. *Journal of Neurophysiology*, 95(2):1099–1114, 2006.

[265] A. Mitra, A. Z. Snyder, E. Tagliazucchi, H. Laufs, J. Elison, R. W. Emerson, M. D. Shen, J. J. Wolff, K. N. Botteron, S. Dager, et al. Resting-state fMRI in sleeping infants more closely resembles adult sleep than adult wakefulness. *PLoS One*, 12(11):e0188122, 2017.

[266] B. Mohar. Some applications of laplace eigenvalues of graphs. In *Graph Symmetry*, pages 225–275. Springer, 1997.

[267] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45(1):3–49, 2003.

[268] H. R. Monday, T. J. Younts, and P. E. Castillo. Long-term plasticity of neurotransmitter release: emerging mechanisms and contributions to brain function and disease. *Annual Review of Neuroscience*, 41:299–322, 2018.

[269] A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, and B. Wold. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature Methods*, 5(7):621, 2008.

[270] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela. Community structure in time-dependent, multiscale, and multiplex networks. *Science*, 328(5980):876–878, 2010.

[271] M. Murphy, M.-A. Bruno, B. A. Riedner, P. Boveroux, Q. Noirhomme, E. C. Landsness, J.-F. Brichant, C. Phillips, M. Massimini, S. Laureys, et al. Propofol anesthesia and sleep: a high-density EEG study. *Sleep*, 34(3):283–291, 2011.

[272] M. Newman. *Networks*. Oxford university press, 2018.

[273] M. E. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[274] M. E. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

[275] M. E. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113, 2004.

[276] C. D. Nguyen, K. J. Gardiner, and K. J. Cios. Protein annotation from protein interaction networks and gene ontology. *Journal of Biomedical Informatics*, 44(5):824–829, 2011.

[277] C. Nicolini, C. Bordier, and A. Bifone. Community detection in weighted brain connectivity networks beyond the resolution limit. *NeuroImage*, 146:28–39, 2017.

[278] M. F. Nolan, G. Malleret, J. T. Dudman, D. L. Buhl, B. Santoro, E. Gibbs, S. Vronskaya, G. Buzsáki, S. A. Siegelbaum, E. R. Kandel, et al. A behavioral role for dendritic integration: Hcn1 channels constrain spatial memory and plasticity at inputs to distal dendrites of CA1 pyramidal neurons. *Cell*, 119(5):719–732, 2004.

[279] F. Oeffner, C. Moch, A. Neundorf, J. Hofmann, M. Koch, and K. Grzeschik. Novel interaction partners of Bardet-Biedl syndrome proteins. *Cytoskeleton*, 65(2):143–155, 2008.

[280] W. C. Oh, L. K. Parajuli, and K. Zito. Heterosynaptic structural plasticity on local dendritic segments of hippocampal CA1 neurons. *Cell Reports*, 10(2):162–169, 2015.

[281] T. Opsahl and P. Panzarasa. Clustering in weighted networks. *Social Networks*, 31(2):155–163, 2009.

[282] M. Ouardouz and B. R. Sastry. Mechanisms underlying LTP of inhibitory synaptic transmission in the deep cerebellar nuclei. *Journal of Neurophysiology*, 84(3):1414–1421, 2000.

[283] E. S. Page. Continuous inspection schemes. *Biometrika*, 41(1/2):100–115, 1954.

[284] G. Palla, I. Derényi, I. Farkas, and T. Vicsek. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, 435(7043):814, 2005.

[285] J. M. Palva, A. Zhigalov, J. Hirvonen, O. Korhonen, K. Linkenkaer-Hansen, and S. Palva. Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proceedings of the National Academy of Sciences*, 110(9):3585–3590, 2013.

[286] J. Pandey, M. Koyutürk, S. Subramaniam, and A. Grama. Functional coherence in domain interaction networks. *Bioinformatics*, 24(16):i28–i34, 2008.

[287] D. Papo. Gauging functional brain activity: from distinguishability to accessibility. *Frontiers in Physiology*, 10:509, 2019.

[288] D. Papo, M. Zanin, J. H. Martínez, and J. M. Buldú. Beware of the small-world neuroscientist! *Frontiers in Human Neuroscience*, 10:96, 2016.

[289] J. Parvizi. Corticocentric myopia: old bias in new cognitive sciences. *Trends in Cognitive Sciences*, 13(8):354–359, 2009.

[290] S. Passow, K. Specht, T. C. Adamsen, M. Biermann, N. Brekke, A. R. Craven, L. Ersland, R. Grüner, N. Kleven-Madsen, O.-H. Kvernenes, et al. Default-mode network functional connectivity is closely related to metabolic activity. *Human Brain Mapping*, 36(6):2027–2038, 2015.

[291] K. Pearson. The problem of the random walk. *Nature*, 72(1865):294–294, 1905.

[292] H. Peng, Z. Ruan, F. Long, J. H. Simpson, and E. W. Myers. V3D enables real-time 3d visualization and quantitative analysis of large-scale biological image data sets. *Nature Biotechnology*, 28(4):348–353, 2010.

[293] C. Pesquita, D. Faria, H. Bastos, A. E. Ferreira, A. O. Falcão, and F. M. Couto. Metrics for go based protein semantic similarity: a systematic evaluation. In *BMC Bioinformatics*, volume 9, page S4. BioMed Central, 2008.

[294] K. H. Pettersen, H. Lindén, T. Tetzlaff, and G. T. Einevoll. Power laws from linear neuronal cable theory: power spectral densities of the soma potential, soma membrane current and single-neuron contribution to the EEG. *PLoS Comput Biology*, 10(11):e1003928, 2014.

[295] D. Picchioni, M. L. Pixa, M. Fukunaga, W. S. Carr, S. G. Horovitz, A. R. Braun, and J. H. Duyn. Decreased connectivity between the thalamus and the neocortex during human nonrapid eye movement sleep. *Sleep*, 37(2):387–397, 2014.

[296] R. A. Poldrack, A. Kittur, D. Kalar, E. Miller, C. Seppa, Y. Gil, D. S. Parker, F. W. Sabb, and R. M. Bilder. The cognitive atlas: toward a knowledge foundation for cognitive neuroscience. *Frontiers in Neuroinformatics*, 5:17, 2011.

[297] T. A. Pologruto, R. Yasuda, and K. Svoboda. Monitoring neural activity and [Ca^{2+}] with genetically encoded Ca^{2+} indicators. *Journal of Neuroscience*, 24(43):9572–9579, 2004.

[298] A. Polsky, B. W. Mel, and J. Schiller. Computational subunits in thin dendrites of pyramidal cells. *Nature Neuroscience*, 7(6):621, 2004.

[299] M. A. Porter, J.-P. Onnela, and P. J. Mucha. Communities in networks. *Notices of the AMS*, 56(9):1082–1097, 2009.

[300] J. Prada, M. Sasi, C. Martin, S. Jablonka, T. Dandekar, and R. Blum. An open source tool for automatic spatiotemporal assessment of calcium transients and local 'signal-close-to-noise'activity in calcium imaging data. *PLoS Computational Biology*, 14(3):e1006054, 2018.

[301] O. Puig, F. Caspary, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm, and B. Séraphin. The tandem affinity purification (tap) method: a general procedure of protein complex purification. *Methods*, 24(3):218–229, 2001.

[302] R. M. Pullon, L. Yan, J. W. Sleigh, and C. E. Warnaby. Granger causality of the electroencephalogram reveals abrupt global loss of cortical information flow during propofol-induced loss of responsiveness. *Anesthesiology*, 133(4):774–786, 2020.

[303] Y. Qi, W. Xu, L. Zhu, and Z. Zhang. Real-world networks are not always fast mixing. *The Computer Journal*, 64(2):236–244, 2021.

[304] I. Rajan and H. T. Cline. Glutamate receptor activity is required for normal development of tectal cell dendrites in vivo. *Journal of Neuroscience*, 18(19):7836–7846, 1998.

[305] E. Ramón-Moliner. The morphology of dendrites. *The structure and function of nervous tissue*, 1:205–267, 1968.

[306] Q. Ran, T. Jamoulle, J. Schaeverbeke, K. Meersmans, R. Vandenberghe, and P. Dupont. Reproducibility of graph measures at the subject level using resting-state fMRI. *Brain and Behavior*, 10(8):2336–2351, 2020.

[307] W. M. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971.

[308] J. Reichardt and S. Bornholdt. Statistical mechanics of community detection. *Physical Review E*, 74(1):016110, 2006.

[309] J. Ren, J. Wang, M. Li, and L. Wang. Identifying protein complexes based on density and modularity in protein-protein interaction network. *BMC Systems Biology*, 7(4):1–15, 2013.

[310] P. A. Robinson, C. Rennie, D. L. Rowe, S. O'Connor, and E. Gordon. Multiscale brain modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):1043–1050, 2005.

[311] C. Rosazza and L. Minati. Resting-state brain networks: literature review and clinical applications. *Neurological Sciences*, 32(5):773–785, 2011.

[312] F. Rosenblatt. Principles of neurodynamics. perceptrons and the theory of brain mechanisms. Technical report, Cornell Aeronautical Lab Inc Buffalo NY, 1961.

[313] G. Rossetti and R. Cazabet. Community discovery in dynamic networks: a survey. *ACM Computing Surveys (CSUR)*, 51(2):1–37, 2018.

[314] M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123, 2008.

[315] G.-C. Rota. The number of partitions of a set. *The American Mathematical Monthly*, 71(5):498–504, 1964.

[316] J.-F. Rual, K. Venkatesan, T. Hao, T. Hirozane-Kishikawa, A. Dricot, N. Li, G. F. Berriz, F. D. Gibbons, M. Dreze, N. Ayivi-Guedehoussou, et al. Towards a proteome-scale map of the human protein–protein interaction network. *Nature*, 437(7062):1173–1178, 2005.

[317] M. Rubinov and O. Sporns. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, 52(3):1059–1069, 2010.

[318] G. Ruffini. An algorithmic information theory of consciousness. *Neuroscience of Consciousness*, 3(1), 2017.

[319] H. Ruffner, A. Bauer, and T. Bouwmeester. Human protein–protein interaction networks and the value for drug discovery. *Drug Discovery Today*, 12(17-18):709–716, 2007.

[320] S. Ryali, K. Supekar, T. Chen, J. Kochalka, W. Cai, J. Nicholas, A. Padmanabhan, and V. Menon. Temporal dynamics and developmental maturation of salience, default and central-executive network interactions revealed by variational bayes hidden markov modeling. *PLoS Computational Biology*, 12(12):e1005138, 2016.

[321] T. Rydén et al. EM versus Markov Chain Monte Carlo for estimation of hidden Markov models: A computational perspective. *Bayesian Analysis*, 3(4):659–688, 2008.

[322] B. L. Sabatini, T. G. Oertner, and K. Svoboda. The life cycle of Ca^{2+} ions in dendritic spines. *Neuron*, 33(3):439–452, 2002.

[323] S. Sadaghiani, J.-B. Poline, A. Kleinschmidt, and M. D'Esposito. Ongoing dynamics in large-scale functional connectivity predict perception. *Proceedings of the National Academy of Sciences*, 112(27):8463–8468, 2015.

[324] S. Sahara, Y. Yanagawa, D. D. O'Leary, and C. F. Stevens. The fraction of cortical gabaergic neurons is constant from near the start of cortical neurogenesis to adulthood. *Journal of Neuroscience*, 32(14):4755–4761, 2012.

[325] S. E. Said and D. A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.

[326] K. D. Sakaki, P. Coleman, T. D. Toth, C. Guerrier, and K. Haas. Automating event-detection of brain neuron synaptic activity and action potential firing in vivo using a random-access multiphoton laser scanning microscope for real-time analysis. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1–7. IEEE, 2018.

[327] M. A. S. Salman and D. V. Borkar. Exponential matrix and their properties. *International Journal of Scientific and Innovative Mathematical Research (IJSIMR) Volume*, 4:53–63, 2016.

[328] P. G. Sämann, R. Wehrle, D. Hoehn, V. I. Spoormaker, H. Peters, C. Tully, F. Holsboer, and M. Czisch. Development of the brain's default mode network from wakefulness to slow wave sleep. *Cerebral Cortex*, 21(9):2082–2093, 2011.

[329] G. Samorodnitsky. *Long range dependence*. Now Publishers Inc, 2007.

[330] P. A. Santi. Light sheet fluorescence microscopy: a review. *Journal of Histochemistry & Cytochemistry*, 59(2):129–138, 2011.

[331] M. H. Schaefer, L. Serrano, and M. A. Andrade-Navarro. Correcting for the study bias associated with protein–protein interaction measurements reveals differences between protein degree distributions from different cancer types. *Frontiers in Genetics*, 6:260, 2015.

[332] M. Schartner, A. Seth, Q. Noirhomme, M. Boly, M.-A. Bruno, S. Laureys, and A. Barrett. Complexity of multi-dimensional spontaneous EEG decreases during propofol induced general anaesthesia. *PloS One*, 10(8):e0133532, 2015.

[333] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676, 2012.

[334] R. Schneggenburger, Z. Zhou, A. Konnerth, and E. Neher. Fractional contribution of calcium to the cation current through glutamate receptor channels. *Neuron*, 11(1):133–143, 1993.

[335] A. M. Schnoes, D. C. Ream, A. W. Thorman, P. C. Babbitt, and I. Friedberg. Biases in the experimental annotations of protein function and their effect on our understanding of protein function space. *PLoS Computational Biology*, 9(5):e1003063, 2013.

[336] L. M. Schriml, E. Mitraka, J. Munro, B. Tauber, M. Schor, L. Nickle, V. Felix, L. Jeng, C. Bearer, R. Lichenstein, et al. Human disease ontology 2018 update: classification, content and workflow expansion. *Nucleic Acids Research*, 47(D1):D955–D962, 2019.

[337] E. M. Schuman and D. V. Madison. Locally distributed synaptic potentiation in the hippocampus. *Science*, 263(5146):532–536, 1994.

[338] D. Schwender, H. Kunze-Kronawitter, P. Dietrich, S. Klasing, H. Forst, and C. Madler. Conscious awareness during general anaesthesia: patients' perceptions, emotions, cognition and reactions. *British Journal of Anaesthesia*, 80(2):133–139, 1998.

[339] G. W. Schwert. Tests for unit roots: A monte carlo investigation. *Journal of Business & Economic Statistics*, 20(1):5–17, 2002.

[340] B. Schwikowski, P. Uetz, and S. Fields. A network of protein–protein interactions in yeast. *Nature Biotechnology*, 18(12):1257, 2000.

[341] M. Segal. Dendritic spines and long-term plasticity. *Nature Reviews Neuroscience*, 6(4):277, 2005.

[342] I. Serebriiskii, J. Estojak, M. Berman, and E. A. Golemis. Approaches to detecting false positives in yeast two-hybrid systems. *Biotechniques*, 28(2):328–337, 2000.

[343] I. G. Serebriiskii and E. A. Golemis. Two-hybrid system and false positives. In *Two-Hybrid Systems*, pages 123–134. Springer, 2001.

[344] A. Sevenius Nilsen, B. E. Juel, and W. Marshall. Evaluating approximations and heuristic measures of integrated information. *Entropy*, 21(5):525, 2019.

[345] P. Shakarian, P. Roos, D. Callahan, and C. Kirk. Mining for geographically disperse communities in social networks by leveraging distance modularity. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1402–1409, 2013.

[346] A. Shanker, J. H. Abel, G. Schamberg, and E. N. Brown. Etiology of burst suppression EEG patterns. *Frontiers in Psychology*, page 2207, 2021.

[347] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11):2498–2504, 2003.

[348] X. Shen and F. G. Meyer. Low-dimensional embedding of fMRI datasets. *NeuroImage*, 41(3):886–902, 2008.

[349] Y. Shen, J. Wang, and S. Navlakha. A correspondence between normalization strategies in artificial and biological neural networks. *bioRxiv*, 2020.

[350] C.-W. Shin and S. Kim. Self-organized criticality and scale-free properties in emergent functional neural networks. *Physical Review E*, 74(4):045101, 2006.

[351] J. M. Shine, P. G. Bissett, P. T. Bell, O. Koyejo, J. H. Balsters, K. J. Gorgolewski, C. A. Moodie, and R. A. Poldrack. The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron*, 92(2):544–554, 2016.

[352] M. F. Shlesinger. Fractal time and 1/f noise in complex systems. *Annals of the New York Academy of Sciences*, 504(1):214–228, 1987.

[353] G. L. Shulman, J. A. Fiez, M. Corbetta, R. L. Buckner, F. M. Miezin, and M. E. Raichle. Common blood flow changes across visual tasks: II. decreases in cerebral cortex. *Journal Cognitive Neuroscience*, 9:648–63.

[354] N. Shulman and M. Feder. The uniform distribution as a universal prior. *IEEE Transactions on Information Theory*, 50(6):1356–1362, 2004.

[355] R. H. Shumway and D. S. Stoffer. Time series analysis and its applications. *Studies In Informatics And Control*, 9(4):375–376, 2000.

[356] J.-B. Sibarita. Deconvolution microscopy. In *Microscopy Techniques*, pages 201–243. Springer, 2005.

[357] K. Sigmund, C. Hauert, and M. A. Nowak. Reward and punishment. *Proceedings of the National Academy of Sciences*, 98(19):10757–10762, 2001.

[358] G. J. Simpson. *Nonlinear optical polarization analysis in chemistry and biology.* Cambridge University Press, 2017.

[359] J. Smallwood, K. Brown, B. Baird, and J. W. Schooler. Cooperation between the default mode network and the frontal–parietal network in the production of an internal train of thought. *Brain Research*, 1428:60–70, 2012.

[360] K. Smith, D. Abásolo, and J. Escudero. Accounting for the complex hierarchical topology of EEG phase-based functional connectivity in network binarisation. *PloS One*, 12(10):e0186164, 2017.

[361] S. L. Smith, I. T. Smith, T. Branco, and M. Häusser. Dendritic spikes enhance stimulus selectivity in cortical neurons in vivo. *Nature*, 503(7474):115–120, 2013.

[362] S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird, et al. Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*, 106(31):13040–13045, 2009.

[363] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, et al. Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23:S208–S219, 2004.

[364] R. D. Smrt and X. Zhao. Epigenetic regulation of neuronal dendrite and dendritic spine development. *Frontiers in Biology*, 5(4):304–323, 2010.

[365] R. V. Solé and S. Valverde. Information theory of complex networks: on evolution and architectural constraints. In *Complex Networks*, pages 189–207. Springer, 2004.

[366] M. Spitzer. *Exploratory analysis and data modeling in functional neuroimaging.* MIT Press, 2003.

[367] D. Sridharan, D. J. Levitin, and V. Menon. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences*, 105(34):12569–12574, 2008.

[368] C. J. Stam. Functional connectivity patterns of human magnetoencephalographic recordings: a 'small-world'network? *Neuroscience Letters*, 355(1-2):25–28, 2004.

[369] A. Stevner, D. Vidaurre, J. Cabral, K. Rapuano, S. F. V. Nielsen, E. Tagliazucchi, H. Laufs, P. Vuust, G. Deco, M. Woolrich, et al. Discovery of key whole-brain transitions and dynamics during human wakefulness and non-REM sleep. *Nature Communications*, 10(1):1035, 2019.

[370] J. T. Stock, M. V. Orna, et al. *Electrochemistry, past and present*, volume 390. ACS Publications, 1989.

[371] S. H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001.

[372] G. Stuart, N. Spruston, and M. Häusser. *Dendrites*. Oxford University Press, 2016.

[373] M. P. Stumpf and M. A. Porter. Critical truths about power laws. *Science*, 335(6069):665–666, 2012.

[374] D. Stutzbach, R. Rejaie, N. Duffield, S. Sen, and W. Willinger. On unbiased sampling for unstructured peer-to-peer networks. *IEEE/ACM Transactions on Networking*, 17(2):377–390, 2008.

[375] H.-I. Suk, C.-Y. Wee, S.-W. Lee, and D. Shen. State-space model with deep learning for functional dynamics estimation in resting-state fMRI. *NeuroImage*, 129:292–307, 2016.

[376] M. R. Sullivan, A. Nimmerjahn, D. V. Sarkisov, F. Helmchen, and S. S.-H. Wang. In vivo calcium imaging of circuit activity in cerebellar cortex. *Journal of Neurophysiology*, 94(2):1636–1644, 2005.

[377] S. Suthram, J. T. Dudley, A. P. Chiang, R. Chen, T. J. Hastie, and A. J. Butte. Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. *PLoS Computational Biology*, 6(2):e1000662, 2010.

[378] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, et al. String v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, 43(D1):D447–D452, 2014.

[379] T. Takeuchi, A. J. Duszkiewicz, and R. G. Morris. The synaptic plasticity and memory hypothesis: encoding, storage and persistence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1633):20130288, 2014.

[380] J. Tee and D. P. Taylor. Is information in the brain represented in continuous or discrete form? *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 6(3):199–209, 2020.

[381] M. S. Thanawala and W. G. Regehr. Presynaptic calcium influx controls neurotransmitter release in part by regulating the effective size of the readily releasable pool. *Journal of Neuroscience*, 33(11):4625–4633, 2013.

[382] G. J. Thompson. Neural and metabolic basis of dynamic resting state fMRI. *NeuroImage*, 180:448–462, 2018.

[383] Z. Tian, M. Guo, C. Wang, X. Liu, and S. Wang. Refine gene functional similarity network based on interaction networks. *BMC bioinformatics*, 18(16):183–193, 2017.

[384] D. G. Tomasi, E. Shokri-Kojori, C. E. Wiers, S. W. Kim, Ş. B. Demiral, E. A. Cabrera, E. Lindgren, G. Miller, G.-J. Wang, and N. D. Volkow. Dynamic brain glucose metabolism identifies anti-correlated cortical-cerebellar networks at rest. *Journal of Cerebral Blood Flow & Metabolism*, 37(12):3659–3670, 2017.

[385] C. Tong, Y. Lian, J. Niu, Z. Xie, and Y. Zhang. A novel green algorithm for sampling complex networks. *Journal of Network and Computer Applications*, 59:55–62, 2016.

[386] H. Tong, C. Faloutsos, and J.-Y. Pan. Fast random walk with restart and its applications. *Proceedings of the 6th International Conference on Data Mining*, pages 613–622, 2006.

[387] G. Tononi. An information integration theory of consciousness. *BMC Neuroscience*, 5(1):1–22, 2004.

[388] V. A. Traag, L. Waltman, and N. J. Van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports*, 9(1):1–12, 2019.

[389] B. Trabold and T. Metterlein. Postoperative delirium: risk factors, prevention, and treatment. *Journal of Cardiothoracic and Vascular Anesthesia*, 28(5):1352–1360, 2014.

[390] A. Tran-Van-Minh, R. D. Cazé, T. Abrahamsson, L. Cathala, B. S. Gutkin, and D. A. Digregorio. Contribution of sublinear and supralinear dendritic integration to neuronal computations. *Frontiers in Cellular Neuroscience*, 9:67, 2015.

[391] S. Tripathi, S. Moutari, M. Dehmer, and F. Emmert-Streib. Comparison of module detection algorithms in protein networks and investigation of the biological meaning of predicted modules. *BMC Bioinformatics*, 17(1):129, 2016.

[392] A. Tung and W. B. Mendelson. Anesthesia and sleep. *Sleep Medicine Reviews*, 8(3):213–225, 2004.

[393] L. Q. Uddin, B. T. Yeo, and R. N. Spreng. Towards a universal taxonomy of macro-scale functional human brain networks. *Brain Topography*, pages 1–17.

[394] B. Urban and M. Bleckwenn. Concepts and correlations relevant to general anaesthesia. *British Journal of Anaesthesia*, 89(1):3–16, 2002.

[395] M. Vaiana, E. M. Goldberg, and S. F. Muldoon. Optimizing state change detection in functional temporal networks through dynamic community detection. *Journal of Complex Networks*, 7(4):529–553, 2019.

[396] A. Valdeolivas, L. Tichit, C. Navarro, S. Perrin, G. Odelin, N. Levy, P. Cau, E. Remy, and A. Baudot. Random walk with restart on multiplex and heterogeneous biological networks. *Bioinformatics*, 35(3):497–505, 2019.

[397] B. van Bommel and M. Mikhaylova. Talking to the neighbours: the molecular and physiological mechanisms of clustered synaptic plasticity. *Neuroscience & Biobehavioral Reviews*, 71:352–361, 2016.

[398] M. P. van den Heuvel, C. J. Stam, M. Boersma, and H. H. Pol. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *NeuroImage*, 43(3):528–539, 2008.

[399] V. van Noort, B. Snel, and M. A. Huynen. Predicting gene function by conserved co-expression. *Trends in Genetics*, 19(5):238–242, 2003.

[400] G. Varvatsoulias. The physiological processes underpinning pet and fMRI techniques with an emphasis on the temporal and spatial resolution of these methods. 2013.

[401] F. Vecchio, F. Miraglia, E. Judica, M. Cotelli, F. Alù, and P. M. Rossini. Human brain networks: A graph theoretical analysis of cortical connectivity normative database from EEG data in healthy elderly subjects. *GeroScience*, 42(2):575, 2020.

[402] D. Vella, I. Zoppis, G. Mauri, P. Mauri, and D. Di Silvestre. From protein-protein interactions to protein co-expression networks: a new perspective to evaluate large-scale proteomic data. *EURASIP Journal on Bioinformatics and Systems Biology*, 2017(1):1–16, 2017.

[403] K. Venkatesan, J.-F. Rual, A. Vazquez, U. Stelzl, I. Lemmens, T. Hirozane-Kishikawa, T. Hao, M. Zenkner, X. Xin, and K.-I. Goh. An empirical framework for binary interactome mapping. *Nature Methods*, 6(1):83, 2009.

[404] R. Veselis. Complexities of human memory: relevance to anaesthetic practice. *British Journal of Anaesthesia*, 121(1):210–218, 2018.

[405] T. Vettenburg, H. I. Dalgarno, J. Nylk, C. Coll-Lladó, D. E. Ferrier, T. Čižmár, F. J. Gunn-Moore, and K. Dholakia. Light-sheet microscopy using an Airy beam. *Nature Methods*, 11(5):541–544, 2014.

[406] D. Vidaurre, R. Abeysuriya, R. Becker, A. J. Quinn, F. Alfaro-Almagro, S. M. Smith, and M. W. Woolrich. Discovering dynamic brain networks from big data in rest and task. *NeuroImage*, 180:646–656, 2018.

[407] D. Vidaurre, L. T. Hunt, A. J. Quinn, B. A. Hunt, M. J. Brookes, A. C. Nobre, and M. W. Woolrich. Spontaneous cortical activity transiently organises into frequency specific phase-coupling networks. *Nature Communications*, 9(1):2987, 2018.

[408] D. Vidaurre, A. J. Quinn, A. P. Baker, D. Dupret, A. Tejero-Cantero, and M. W. Woolrich. Spectrally resolved fast transient brain states in electrophysiological data. *NeuroImage*, 126:81–95, 2016.

[409] D. Vidaurre, S. M. Smith, and M. W. Woolrich. Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48):12827–12832, 2017.

[410] S. Vijayan, S. Ching, P. L. Purdon, E. N. Brown, and N. J. Kopell. Thalamocortical mechanisms for the anteriorization of alpha rhythms during propofol-induced unconsciousness. *Journal of Neuroscience*, 33(27):11070–11075, 2013.

[411] N. Vijesh, S. K. Chakrabarti, J. Sreekumar, et al. Modeling of gene regulatory networks: A review. *Journal of Biomedical Science and Engineering*, 6(02):223, 2013.

[412] A. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE transactions on Information Theory*, 13(2):260–269, 1967.

[413] I. Voineagu, X. Wang, P. Johnston, J. K. Lowe, Y. Tian, S. Horvath, J. Mill, R. M. Cantor, B. J. Blencowe, and D. H. Geschwind. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*, 474(7351):380, 2011.

[414] C. Von Mering, R. Krause, B. Snel, M. Cornell, S. G. Oliver, S. Fields, and P. Bork. Comparative assessment of large-scale data sets of protein–protein interactions. *Nature*, 417(6887):399, 2002.

[415] A. Wagner and D. A. Fell. The small world inside large metabolic networks. *Proceedings of the Royal Society of London B: Biological Sciences*, 268(1478):1803–1810, 2001.

[416] C. Walz, K. Jüngling, V. Lessmann, and K. Gottmann. Presynaptic plasticity in an immature neocortical network requires NMDA receptor activation and bdnf release. *Journal of Neurophysiology*, 96(6):3512–3516, 2006.

[417] H. Wang, S. Zhao, Q. Dong, Y. Cui, Y. Chen, J. Han, L. Xie, and T. Liu. Recognizing brain states using deep sparse recurrent neural network. *IEEE Transactions on Medical Imaging*, 38(4):1058–1068, 2018.

[418] P. Wang and P. Blunsom. Collapsed variational bayesian inference for hidden markov models. In *Artificial Intelligence and Statistics*, pages 599–607, 2013.

[419] Z. Wang, Y. Chen, and Y. Li. A brief review of computational gene prediction methods. *Genomics, Proteomics & Bioinformatics*, 2(4):216–221, 2004.

[420] Z. Wang, M. Gerstein, and M. Snyder. RNA-seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1):57, 2009.

[421] J. H. Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963.

[422] C. E. Warnaby, J. W. Sleigh, D. Hight, S. Jbabdi, and I. Tracey. Investigation of slow-wave activity saturation during surgical anesthesia reveals a signature of neural inertia in humans. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 127(4):645–657, 2017.

[423] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world'networks. *Nature*, 393(6684):440–442, 1998.

[424] R. Weisskoff, J. Baker, J. Belliveau, T. Davis, K. Kwong, M. Cohen, and B. Rosen. Power spectrum analysis of functionally-weighted mr data: what's in the noise. In *Proc. Soc. Magn. Reson. Med*, volume 1, 1993.

[425] D. B. West et al. *Introduction to graph theory*, volume 2. Prentice Hall Upper Saddle River, 2001.

[426] E. L. Whitlock, A. Vannucci, and M. S. Avidan. Postoperative delirium. *Minerva Anestesiologica*, 77(4):448, 2011.

[427] A. Witt and B. D. Malamud. Quantification of long-range persistence in geophysical time series: Conventional and benchmark-based improvement techniques. *Surveys in Geophysics*, 34(5):541–651, 2013.

[428] M. W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. M. Smith. Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, 45(1):S173–S186, 2009.

[429] J. Wright and P. Bourke. The growth of cognition: Free energy minimization and the embryogenesis of cortical computation. *Physics of Life Reviews*, 36:83–99, 2021.

[430] G. Wu, X. Feng, and L. Stein. A human functional protein interaction network and its application to cancer data analysis. *Genome Biology*, 11(5):1–23, 2010.

[431] F. Xia, J. Liu, H. Nie, Y. Fu, L. Wan, and X. Kong. Random walks: A review of algorithms and applications. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 4(2):95–107, 2019.

[432] H. Xiao and H. Peng. APP2: automatic tracing of 3d neuron morphology based on hierarchical pruning of a gray-weighted image distance-tree. *Bioinformatics*, 29(11):1448–1454, 2013.

[433] X. Xu, S. Xu, L. Jin, and E. Song. Characteristic analysis of otsu threshold and its applications. *Pattern Recognition Letters*, 32(7):956–961, 2011.

[434] R. Xulvi-Brunet and H. Li. Co-expression networks: graph properties and topological comparisons. *Bioinformatics*, 26(2):205–214, 2010.

[435] G. Yang, F. Pan, and W.-B. Gan. Stably maintained dendritic spines are associated with lifelong memories. *Nature*, 462(7275):920, 2009.

[436] T. Yarkoni, R. A. Poldrack, T. E. Nichols, D. C. Van Essen, and T. D. Wager. Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8(8):665–670, 2011.

[437] S.-H. Yook, Z. N. Oltvai, and A.-L. Barabási. Functional and topological characterization of protein interaction networks. *Proteomics*, 4(4):928–942, 2004.

[438] S. Yoon, S. Lee, S.-H. Yook, and Y. Kim. Statistical properties of sampled networks by fks. *Physical Review E*, 75(4):046114, 2007.

[439] W. Yu, X. Lin, and W. Zhang. Towards efficient simrank computation on large networks. In *2013 IEEE 29th International Conference on Data Engineering (ICDE)*, pages 601–612. IEEE, 2013.

[440] S. S. Zakharenko, L. Zablow, and S. A. Siegelbaum. Visualization of changes in presynaptic function during long-term synaptic plasticity. *Nature Neuroscience*, 4(7):711–717, 2001.

[441] N. T. Zewde. Multiscale solutions to quantitative systems biology models. *Frontiers in Molecular Biosciences*, 6:119, 2019.

[442] F. Zhang, L.-P. Wang, M. Brauner, J. F. Liewald, K. Kay, N. Watzke, P. G. Wood, E. Bamberg, G. Nagel, A. Gottschalk, et al. Multimodal fast optical interrogation of neural circuitry. *Nature*, 446(7136):633–639, 2007.

[443] Z. Zhang and J. Zhang. A big world inside small-world networks. *PloS One*, 4(5):e5686, 2009.

[444] K. Zito and V. N. Murthy. Dendritic spines. *Current Biology*, 12(1):R5, 2002.

A | Appendices:

Resting State Networks

This is a brief primer on the resting state networks which are important to the study of network neuroscience in fMRI.

A.1 Regional Activity in Resting State Networks

Resting State Networks (RSNs) are a type of brain network which are mostly used to characterise the brain state of awake subjects using fMRI data. There are three primary RSNs, namely the DMN, Salience and Central Executive Networks (SN and CEN). These networks are commonly active during rest and task while awake (see Appendix Table A.1). The RSNs have a dynamic relationship in wakefulness that is known to be disrupted to varying degrees during periods of altered conscious awareness such as sleep, or anaesthesia [170], as well as pathological conditions such as coma or minimally conscious states caused by brain damage [129].

Resting state networks are characterised by their co-activation and high level of functional connectivity (shared network edges) between the constituent regions. There are three primary RSNs, the Default Mode Network (DMN), Salience Network (SN) and Central Executive Network (CEN). The DMN is an RSN associated with unfocused attention as seen in daydreaming or self-referential thinking. It is anti-correlated with task performance [243]. DMN activity is most associated with co-recruitment of prefrontal and precuneous cortices, as well as the angular and (posterior) cingulate gyri. In contrast to the DMN, the CEN is more closely linked to focused attention when performing a task. CEN activity is associated with co-activation of the middle frontal gyrus and inferior parietal areas (that include the supramagrginal and angular gyri) Though shared regions do exist between these networks, the SN is more closely linked with

the (anterior) cingulate cortex, while the CEN is seen to recruit the cingulate regions, supramarginal gyri (part of the parietal lobe) and more posterior frontal regions than in the DMN. Switching between the DMN and CEN is common and may be moderated by the SN in response to stimuli [261, 393].

The brain regions in Table A.1 are some of those that are typically associated with each of the RSNs, however, since the exact network observed is task and subject-specific, different studies characterise each RSN differently, including different regions and to different spatial extents. Other regions commonly associated with the wakeful brain during fMRI include auditory regions (such as Heschl's Gyrus) in the temporal lobes (located medially, to the left and right), these are most commonly activated by the loud noise of the scanner but may also sometimes by superficially active even in lower levels of consciousness and possibly superficially even in an unconscious state [111].

DMN	SN	CEN
Inferior Frontal Gyrus, pars triangularis	Insular Cortex	Middle Frontal Gyrus
Angular Gyrus	Frontal Medial Cortex	Supramarginal Gyrus, anterior
Cingulate Gyrus, posterior	Cingulate Gyrus, anterior	Supramarginal Gyrus, posterior
Precuneous Cortex	Parietal Operculum Cortex	Superior Parietal Lobule
		Inferior Temporal Gyrus, posterior

Table A.1: Resting state networks classified by several key brain regions as defined by the Harvard-Oxford (HO) brain parcellation. Brain regions were conservatively chosen for the Default Mode Network (DMN), the Salience Network (SN) and the Central Executive Network (CEN) based on a review of regions most commonly implicated in their core functions [21, 261, 393]. Due to the relatively low resolution of the HO parcellation, chosen regions may encompass a larger region than was initially implicated to be involved in the circuit.

B | Appendices: CommFinedWalker

This appendix details important tools and methods for co-expression and PIN analysis as well as supplementary results.

B.1 Snowball Sampling

We give a detailed description of snowball sampling due to its importance to later sections. Snowball sampling is a recursive sampling technique that works by considering the union of repeatedly applying the neighbourhood operator defined in Section 2.1.1. That is, given a node i, we define the n^{th} order snowball sample

$$O_n(i) = \bigcup_{j \in O_{n-1}(i)} N(j),$$

where is the singleton $O_0 = \{i\}$. Put another way, a snowball sample is the union of all nodes that are at most distance n from i. Snowball sampling can also be considered a generalisation of the neighbourhood concept to higher orders and $O_n(i)$ can also be termed the n-neighbourhood of i.

B.2 Hierarchical Gene Set Enrichment Analysis: topGO

The purpose of gene set enrichment analysis is to find terms which are disproportionately annotated to an otherwise interesting subset of genes or in this case the genes' associated proteins $I \subset V(G)$ (e.g. a community). Enrichment is in comparison to the universe of proteins from which the interesting subset was selected. In the context of PINs this universe is formed from

the full set of annotated network proteins $\pi_{\mathscr{A}}[V(G)]$ (recall $\pi_A[B]$ returns the proteins in B associated with annotations in the set A). A common way of testing term enrichment in GO is to use Fisher's exact test to compute the probability of observing a specific pattern of protein annotation for each term $a \in \{a : \pi_{\{a\}}(I) \neq \emptyset\}$. In order to compute each probability the test operates under an assumption of independence between: the probability of the observing the given number of proteins associated with a inside I, $|\pi_{\{a\}}[I]|$ and outside of I, $|\pi_{\{a\}}[V(G) - I]|$. The p-value for the test is an exact probability drawn from the hypergeometric distribution based on the null hypothesis of independence between the in-group and out-group sampling. The probability is given

$$P(\text{observing occurences of } a|\text{independent in/out-groups}) = \frac{\binom{|\pi_{\mathscr{A}}[I]|}{|\pi_{\{a\}}[I]|}\binom{|\pi_{\mathscr{A}}[V(G)-I]|}{|\pi_{\{a\}}[V(G)-I]|}}{\binom{|\pi_{\mathscr{A}}[V(G)]|}{|\pi_{\{a\}}[V(G)]|}}. \quad \text{(B.1)}$$

the null hypothesis can be rejected with p-value equal to the probability in Eq. (B.1) when the probability is sufficiently small (e.g. < 0.05).

This test is repeated for each term and then for all of the terms' ancestors as well. Since, by the nature of specificity, child annotations are related to their parents, significance testing of child terms effectively re-tests parent terms. Therefore when only using standard multiple testing corrections, this exhaustive testing is still biased by the dependence between parent and child annotations [12]. A method that respects the topology of GO (termed topGO) can be useful in mitigating this bias.

Using topGO, each of the three GO DAGs are organised hierarchically into levels, loosely corresponding to the specificity of the terms. The level of a term is defined as the maximum number of repeated parental calls possible from that term for each call to only contain the root term (see Section 3.1.4) [11]. Enrichment is then tested at each of these levels starting from the most specific terms in the highest level of specificity. If a term, a, is found to be significantly over-represented in I under the Fisher test (Eq. (B.1)) then all proteins in the protein universe associated with it are deleted from further tests of a's ancestors. See Figure B.1 for an example based on Figure 3.1. The result is that each term enrichment test is performed on term sets not directly related to previous significant terms, mitigating some of the bias caused by parent-child term correlations.

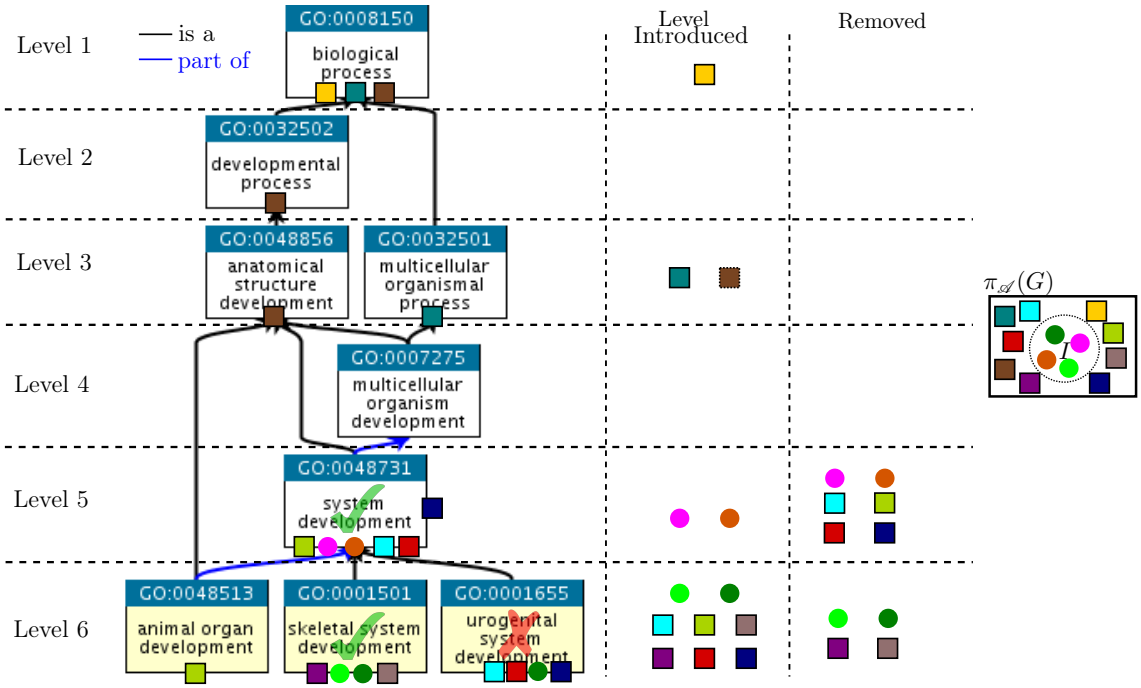
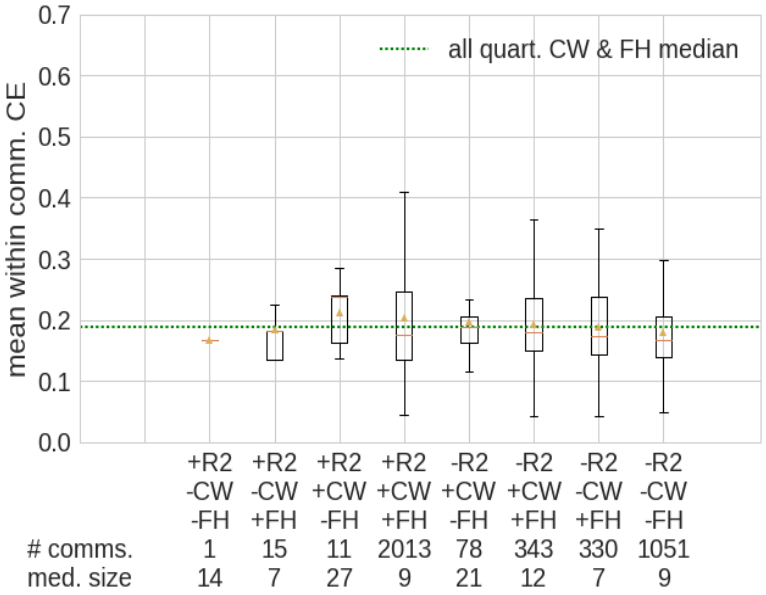
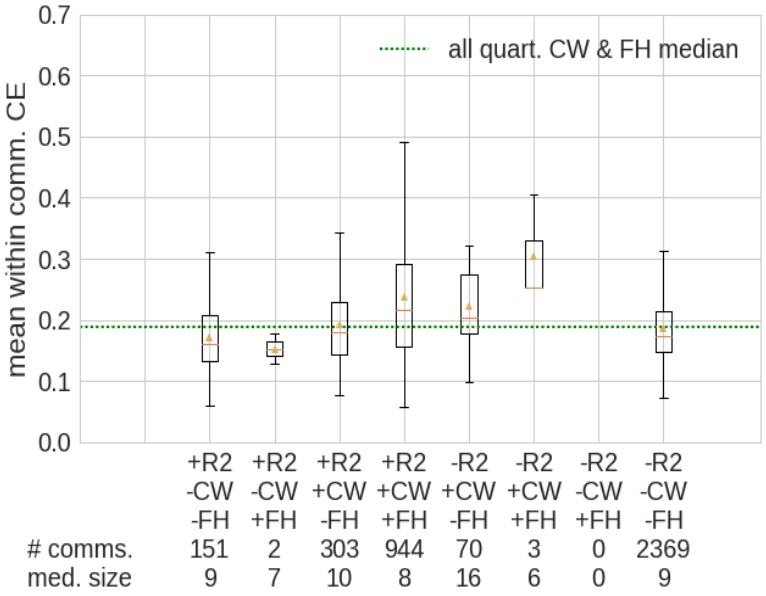


Figure B.1: This Figure shows the specificity levels of the GO hierarchy with topGO testing of terms associated with the proteins in the interest set I (right, dotted circle). This is an example data set with an interest set of $|I| = 4$ and protein universe $|\pi_{\mathscr{A}}(G)| = 13$. Testing only takes place for terms associated with proteins in I and begins with terms that are in the highest level for which annotations are available (Level 6, bottom). Different proteins annotated with each term are sourced from the protein universe $\pi_{\mathscr{A}}(G)$ (solid rectangle on right). They are represented by colour and by shape coresponding to proteins in the set of interest (coloured circles), I, and proteins outside I (bordered squares). Proteins appear in the GO heirarchy (left) in the text boxes of the terms which annotate them. The two columns show the level a protein is first introduced to be tested (i.e. its highest level of annotation) and the level it is subsequently removed from testing of its parents after a significant test. In Level 6, I has been found to be enriched for *skeletal system development* (green tick) while the null hypothesis was not rejected for the other term with I proteins, *urogenital system development* (red cross). Thus all four of the proteins annotated with *skeletal system development* are removed from further tests of ancestor terms, while proteins associated with only *urogenital system development* or *animal organ development* are not removed. At Level 5 testing for *systems development* is significant and so all proteins, including the two just introduced, are eliminated from further parental tests. Terms introduced in lower levels are associated with proteins in $\pi_{\mathscr{A}}(G)$ but not I and so are not tested or removed.

B.3 Co-expression Over Local Homogeneity Quartiles

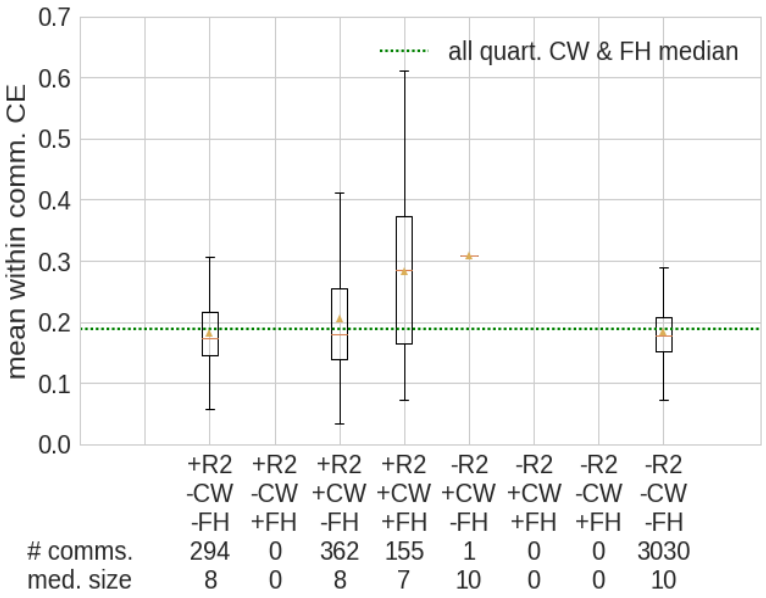


(a) HINT-P: Quantile 1

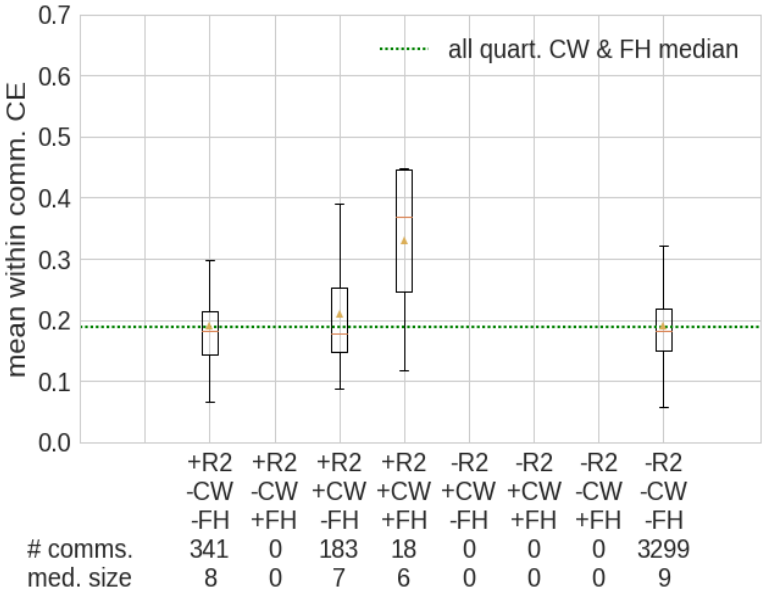


(b) HINT-P: Quantile 2

Figure B.2: Box plots of HINT-P community co-expression statistics for the two highest LHM quartiles showing all combinations of communities selected (+) or not (-) by each combination of methods (means for each combination marked with a triangle). The text below each plot shows the number of communities above the median size of communities in that class. Again, CW&FH median co-expression is used as benchmark value (dotted line). Some of the classes with only a few communities may have unreliable co-expression distributions, in quartile most of these invole CFW-R2 as it selects far fewer communities at higher homogeneity, but even for quartile two compartively few FH communities.

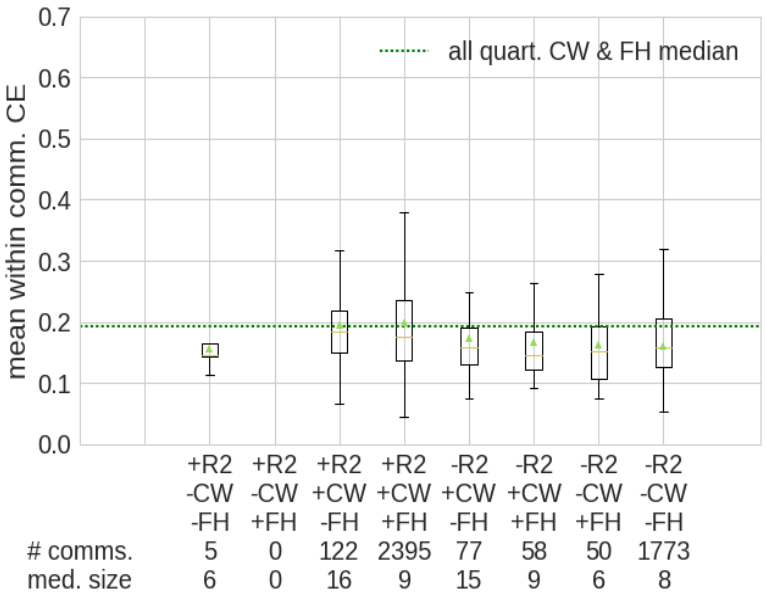


(a) HINT-P: Quantile 3

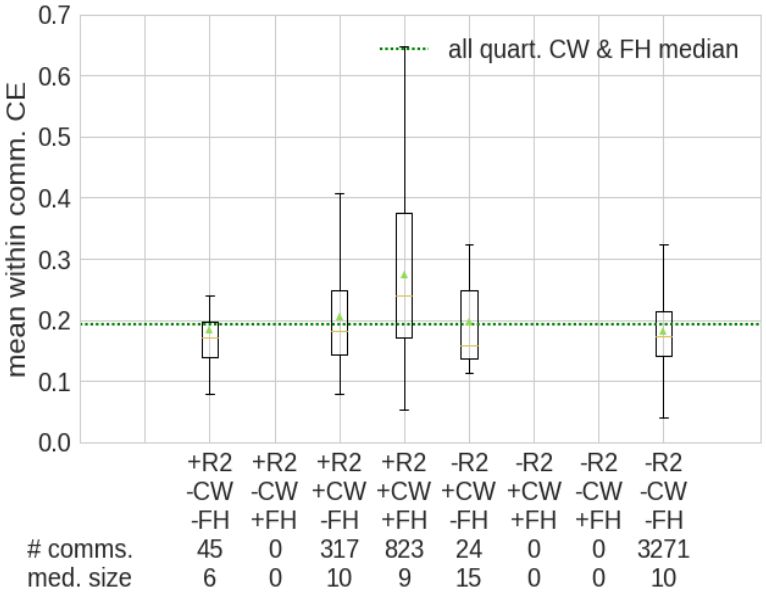


(b) HINT-P: Quantile 4

Figure B.3: Box plots of HINT-P community co-expression statistics for communities in the two highest LHM quartiles showing all communities selected (+) or not (-) by each combination of methods (means for each combination marked with a triangle). The text below each plot shows the number of communities above the median size of communities in that class. Again, CW&FH median co-expression is used as benchmark value (dotted line). Similar co-expression trends exist in both HINT-P and BIOGRID-AP (Figure B.5) with the highest community co-expression levels occurring for cases where two or more of the random walk-based methods intersect.

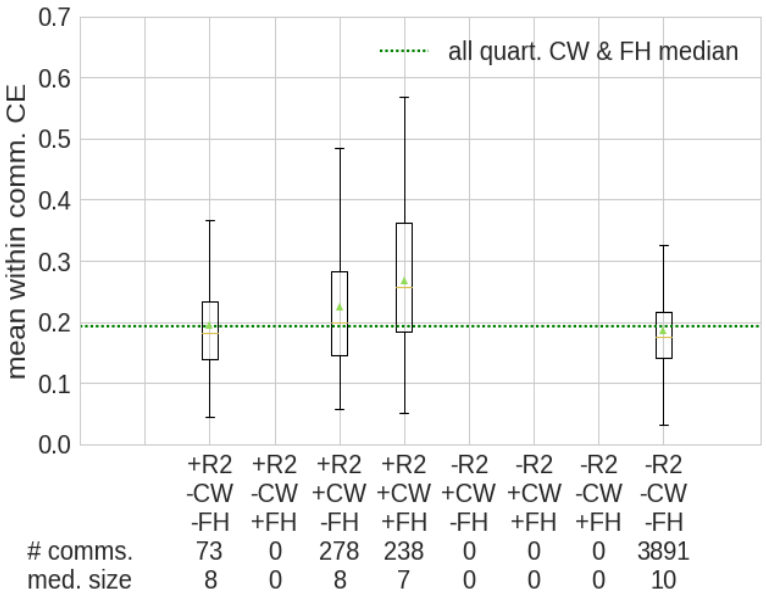


(a) BIOGRID-AP: Quantile 1

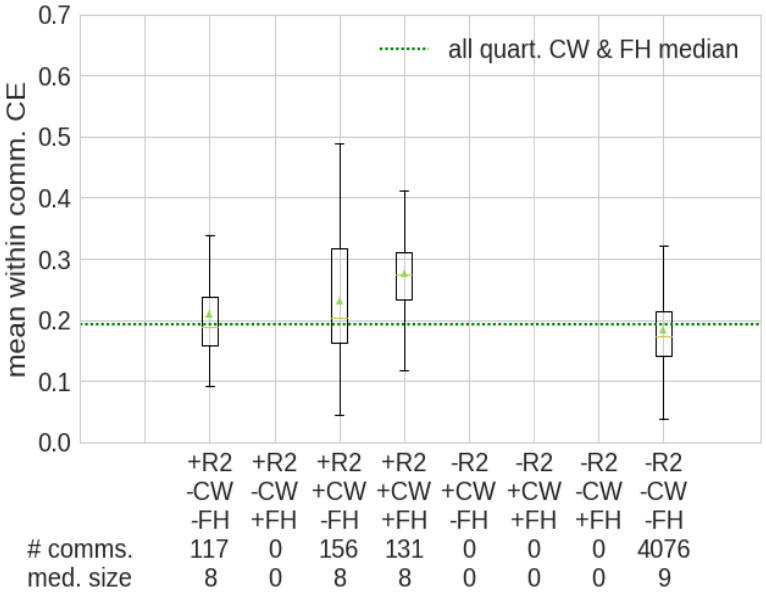


(b) BIOGRID-AP: Quantile 2

Figure B.4: Box plots of BIOGRID-AP community co-expression statistics across community LHM quartiles for all communities selected (+) or not (-) by each possible combination of methods (means for each combination marked with a triangle). The text below each plot shows the number of communities in that class above the median size of communities in that class. Again, CW∩FH median coexpression is used as benchmark value (dotted line). In these higher quartiles most communities are selected by all three methods, with less variation between classes. Note that by quartile two there are very few CW only and no FH only communities.



(a) BIOGRID-AP: Quantile 3



(b) BIOGRID-AP: Quantile 4

Figure B.5: Box plots of BIOGRID-AP community co-expression statistics across community LHM quartiles for all communities selected (+) or not (-) by each possible combination of methods (means for each combination marked with a triangle). The text below each plot shows the number of communities in that class above the median size of communities in that class. Again, CW∩FH median coexpression is used as benchmark value (dotted line). In the lower quartiles CW and FH communities are largely subsumed by CFW-R2 (R2) with medium to high corresponding co-expression scores.

B.4 BBS Neighbourhood Network

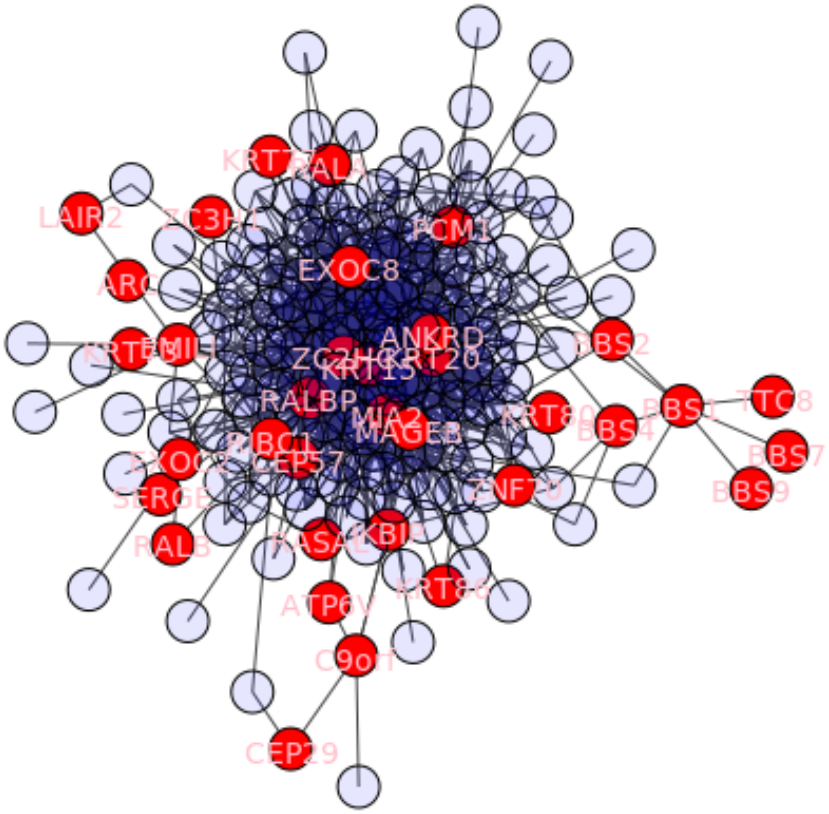


Figure B.6: Graph of the community (red nodes) with the lowest T-value for CFW-R2 not also selected by CW in HINT-P, including neighbouring proteins (blue). Proteins in the community are labelled by their gene symbols. Not all proteins that are thought to have a mechanistic relationship interact directly in the PIN (e.g. BBS4 and BBS7).

C | Hidden Markov Modelling for fMRI Data

This appendix provides additional information that did not appear in in the original methods paper in Chapter 4: *Ranking of Communities in Multiplex Spatiotemporal Models of Brain Dynamics*, including methodological details about regarding pre-processing and training of Hidden Markov Models (HMMs). Figure C.1 shows the Harvard-Oxford Parcellation used throughout Chapters 4 and 5.

C.1 Training of the HMM with Variational Bayes in HMM-MAR

Explicit estimation of the full likelihood by maximisation is often intractable for high dimensional time series data. Approximate methods using Variational Bayes, or Markov-Chain Monte Carlo methods are often used [321]. HMM-MAR is a well-established package for modelling fMRI and EEG data, implemented in Matlab [408], which includes simulation, summary statistics and approximate model inference methods. HMM-MAR employs Variational Bayes (VB), as it is one of the most computationally inexpensive methods for model inference which requires very few additional hyperparameters to be selected, making it simple to train and use in practice.

In brief, the VB framework defines a latent distribution $q(\mathbf{S})$, over the hidden states up to the final observed time T, where $\mathbf{S} = \{S_1, ..., S_T\}$ is an element of the set of possible state trajectories $\mathcal{S}$. The model of the observed data $\mathbf{X} = \{X_1, ..., X_T\}$ given the parameters Θ can be decomposed into

$$\log p(\mathbf{X}|\boldsymbol{\Theta}) = \sum_{\mathbf{S}\in\mathcal{S}} q(\mathbf{S}) \log \frac{p(\mathbf{X}, \mathbf{S}|\boldsymbol{\Theta})}{q(\mathbf{S})} - \sum_{\mathbf{S}\in\mathcal{S}} q(\mathbf{S}) \log \frac{p(\mathbf{S}|\boldsymbol{\Theta}, \mathbf{X})}{q(\mathbf{S})} \tag{C.1}$$

$$= \mathcal{L}(q|\boldsymbol{\Theta}) + KL(q||p_{\mathbf{S}|\mathbf{X},\boldsymbol{\Theta}}), \tag{C.2}$$

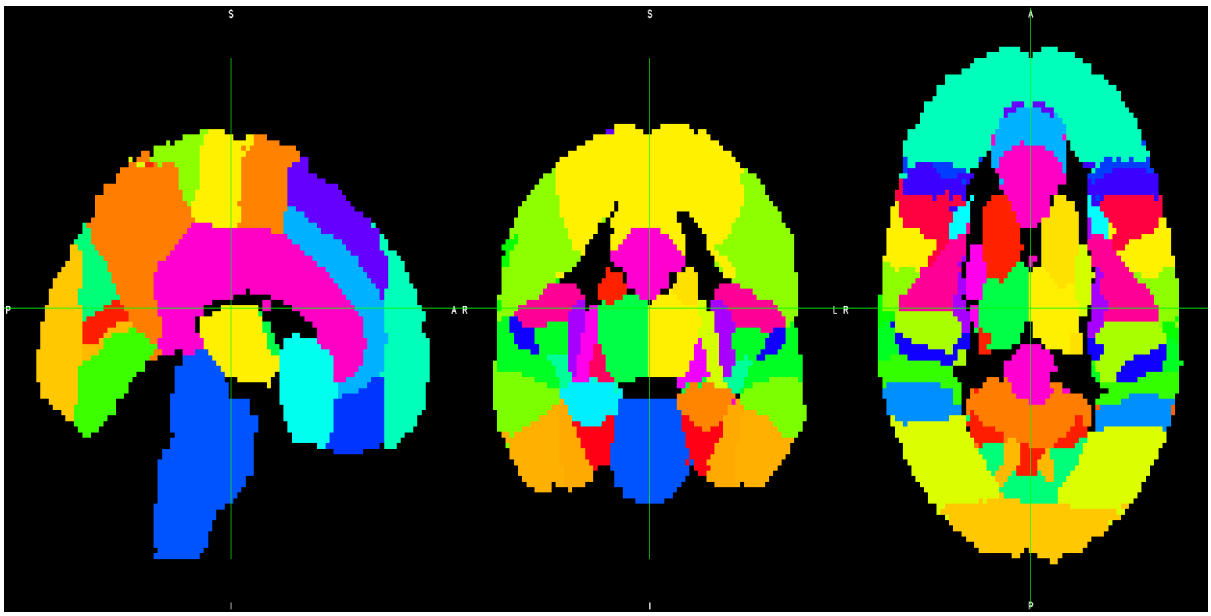


Figure C.1: Cross-sectional views of the combined Harvard-Oxford Cortical-Subcortical (HOCHOS) atlas formed by taking the maximum probability across both atlases to assign a brain region label to each voxel. These three cross-sectional views show the brain region parcellation of the atlas in MNI152 standard space, assigning random colours to each region in the parcellation. The views are along three axes dorso-ventral (D-V), left-right (L-R) and anterior-posterior (A-P). The left cross-section shows a view along the central A-P (oriented right to left) and D-V (oriented top to bottom) axes in which cortical and subcortical structures (including the brain stem in the most posterior portion) are clearly visible. The centre shows a D-V (oriented top to bottom) and left to right (oriented as expected) view in which both thalamii are visible and coloured separately. The last cross section is of the A-P (top to bottom) and L-R view and shows a number of cortical regions common to both hemispheres (such as the frontal pole). This image was generated in FSL.

where log is the natural logarithm and $\log p(\mathbf{X}|\mathbf{\Theta})$ is the state-marginalised log-likelihood of the data. The parameter tuple $\mathbf{\Theta} = (\mathbf{P}, \mu, \Sigma)$ contains all model parameters. Here, q can be viewed as an approximation of $p(\mathbf{S}|\mathbf{X}, \mathbf{\Theta})$ (denoted $p_{\mathbf{S}|\mathbf{X},\mathbf{\Theta}}$). The quantity $\mathcal{L}(q|\mathbf{\Theta})$ is the so-called Evidence Lower Bound (ELBO) for q given the model parameters. It is equivalent to the first term in Eq. (C.1) and can also be written as

$$\mathcal{L}(q|\mathbf{\Theta}) = \sum_{\mathbf{S}\in\mathcal{S}} q(\mathbf{S}) \left[\log p(\mathbf{X}, \mathbf{S}|\mathbf{\Theta}) - \log q(\mathbf{S})\right].$$

The ELBO has the property that $\log p(\mathbf{X}|\mathbf{\Theta}) \leq \mathcal{L}(q|\mathbf{\Theta})$, with the special case $q = p_{\mathbf{S}|\mathbf{X},\mathbf{\Theta}}$ giving $\mathcal{L}(q) = p(\mathbf{X}|\Theta)$, the true model evidence. So, maximising $\mathcal{L}(q|\mathbf{\Theta})$ effectively optimises the log-likelihood $\ell(\mathbf{X}|\mathbf{\Theta}) = \log p(\mathbf{X}|\mathbf{\Theta})$.

The value $KL(q||p_{\mathbf{S}|\mathbf{X},\mathbf{\Theta}})$ is the model Kulback-Leibler divergence between q and $p_{\mathbf{S}|\mathbf{X},\mathbf{\Theta}}$, given by the second term in Eq. (C.1). In general for distributions p and q with shared discrete

support $\mathbf{Z}$,

$$KL(q||p) = -\sum_{\mathbf{z}\in\mathbf{Z}} q(z) \log \frac{p(z)}{q(z)}.$$

It measures the divergence between the posterior distribution for $\mathbf{S}$ given $\mathbf{X}$, and q. The divergence vanishes only when $q = p_{\mathbf{S}|\mathbf{X},\boldsymbol{\Theta}}$, implying that to minimise the divergence is to determine the VB estimator of the model parameters. It can be demonstrated using the law of large numbers, that minimising the divergence is equivalent to maximising $\mathcal{L}(q|\boldsymbol{\Theta})$, the desired outcome [418].

By convention, we obtain an estimate for $p(\mathbf{X}|\boldsymbol{\Theta})$ by iteratively minimising the Variational Free Energy (FE),

$$FE = KL(q||p_{\mathbf{S}|\mathbf{X},\boldsymbol{\Theta}}) - \log p(\mathbf{X}|\Theta) = -\mathcal{L}(q|\boldsymbol{\Theta}).$$

For further details on the iterative minimisation procedure and choice of q see [173, 408].

The result of FE minimisation is a posterior distribution that can be decomposed into an emission model, $p(X_t|S_t, \mu, \Sigma)$, and $p(S_t|S_{t-1}, \mathbf{P})$, the state transition model. Each iteration includes an independent run of the minimisation algorithm for each trial across multiple cycles by updating both components of the FE. In HMM-MAR, iterations are continued until the FE of the model is deemed to have converged, measured by a decrease in the FE by a factor of 10^{-6} or less at the end of fully iterating over all trials.

The Viterbi Path

The Viterbi path is the most likely sequence of states to have generated the sequence of observations $\mathbf{X}$ from the model $\mathcal{M}$. In brief, the algorithm for generating the most likely state sequence $\hat{\mathbf{S}}$ is shown in Algorithm 1 [412]. Note that if breaks exist between patients or phases then the algorithm must be run independently for each T_{kj} where k is the patient and j the phase. While the most likely path is not necessary unique, in the case where the observational model at time t, O_t, is MVN distributed, the support of O_t is $\mathbb{R}^d$ and so the probability of observing two paths with exactly the same likelihood is vanishingly small.

Algorithm 1: Viterbi algorithm for calculating the most probable state trajectory

function Viterbi $(\mathbf{X}, \mathcal{M})$;
Input : $\mathbf{X}$ a sequence of observations, HMM $\mathcal{M} = (\mathbf{S}, \mathbf{P}, \Sigma, \mu, p)$
Output: $\hat{\mathbf{s}}$ the most probable state sequence
define two $|\mathbf{S}| \times T$ (T the length of $\mathbf{X}$) tables, H_1 and H_2
for $s \in \mathbf{S}$ **do**
    $H_1[s, 1] \leftarrow p_s \cdot f(X_1|\Sigma(s), \mu(s))$ where f is the Multivariate Normal (MVN) distribution with mean $\mu(s)$ and covariance $\Theta(s)$
    $H_2[s, 1] \leftarrow 0$
end
for *each observation* $j = 2, 3, 4, .., T$ **do**
    for *each state* s **do**
        $H_1[s, j] = \max_k(H_1[k, j-1] \cdot P_{k,s} \cdot f(X_i|\Sigma(s), \mu(s)))$
        $H_2[s, j] = \arg\max_k(H_1[k, j-1] \cdot P_{k,s} \cdot f(X_i|\Sigma(s), \mu(s)))$
    end
end
define a $T \times 1$ dimensional table $\mathbf{z}$
$z_T \leftarrow \arg\max_k H_1[k, T]$
$\hat{s}_T \leftarrow s_{z_T}$, the z_T^{th} state in state space $\mathbf{S}$
for *each observation* $j = T, T-1, .., 2$ **do**
    $z_{j-1} \leftarrow H_2[z_j, j]$
    $\hat{s}_{j-1} \leftarrow s_{z_{j-1}}$
end
return $\hat{\mathbf{s}} = (\hat{s}_1, \hat{s}_2, .., \hat{s}_T)$

D | Supplementary Information: Ranking of Communities in Multiplex Spatiotemporal Models of Brain Dynamics

D.0.1 Parallel Analysis for Dimensionality Reduction

Let X be a $F \times T$ multivariate matrix with F features and T time points. Dimensionality reduction of X by Principal Component Analysis (PCA) requires the selection of the reduced dimension $d < D$. Parralel analysis allows this to be done in a data driven way by comparing the original data set to surrogate data [175]. In parallel analysis, PCA is first performed on X. The resulting eigenvalues can be ordered $\lambda_{(d)}$ so that d is the d largest eigenvalue.

Next, the columns (time points) of X are permuted within each row removing structure from the dataset and this process is repeated $R = 10000$ times producing $X_1, .., X_R$ surrogate data sets with the same row-wise distribution as X. For each X_r we can obtain a corresponding d largest eigenvalue $\hat{\lambda}_{(d),r}$.

The optimum choice for d is given by the smallest d satisfying

$$\lambda_{(d+1)} < \hat{\lambda}^{P^{th}}_{(d+1)}, \tag{D.1}$$

where $\hat{\lambda}^{P^{th}}_{(d)}$ is the P^{th} of the permuted eigenvalues $\hat{\lambda}_{(d),1}, ..., \hat{\lambda}_{(d),R}$. The value of P determines how much the eigenvalues of the components of X must dominate the eigenvalues of the permuted datasets. We choose the percentile $P = 99$. In other words, the first d components in the original

dataset must each account for more variance than 99% of the permuted components. This was chosen rather than the standard $P = 95$ in order include as much of the signal in X as possible for HMM model training. The dimensionally reduced data set X^* is thus given by

$$X^* = AX \tag{D.2}$$

where A is the eigenmatrix of the first d columnwise eigenvectors of X.

D.0.2 Directed, Weighted Modularity Score

The modularity score $Q(\mathcal{C})$ for a given partition $\mathcal{C} \subset 2^V$ of a weighted, directed graph $G = (V, A)$ with node set V and adjacency matrix A is a measure of how well the partition separates nodes into modules by highly scoring partitions with lower weights on between community edges and higher weights on within community edges. It is calculated as

$$Q(\mathcal{C}) = \frac{1}{m} \sum_{v,v' \in V} \left[A_{v,v'} - \gamma \frac{k_v^{out} k_{v'}^{in}}{m} \right] I(\exists C \in \mathcal{C} \text{ s.t. } v, v' \in C), \tag{D.3}$$

for $m = \sum_{w,w' \in V} A_{w,w'}$, $k_{v'}^{in} = \sum_{u \in V} A_{u,v'}$, $k_v^{out} = \sum_{u \in V} A_{v,u}$ and I the standard indicator function that is one if and only if the condition holds and zero otherwise [274]. The Louvain optimiser require that the inputted modularity matrix $Q'(\mathcal{C})$ to be optimised is symmetric. This is achieved by inputting $Q'(\mathcal{C}) = (Q(\mathcal{C}) + Q(\mathcal{C})^T)/2$.

D.0.3 Community Centrality

The community centrality for the temporal graph $G(P) = (\mathcal{S}, P)$ is calculated using the symmetric undirected version of the transition matrix P to obtain the within community degree centrality z-score, $z(s)$ for $s \in \mathcal{S}$ [160]. Given a partition $\mathcal{U}$ of the temporal graph into communities, this statistic measures how well connected $s \in U \in \mathcal{U}$ is to the rest of its community. The undirected network is based on $G(P') = (\mathcal{S}, P')$, where $P' = (P + P^T)/2$. The score is

$$z(s) = \frac{\nu_s - \nu_U}{\tau_U}, \tag{D.4}$$

where ν_s is the community-specific degree

$$\nu_s = \sum_{s' \in U - s} P'_{s,s'}, \tag{D.5}$$

and ν_U is the the expected centrality over all other nodes in U,

$$\nu_U = \frac{1}{|U|} \sum_{s \in U} \nu_s. \tag{D.6}$$

Lastly τ_U is the standard deviation of ν_s for $s \in U$, so

$$\tau_U = \sqrt{\frac{1}{|U|} \sum_{s,s' \in U} (\nu_s - \nu_U)^2}. \tag{D.7}$$

D.1 Clique-Recovery Simulation Experiments

In order to determine whether state Markov Information matrices (see Eq. (4.4) of the main text) could reasonably be used to recover the spatial community structure of a state, we tested this method on a clique community recovery task in which clique communities are embedded in a noisy covariance matrix generated from an inverse Wishart distribution with scale matrix Ψ of size $D = 63$ (the number of ROIs). We then compare the performance to a simple method for using absolute correlation to generate an undirected graph from a covariance matrix. We also examine the effects of model parameters on clique recovery performance, notably the degrees of freedom (a proxy for the level of noise in estimating the covariance matrix), the within-clique correlation r and the number of principal components used to reduce the matrix dimension (as in Section 4.2.1 of the main text).

The scale matrix Ψ for a predefined partition $\mathcal{C}$ on a graph of order D is given by

$$\Psi_{i,j} = \begin{cases} r & \exists C \in \mathcal{C} : i, j \in C \text{ and } i \neq j \\ 0 & \text{otherwise.} \end{cases} \tag{D.8}$$

Community detection algorithm performance was computed using the Adjusted Rand Index (ARI) [307], which measures the similarity between the true partition $\mathcal{C}$ and that calculated by

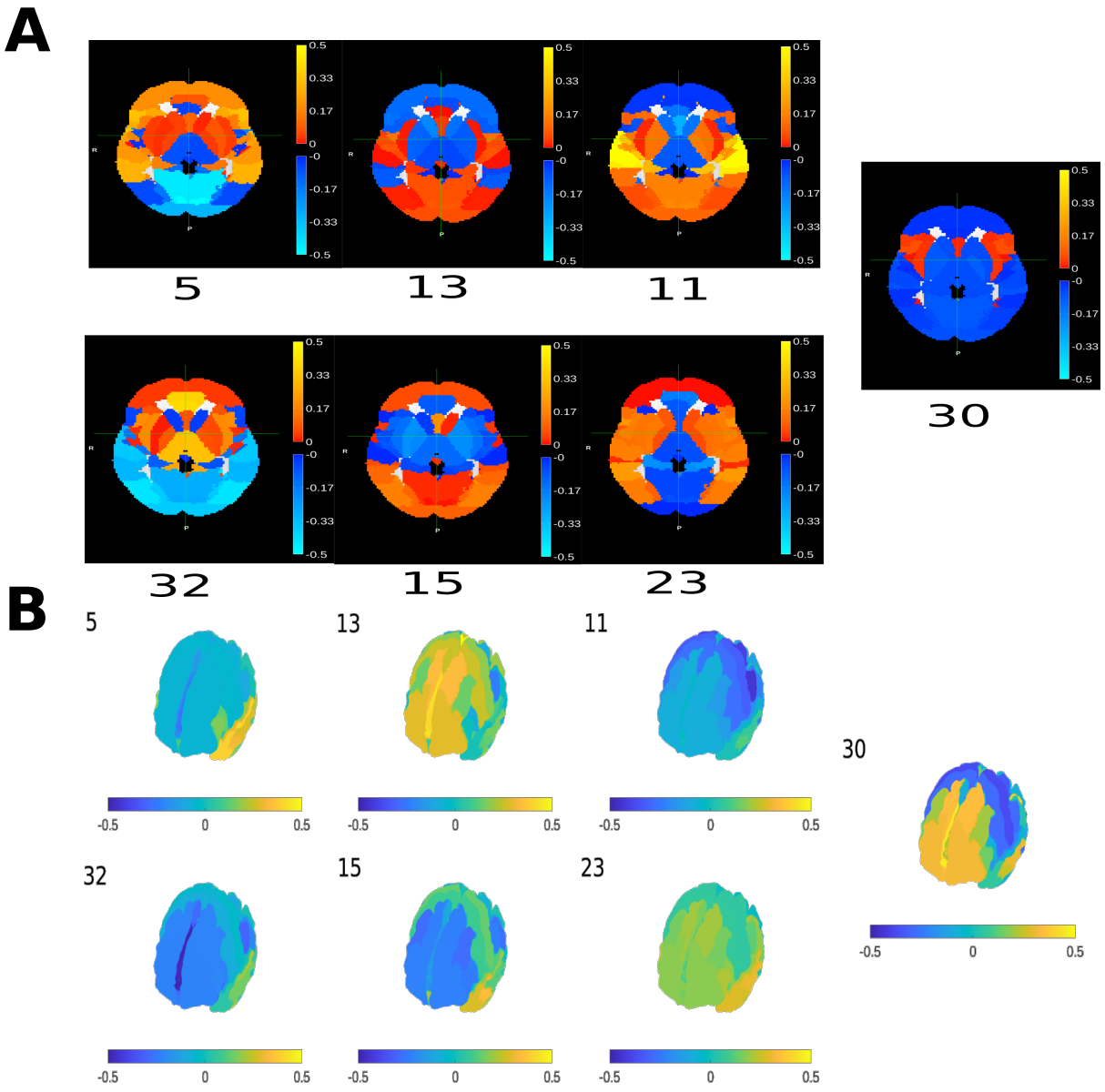


Figure D.1: Hub state mean activity brain maps. (A) The figure shows activity for a central axial slice of the 3D mean activity brain maps of all hub states (the index is determined by the original 33 state HMM). (B) Plots of a surface representation of the same mean activity brain maps for the same hub states.

the Louvain algorithm. A moderate ARI score is 0.6 or above.

In our model preprocessing procedure, data is first dimensionally reduced to reduce noise and complexity. We thus perform the same transformation on the sampled covariance matrix as in Eq. (4.3) of the main text. In order to approximate real partitioning, true partitions were sampled from the set of $K = 27$ partitions determined from data in the main text and 1000 $D \times D$ matrix realisations of the inverse Wishart distribution were generated for each setting of the simulation parameters: degrees of freedom ($df > D + 2$), principal components (PC) and

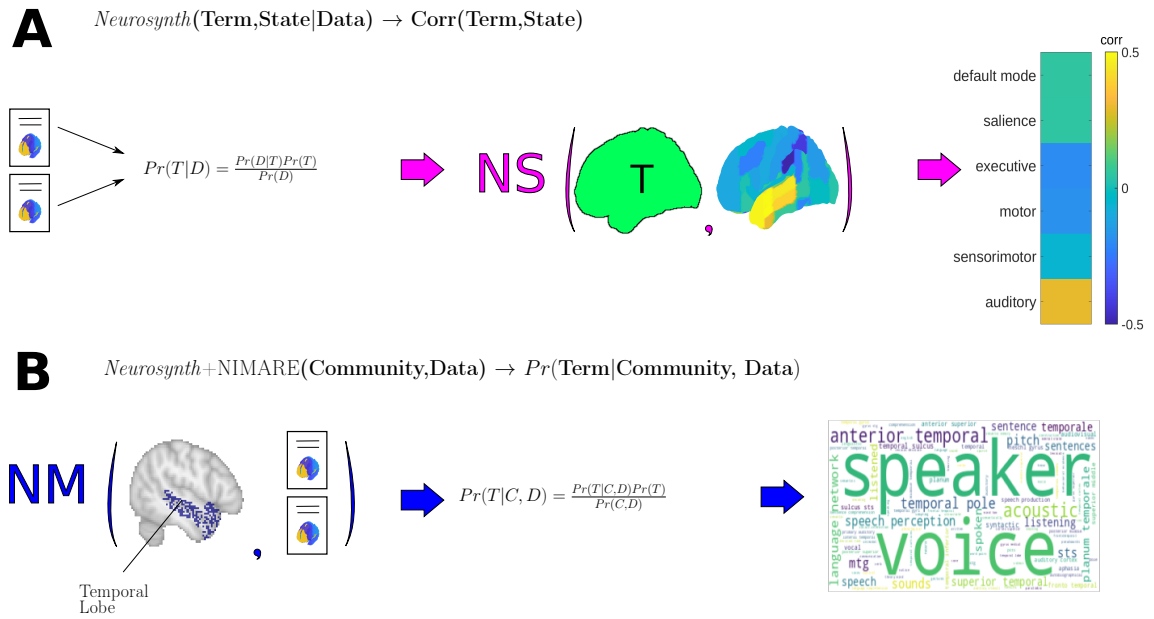


Figure D.2: Diagram overviewing the two *Neurosynth* methods used in our analysis, using the specific example of State 11, Community 1. (A) The first method shows how *Neurosynth* (NS) can be used to generate correlation scores from brain state activity maps and a set of predefined brain terms. In order to perform such a correlation analysis *Neurosynth* requires a corpus of data (D) composed of abstracts with associated brain activation coordinates in (x, y, z) voxel space. We used data from Version 7, dated July 2018 which includes 14,371 studies with 3,178 terms drawn from the data (after removal of numeric characters and words with a frequency less than 1 in 1,000). From this corpus a posterior distribution of term (T) association at each voxel is produced through a Naive Bayes scheme. Term activity maps for a prespecified term are then correlated with the state mean activity brain map and the whole process is repeated for each of the terms of interest. Finally, a full profile of correlations for each term is outputted. (B) For this method we employ the *NiMARE* tool (NM) which uses the *Neurosynth* algorithm to produces a posterior probability over all terms in the corpus for a given selection of voxels. We use as input the voxels defined by the spatial community of brain regions outputted by the community ranking procedure (spatial Community 1 of State 11). The spatial community (C) has been binarised and projected onto the brain map. We show as reference that community activity seems mostly to be located in the temporal lobe, a key region for auditory processing. The final posterior probability of term association is shown as a word cloud where the size of terms is proportional to their predicted probability of association. This method is available in the *NiMARE* Python package.

within-clique correlation (r). Figure D.4 shows the result of these experiments for both methods. The results show that for moderate levels of within-clique correlation, both graph construction methods are able to recover true community activity for a reasonably large range of parameters. Notably, when the number of principal components is either too high (including too much noise) or too low (removing too much signal), community detection performance suffers, underscoring the importance of reasonably choosing this parameter.

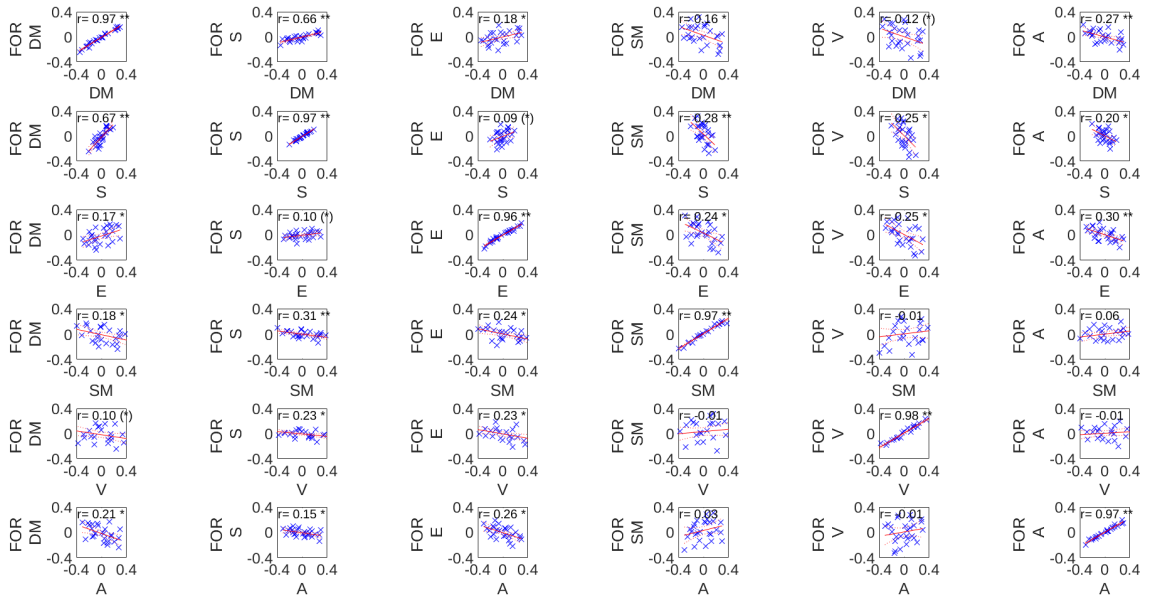


Figure D.3: Grid of scatter plots showing the relationships between *Neurosynth* terms according to their scores for each state using the terms "default mode" (DM), "salience" (S), "executive" (E), "sensorimotor" (SE), "visual" (V) and "auditory" (A) contrasted against the expected score according to the HMM state transition probabilities. Each plot in the grid shows the score for each term associated with a state activity map, plotted against the mean expected score (under the HMM transition probability P) of the next forward (FOR) timestep. The comparison between state score and expected score forward in time demonstrates spatiotemporal relationships between the states (layers) of the network.

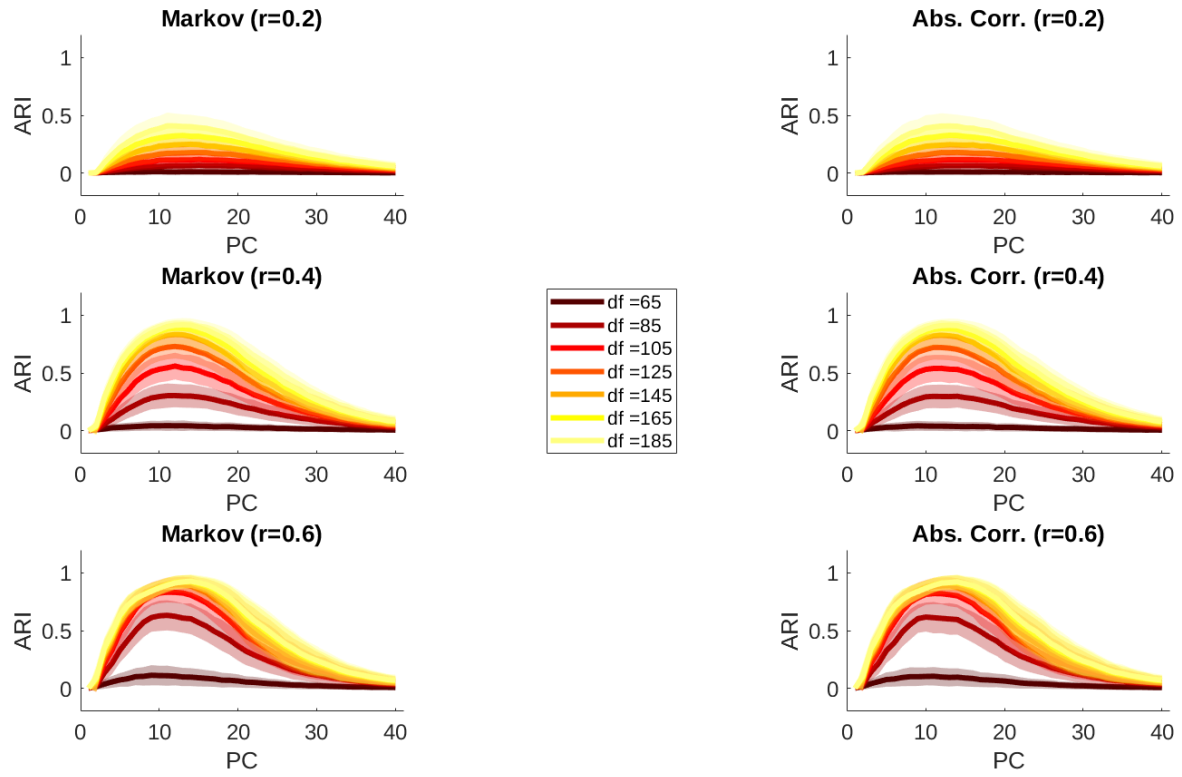


Figure D.4: This figure compares the performance of the Louvain community detection algorithm on clique-recovery tasks for two different methods of constructing graph models from simulated state covariance matrices of size $D = 63$. The first method is the directed Markov Information matrix method (left) and the second is the method of undirected graph construction by absolute correlation (right). ARI scores for a fixed number of degrees of freedom (*df*) are plotted across a variable number of principal components (linear embedding dimension), with higher *df* corresponding to reduced noise in the covariance matrix. Performance is almost identical for both methods with performance improving as either the clique strength or degrees of freedom increases. The relationship between linear embedding dimension (PC) and model performance is not monotonic but rather, improves as more signal-rich components are included, deteriorating when higher noisy components are included.

E | Supplementary Information: Applications of HMGMs to Conditions of Consciousness

Supplementary figures and tables to the results in Chapter 5 on the comparative analysis and modelling of anaesthesia and wakeful brain state activity is provided here. The four phase experimental paradigm alluded to in Chapter 4 (from where resting wakeful fMRI data was acquired) and performed in [262] is shown in more detail.

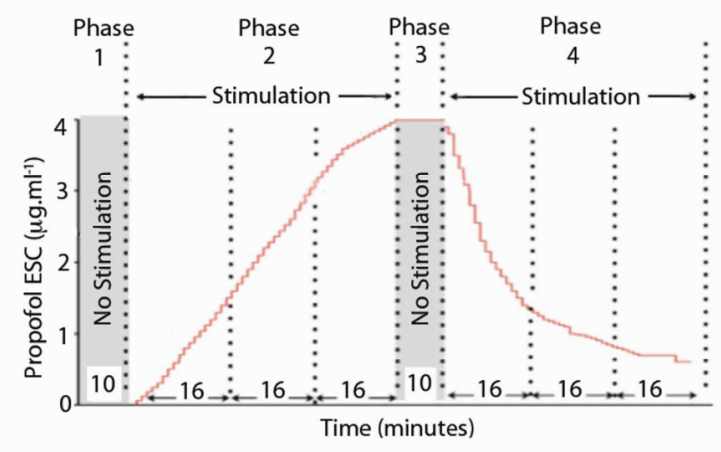


Figure E.1: This figure shows the full four phase protocol for propofol anaesthesia induction and recovery, as a function of anaesthesia Effect Site Concentration (ESC). In the first phase stimulation free eyes closed and resting wakeful activity is recorded. In the second anaesthesia is gradually induced. During this phase, subjects were exposed to noxious and auditory stimulation. In phase three, stimulation was halted for ten minutes and subjects were held at an ESC of 4μg/ml. Finally in phase 4 stimulation was reinitiated and subjects were allowed to recover from anaesthetisation. Our analysis focuses on the two no stimulation periods at peak anaesthetic dose (deep anaesthesia) and wakefulness, however Slow Wave Activity Saturation (SWAS) and Loss of Behavioural Responsiveness (LOBR), as measured by active stimulus response by button press, both occurred in phase 2.This figure was reproduced from [262] by Mhuirchaertaigh et al.

wakefulness ID	comparitive ID
1	1
2	2
3	3
4	4
5	5
6	-
7	6
8	7
9	8
10	9
11	-
12	10
13	11
14	12
15	13

Table E.1: This table shows the IDs of subjects as they appear in the study of wakefulness (Chapter 4) and the comparative analysis of anaesthesia and wakefulness (Chapter 5), where 2 subjects were removed due to extensive burst suppression under anaesthesia.

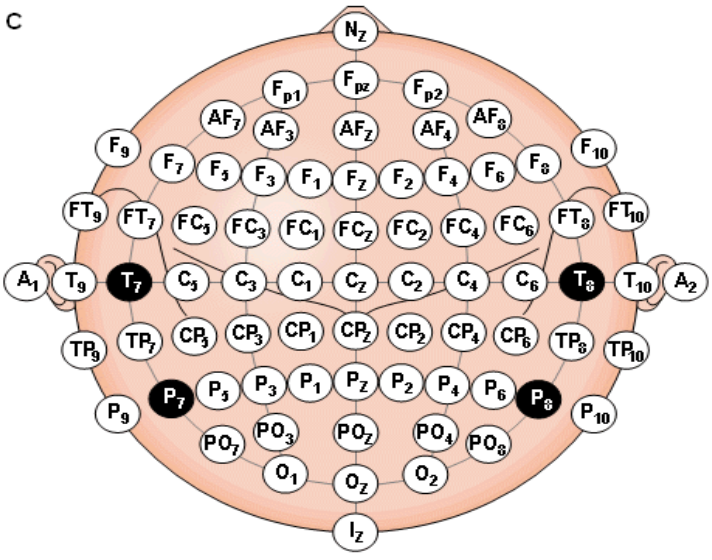


Figure E.2: This figure shows the scalp locations of a standard EEG electrode configuration with the head positioned anterior (top) to posteriorly (bottom). This figure was reproduced from [249] by Malmivuo and Plomsey.

number	labels
1	Fp1
2	Fp2
3	F3
4	F4
5	C3
6	C4
7	P3
8	P4
9	O1
10	O2
11	F7
12	F8
13	T7
14	T8
15	P7
16	P8
17	Fz
18	Cz
19	Pz
20	Oz
21	FC1
22	FC2
23	CP1
24	CP2
25	FC5
26	FC6
27	CP5
28	CP6
29	TP9
30	TP10
31	POz
32	FCz

Table E.2: This table lists the electrode locations of the EEG set-up for both EEG and EEG/fMRI recordings in anaesthesia and wakefulness.

K	20	21	22	23	24	25	26	27	28	29	30
ER_{anaes}	1.145	0.809	0.998	1.131	1.104	1.392	1.174	1.555	1.291	1.586	1.188
ER_{wake}	1.392	1.359	1.388	1.499	1.444	1.529	1.469	1.558	1.644	1.584	1.563

Table E.3: This table shows the Entropy Rates (ER) of Markov chains in anaesthesia and wakefulness HMGM models with different numbers of states (K). We see that even when $K_{anaes} = K_{wake} = K$, $ER_{wake} > ER_{anaes}$ for most models, suggesting that for plausible K, anaesthesia models have less complex dynamics than wakeful ones.

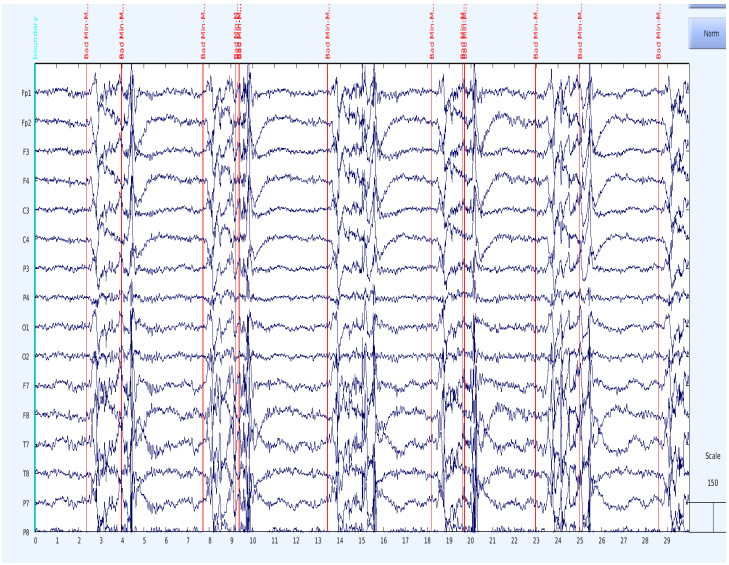


Figure E.3: This figure shows the burst suppressed activity in a subset of electrodes recorded for subject 6 in Chapter 4 (see Table E.1). The red Min-Max markers show the start location of local peaks in activity which are common in the burst regime of burst suppression. This figure was generated in EEGLAB [244].

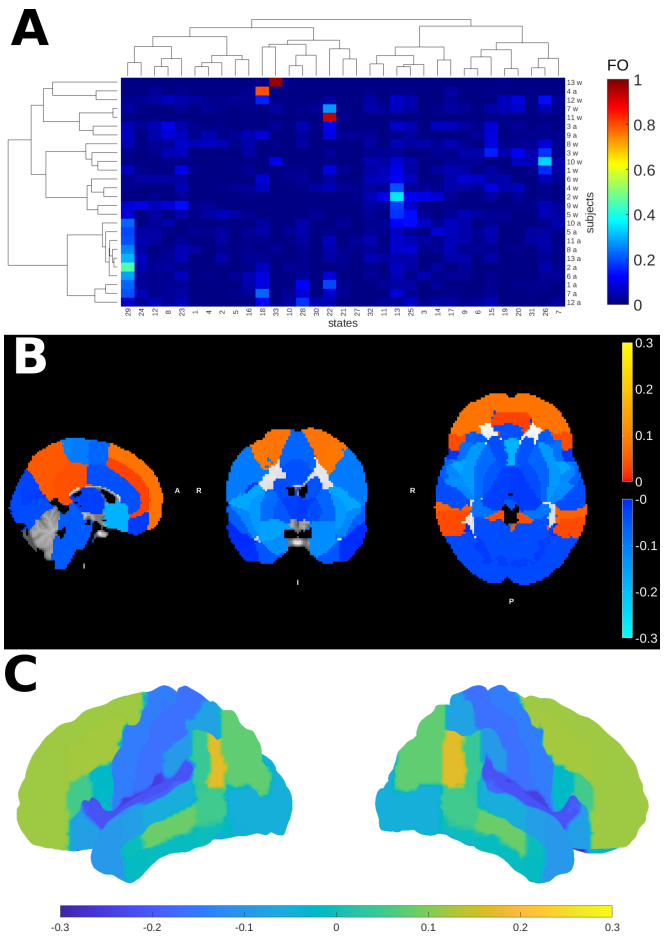


Figure E.4: This figure summarises the results of the anaesthesia-wakefulness model with $K = 33$ states. **(A)** This figure shows a clustergram of Fractional Occupancy (FO) over subject-trial with "w" referring to wakeful trials and "a" for deep anaesthesia trials. As in the original model, anaesthesia trials are dominated by a single state, state 29. As in the original case. The hub state activity is positively correlated with age ($r = 0.439$) and negatively correlated with C_{SWAS} ($r = -0.5050$). **(B)** Cross-sectional activity maps of state 29 also show generally low activity but exhibit clear increases in activity above baseline in the precuneous and prefrontal cortices. This is consistent with networks we found to be active in the original hub states. It is notable that the baseline here includes wakeful activity as well as deep anaesthesia actvity and so cannot be readily compared to the original model. **(C)** The surface plot of state 29 shows the extent of prefrontal activation in the dorsal prefrontal cortex relative to baseline.

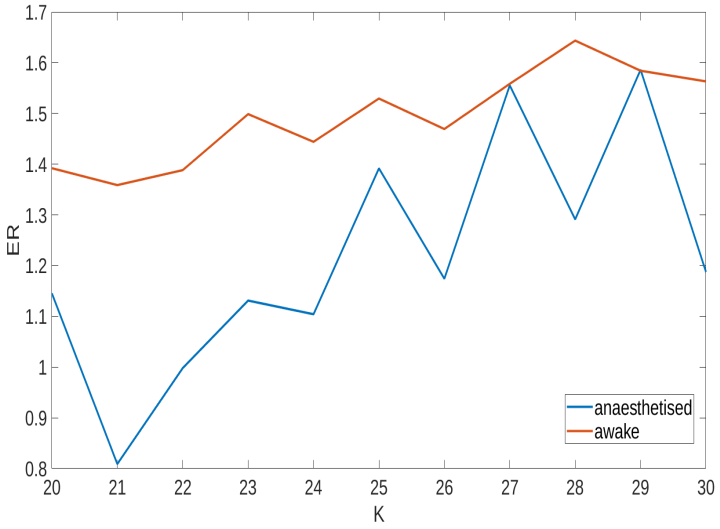


Figure E.5: This figure shows the Entropy Rate (ER) in anaesthesia and wakefulness as a function of the number of states $K \in [20, 30]$. For most models $ER_{wake} > ER_{anaes}$ as shown by the $ER_{wake} = ER_{anaes}$ dotted line. The colour bar indicates the number of K states in the models. Generally higher K correlates with lower difference in $ER_{wake} - ER_{anaes}$ ($r = 0.4350$), as the anaesthesia model becomes more saturated with states.

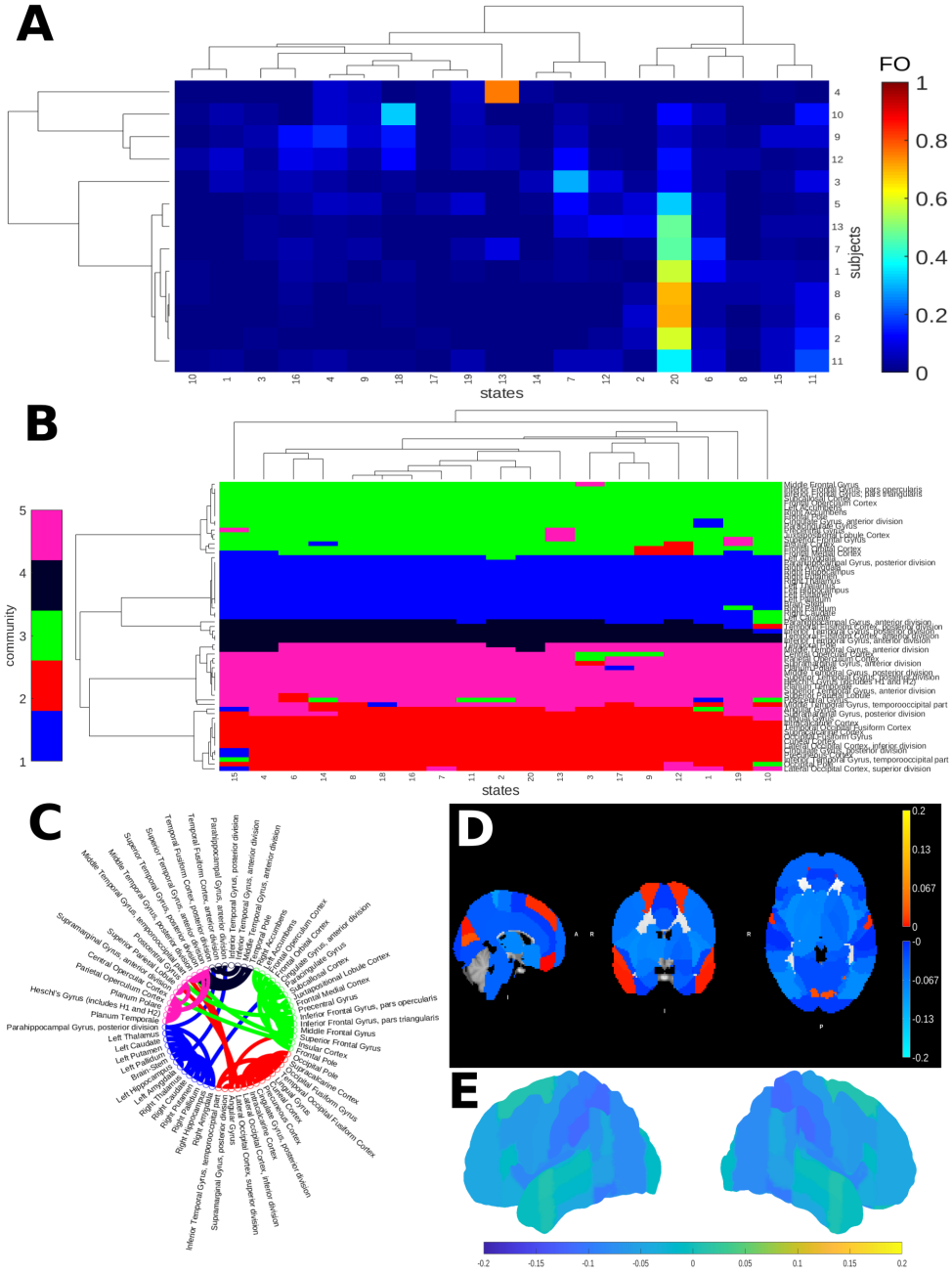


Figure E.6: This figure summarises a model produced by selecting the number of states in the anaesthesia model $K_{anas} = 20$. **(A)** This figure shows a clustergram of Fractional Occupancy (FO) over subjects for the model. As in the original model, subjects are dominated by a single state, state 20. **(B)** The consensus community detection performed across brain state networks again reveals stratification of activity in deep anaesthesia, with the notable difference that most subcortical regions have combined into a single consensus community 1. **(C)** Hub state 20 connectivity is again relatively low between communities with low functional activity across the brain ((**D**) and (**E**)).

	Estimate	SE	tStat	pValue
(Intercept)	-0.1527	0.338	-0.45178	0.66212
age	-0.0095291	0.0066622	-1.4303	0.18641
C_SWAS	0.12753	0.047946	2.66	0.026048
P_SWAS	0.0055897	0.0074109	0.75425	0.46998

Figure E.7: This figure gives model details for the multivariate linear model of state 12 fractional occupancy. The first value rows are for each of the regressors including a constant intercept term as well as subject age as a possible confounder, C_{SWAS}, P_{SWAS}. The first column shows the coefficient estimates, the second the standard error of the estimate and the third the p-value of the Student's t-test for significance of each coefficient. In this model only C_{SWAS} is a significant predictors of state 17 occupancy ($p = 0.0260$), while the others (including age) are not ($p > 0.05$). The model explains significantly more of the variance than the constant model with just an intercept term ($p = 0.0316$, F-test). This model summary was generated in MATLAB.

	Estimate	SE	tStat	pValue
(Intercept)	0.54222	0.69453	0.78071	0.45502
age	0.035315	0.013689	2.5797	0.029709
C_SWAS	-0.24829	0.09852	-2.5202	0.032755
P_SWAS	-0.01557	0.015228	-1.0224	0.33327

Figure E.8: This figure gives model details for the multivariate linear model of state 17 fractional occupancy. The first value rows are for each of the regressors including a constant intercept term as well as subject age as a possible confounder, C_{SWAS}, P_{SWAS}. The first column shows the coefficient estimates, the second the standard error of the estimate and the third the p-value of the Student's t-test for significance of each coefficient. In this model both age and C_{SWAS} are significant predictors of state 17 occupancy ($p < 0.05$), while P_{SWAS} is not ($p = 0.333$). The model explains significantly more of the variance than the constant model with just an intercept term ($p < 0.01$, F-test). This summary was generated in MATLAB.

F | Appendices: Calcium Imaging Analysis, Neuron Extraction and Simulation

F.1 Initial LSM Set-up

The initial LSM set-up by Haslehurst et al. [167] is designed to facilitate fast volume scanning of a brain slice by coordinating the movement of a galvo mirror with an electrically tunable (focal) lens (ETL) which works with the camera. 2D sheets of illuminating laser light are passed through the slice, parallel to the stage at a variable depth. The sheets are produced by focusing collimated laser beams through a cylindrical lens which flattens them out. The beams are generated at the correct wavelength to elicit a fluorescence response from dyes present in the brain slice. The depth of the horizontal cross-section is varied by the movement of the mirror. Images are captured by the camera positioned above the stage.

Structural images stacks can be generated directly before or after an experiment by moving the mirror at regular increments vertically, passing through the slice, stopping to image the dendrite structure at a specific depth for a full 200ms exposure and continuing. Regularly spaced images, 0.5μm apart, generated in this way and stacked together form a 3D (XYZ) structural representation of a dendrite.

Dynamic volume-scanning images require careful coordination of the galvo mirror and ETL (see Figure a in F.1). By moving through a one dimensional continuous (sinusoidal or saw-tooth) trajectory (see b) from the top to the bottom of the slice, illuminating light can be passed over the full slice volume in a single exposure, fluorescing all dyed regions. In order to ensure the

image is in focus, the ETL follows the trajectory of the mirror but at a constant lag (P) to allow for re-focusing. By repeating this pass through the slice multiple times, a stack of 2D (XYT) images with a fixed depth of field and time window (20ms) is generated. These images form the basis of Ca^{2+} dynamics analysis with the fast volume-scanning method.

Figure F.1, c, shows the relative positioning of the camera, ETL and galvo mirror as well as the other components such as the fluorescent light source and cylindrical (sheeting) lens. Continued light exposure causes bleaching of the dye and has phototoxic effects that damage the dendrite (see Section 6.2.4). Future utilised an improved version of this set-up that causes less bleaching and will be capable of completing multiple exposures in one depth scan with the galvo mirror. Future experiments will therefore be able to produce 3D representations of Ca^{2+} dynamics by stacking these images (XYZT stacks).

F.2 Aikake Information Criterion

The Aikake Information Criterion (AIC) is a metric used in model selection with lower AIC corresponding to models with. The AIC uses a penalty term to approximate the true likelihood function from a set of candidate models. Given a set of possible models $\mathcal{M}_1, \mathcal{M}_2..., \mathcal{M}_n$ for a given set of observations X, each model having corresponding number of parameters p_i, the AIC will tend to converge to the model (lowest AIC) that best approximates the true model of the data. The AIC_i for an individual candidate model $\mathcal{M}_i$ is

$$\text{AIC}_i = -\ell(X|\mathcal{M}_i) + 2p_i$$

where $\ell(X|\mathcal{M}_i)$ is the log-likelihood of model i.

F.3 APP2: Automated 3D Neurite Tracing with the Fastmarching Algorithm

APP2 is an improved version of the original APP algorithm that utilizes an improved pre-processing method, the Grey-scale Image Distance Transform (GSDT) and an augmented pruning procedure. Key to APP2 and APP is the fastmarching (FM) algorithm which is a tree growing procedure that functions like a cellular automaton that starts from a seed set of alive

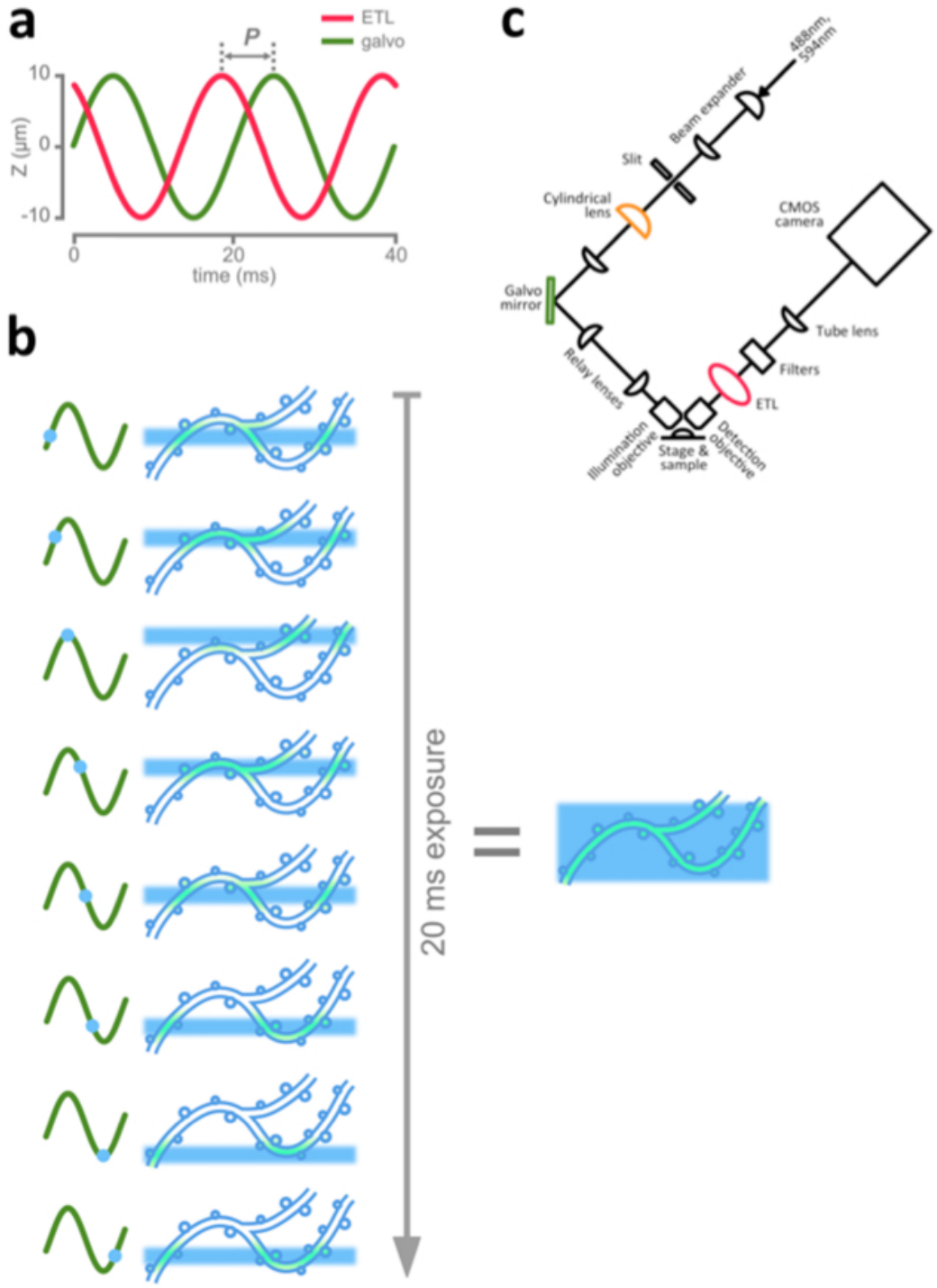


Figure F.1: Initial set-up of the Light Sheet Microscope (LSM) in the Emptage Lab with fast volume scanning capabilities. a.) Shows how the movement of the galvo mirror and Electrically tunable Lens (ETL) are kept out of phase to give sufficient lag time for the lens to adjust to the signal. b.) Shows the movement of the galvo mirror to fluoresce a thin cross-section of the slice at it moves through 160μm in the Z-axis over a 20ms exposure. The result of this single exposure is a static 2D composite image with a depth of 160μm. c.) Is a schematic diagram of the LSM showing the CCMOS light source, ETL and galvo mirror in relation to the stage (where the slice is placed for imaging) and other stationary components [167].

```
VAR Estimation Results:
========================
Endogenous variables: DF1, DF2
Deterministic variables: both
Sample size: 493
Log Likelihood: 780.308
Roots of the characteristic polynomial:
0.8098 0.8098 0.6858 0.6858 0.6356 0.6356 0.5907 0.5907 0.4763 0.4763
Call:
VAR(y = ROIs.frame[dfs], p = which.min(varselect$criteria[1,
]), type = "both")


Estimation results for equation DF1:
====================================
DF1 = DF1.l1 + DF2.l1 + DF1.l2 + DF2.l2 + DF1.l3 + DF2.l3 + DF1.l4 + DF2.l4 + DF1.l5 + DF2.l5 + const + trend

Estimate Std. Error t value Pr(>|t|)
DF1.l1 -8.998e-03  4.522e-02  -0.199  0.84235
DF2.l1 -1.174e-02  8.347e-02  -0.141  0.88816
DF1.l2  1.138e-01  4.457e-02   2.554  0.01095 *
DF2.l2 -7.221e-03  8.416e-02  -0.086  0.93167
DF1.l3  9.864e-02  4.463e-02   2.210  0.02756 *
DF2.l3 -5.043e-02  8.246e-02  -0.612  0.54112
DF1.l4  1.180e-01  4.483e-02   2.632  0.00876 **
DF2.l4  1.003e-01  8.342e-02   1.202  0.22999
DF1.l5  8.571e-02  4.519e-02   1.897  0.05846 .
DF2.l5  1.735e-01  8.281e-02   2.095  0.03669 *
const  -8.171e-04  1.379e-02  -0.059  0.95276
trend   3.337e-06  4.764e-05   0.070  0.94418
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.1505 on 481 degrees of freedom
Multiple R-Squared: 0.07213,Adjusted R-squared: 0.05091
F-statistic: 3.399 on 11 and 481 DF,  p-value: 0.0001484


Estimation results for equation DF2:
====================================
DF2 = DF1.l1 + DF2.l1 + DF1.l2 + DF2.l2 + DF1.l3 + DF2.l3 + DF1.l4 + DF2.l4 + DF1.l5 + DF2.l5 + const + trend

Estimate Std. Error t value Pr(>|t|)
DF1.l1  1.881e-03  2.461e-02   0.076 0.939094
DF2.l1  2.075e-01  4.543e-02   4.568 6.27e-06 ***
DF1.l2 -4.115e-03  2.426e-02  -0.170 0.865354
DF2.l2  6.227e-02  4.580e-02   1.360 0.174614
DF1.l3 -6.212e-02  2.429e-02  -2.557 0.010849 *
DF2.l3  1.591e-01  4.488e-02   3.546 0.000429 ***
DF1.l4  6.280e-02  2.440e-02   2.574 0.010354 *
DF2.l4  1.109e-01  4.540e-02   2.442 0.014955 *
DF1.l5 -2.993e-02  2.459e-02  -1.217 0.224133
DF2.l5  5.446e-02  4.507e-02   1.208 0.227476
const   6.200e-04  7.504e-03   0.083 0.934183
trend  -1.043e-06  2.593e-05  -0.040 0.967934
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.08191 on 481 degrees of freedom
Multiple R-Squared: 0.1723,Adjusted R-squared: 0.1534
F-statistic: 9.104 on 11 and 481 DF,  p-value: 7.354e-15



Covariance matrix of residuals:
DF1       DF2
DF1 0.0226503 0.0001256
DF2 0.0001256 0.0067088

Correlation matrix of residuals:
DF1     DF2
DF1 1.00000 0.01019
DF2 0.01019 1.00000
```

Figure F.2: Output showing the ANOVA results of fitting a VAR(5) model (selected by minimised AIC) to the normalised fluorescence time series associated with ROIs c and d. The time series DF1 is associated with ROI d and DF2 with ROI c. The roots of the characteristic polynomial, like for the univariate case, indicate that there are no unit roots and thus no co-integration. Coefficients associated with DFx at lag y are denoted DFx.ly. Equations for both DF1 and DF2 admit significant cross-coefficients from the other time series (t-test, $p < 0.05$). The fit to both time series includes a number of highly significant terms and the F-tests for any significant model coefficients are both highly significant ($p < 0.001$).

pixels (A) and grows by testing trial pixels (T) which then become alive if they are sufficiently close in distance and pixel intensity. I describe here the GSDT and tracing algorithms, both presented by Xiao et al [432], which follow the FM framework.

F.3.1 GSDT

The goal of the transform is to obtain the minimum distance between the foreground pixels and background pixels in a way that takes into account both Euclidian pixel distance and pixel intensity. The foreground is defined as all voxels above a generally low, user-defined threshold. A low threshold is preferred, since pruning, see Section F.2, will remove many more voxels from the tree. This distance function is then fed into the tracing procedure rather than a simple euclidian distance.

In GSDT pre-processing pixels are thought of as existing in a lattice in which immediate (not diagonal) neighbours are connected by directed edges. The edge-weights are given by the function

$$e(\mathbf{v}_1, \mathbf{v}_2) = ||\mathbf{v}_2 - \mathbf{v}_1||_E . \mathbf{I}(\mathbf{v}_2)$$

where $\mathbf{v}_1 = (x_1, y_1, z_1)$ and $\mathbf{v}_2 = (x_2, y_2, z_2)$ and $\mathbf{I}$ is the pixel value (intensity) function in 3D structural space taking on the intensity values of the deconvolved image. The algorithm proceeds recursively by updating a distance function. The distance function, $d(\mathbf{v})$, is defined for all pixels with initial value

$$d(\mathbf{v}) = \begin{cases} \mathbf{I}(\mathbf{v}) & \mathbf{v} \in \text{background} \\ \infty & \mathbf{v} \in \text{foreground} \end{cases}$$

In GSDT all background pixels are alive seeds (A) and there foreground neighbours are initially set to trial (T). From the set of T pixels the method extracts the one pixel $\mathbf{v}^*$ that has the current minimum distance to the alive set (the background at this point) which then becomes A and any not A neighbours of $\mathbf{v}^*$ are set to T. Finally the distance function is updated to

$$d(\mathbf{v}^*) = \min_{\mathbf{v} \in nb(\mathbf{v}^*)} d(\mathbf{v}) + f(\mathbf{v}^*)$$

where $d(\mathbf{v})$ is given by

$$d(\mathbf{v}) = \min[d(\mathbf{v}), d(\mathbf{v}^*) + e(\mathbf{v}, \mathbf{v}^*)] \tag{F.1}$$

The recursive update call continues until the trial set is exhausted (i.e. foreground depleted). The FM framework is the selection of the nearest $\mathbf{v}^*$ and change from A to T of its neighbours followed by the recursive distance update using the general form in Eq. (F.1).

It has been observed that the soma is generally the brightest and point in the image that is furthest from background and so the GSDT provides a convenient automatic soma detector using $\arg\max d(\mathbf{v})$ which can be used a seed point in the automatic tracing algorithm.

F.3.2 Neuron Reconstruction

Neuron reconstruction proceeds by constructing a parental map, alongside the distance tree, starting from the seed pixel $\mathbf{s}$ and going to the rest of the foreground pixels. The initial distances are given by GSDT. For convenience, pixels in this FM-style algorithm that are neither T or A will be labelled as far (F). In the initiation step, the parent of each pixel is set to itself, i.e. $par(\mathbf{v}) = \mathbf{v}$. Then, as before, $\mathbf{s}$ is set to A and each of its neighbours $\mathbf{v}$ is to T and at the same time $par(\mathbf{v}) = \mathbf{s}$. Then, in the recursive step the method again selects the $\mathbf{v}^*$ with minimum distance from amongst these neighbours and set its neighbours to T and their parent to $\mathbf{v}^*$ if they are F. If they are not F then only change the parent to $\mathbf{v}^*$ if the update step is successful, i.e. if

$$d(\mathbf{v}) = \min[d(\mathbf{v}), d(\mathbf{v}^*) + e(\mathbf{v}, \mathbf{v}^*)] = d(\mathbf{v}^*) + e(\mathbf{v}, \mathbf{v}^*).$$

Again, the updates are run until there are no more T pixels. Here the edge weights $e(x, y)$ over the pixel lattice is defined differently with

$$e(\mathbf{v}_1, \mathbf{v}_2) = ||\mathbf{v}_1 - \mathbf{v}_2||_2 . \frac{g_I(\mathbf{v}_1) + g_I(\mathbf{v}_2)}{2}$$

where

$$g_I(\mathbf{v}) = \exp\left(\lambda \left(1 - \frac{\mathbf{I}(\mathbf{v})}{\mathbf{I}_{\max}}\right)^2\right)$$

where λ is a parameter found to be stable at 10, and $I_{\max}$ is the maximum pixel value. Updates are then more successful when pixel values are high and edge weights are thus low.

Biological realism of the tree is improved by allowing the width of the tree to vary and unrealistically short branches to be pruned. The width of the tree at a point is estimated by a ball containing a low background fraction (threshold fraction hard-coded). Finally the tree is pruned via a hierarchical pruning procedure to remove short, biologically implausible, dendritic branches.

Hierarchical Pruning

In order to prune the tree the segments are ordered by importance. The ordering is produced as follows. First, the leaf segment is selected that has the longest graphical distance from the leaf node to the seed node (inside the soma). This segment is recorded and deleted producing a new tree. The process is repeated on the new tree to exhaustion creating a decreasing order of importance over all segments.

Segments are pruned in a recursive way, in decreasing order of importance, by calculating the overlap between the union of the volumes covered by all the nodes in the segment, termed the coverage volume. This coverage volume is then compared to the volume already covered by more important segments. A segment is deleted along with all of its children (pruned) if the percentage of in-segment nodes covered by the volume of a more important segment is more than 75%. The voxel intensity is also incorporated so that brighter on-average segments tend to be kept. The rule is that a segment S (a set of node voxels) is pruned if

$$\sum_{\mathbf{v}\in S} \mathbf{I}(\mathbf{v}) \frac{\mathbb{I}_{[\mathbf{v}\in C_S]}}{\sum_{\mathbf{u}\in S} \mathbf{I}(\mathbf{u})} \geq 0.75, \tag{F.2}$$

an intensity-weighted fraction of voxels, where C_s is the set of voxels in the coverage volume of more important segments than S, and $\mathbb{I}_{\cdot}$ is an indicator variable that is one if condition $\cdot$ is true and zero otherwise. The thresholding process is repeated until either the set of segments is exhausted or no more segments satisfy Eq. (F.2). Finally leaf segments are pruned by simple thresholding where all leaf segments shorter than a user-specified length L are deleted. Once all of these steps have been carried out sequentially by the APP2 algorithm, the resultant tree volume can be segmented into somatic, dendritic and axonal components for modelling. I found

that setting $L = 5$ voxel widths resulted in trees that were realistic with fewer noisy branches.

F.4 3D Blur Reduction with Image Deconvolution

The process of image acquisition in fluorescence microscopy usually produces distortions that are dependent on the medium, resolution, position of the object, exposure, shape of the beam and the wavelength of fluorescing light. Some of these factors place a hard limit on the resolution of the image, such as the wavelength of light used. This is termed the diffraction limit and this LSFM system is already close to its diffraction limit for orange light (roughly $0.2 - 0.4\mu m$). Other forms of distortion, such as laser beam shape, produce blur which can be corrected [356].

All fluorescence microscopy images contain blur due to the divergence of the laser beam radius with distance from source to objective (the object being imaged). This leads to a 2D blur effect that effects each image plane independently. The 'spread' of intensity causes the intensity at any one point in the image to be a weighted sum of the intensity of its neighbours. However, in this LSFM set-up the fluorescing beam is curved in 3D space, resulting a curved image 'plane' that leads to blur in all three dimensions [405].

Light intensity at a point in 3D space in curved beam fluorescence microscopy (such as in the case of Airy beam technology) image stacks can be modelled as an an observed intensity $\mathbf{J}(x, y, z)$, which is a caused by the interaction between the objective intensity (of the 'true' unblurred image) $\mathbf{I}(x, y, z)$ and a point spread function (PSF_{xyz}) that models the spread (weighted sum) of intensity (blur) caused by the beam. In continuous 3D space (in the ideal case) this can be expressed as a triple integral

$$\mathbf{J}(x, y, z) = \int\int\int \mathbf{I}(u, v, w) PSF_{xyz}(u, v, w) dudvdw.$$

Note that this PSF is shift-variable that is to say that the form of spread depends on the (x, y, z) coordinates of the pixel in 3D space due to the curved path of the beam passing through the slice. Such a curved beam, helps to uniformly illuminate the image plane, thereby reducing the light intensity needed as well as photo-bleaching (depletion of fluorescence due to the non-reversible photon-fluorophore reaction).

The exact form of the PSF used in the LSM experiment and the way in which the object

intensity $\mathbf{I}(x, y, z)$ is extracted from the 3D image stack is proprietary and developed by M2 Lasers specifically for this LFSM set-up. One hundred iterations of the deconvolution algorithm was used to obtain the final pixel intensity estimate [405]. Note that no PSF approximation method is perfect and the obtained approximate $\mathbf{I}$ may vary in a FOV position (in (x, y, z)) and object dependent way from the true object intensity.

F.4.1 The NEURON Simulation Environment

The NEURON Simulation environment is a widely used package written in C++ with GUI and Python plug-in support. It is built around a compartmental model of voltage diffusion in and between neuronal micro-compartments. The equations that govern this interchange of ions are known as Kirchoff's current laws, for which the voltage flux is described by the differential equation

$$c_j \frac{dv_j}{dt} + i_j = \sum_k \frac{v_k - v_j}{r_{jk}} + i_{ext},$$

where c_j is the membrane capacitance of compartment j. The variable i_{ion_j} is the Coulombic current due to sources or sinks of ion to or from j (generally channels or leaks). The value r_{jk} is the resistance to current from j to k. Lastly, i_{ext} is any external current source or sink. This is a law of conservation of current (the change in voltage at j is due to the current from other compartments plus whatever is injected or channelled). The sign conventions are current flowing into j from k and outward from the cell are positive by convention.

A whole cell is usually modelled as a tree setting up child-parent relationships between compartments. Most compartments are 3D tubes while the soma is generally modelled as a single sphere. NEURON follows a simple numerical integration scheme (defaulting to backward Euler) but different mechanisms may follow different schemes. All mechanisms are integrated in step (usually on the order of 1ms) to maintain agreement. This does not always work for mechanisms operating at vastly different flux or time scales, as occurs for different ion channels, so NEURON provides some forms of mechanism-wise integration.

While Coulombic current and voltage are tracked and their diffusion across compartments calculated, specific ion concentrations are handled by the modeller who writes code describing the sources and sinks of current (broadly ion channels) and the neuronal architecture (including

electrostatic properties of compartments and their channel distributions). NEURON can separately keep track of internal-external ion specific and non-specific (mixed) currents across a compartment with specific support for Na^+, K^+, Cl^- (inhibitory) and Ca^{2+} currents.

F.4.2 The Biophysical Model (Branco, 2010)

The paper our model was originally based on is [61]. This model was updated in [361], and continues to receive updates up until February 2021. Although the neuron modelled is from a different region of the brain, it is also a neocortical pyrimidal cell with similar channel compositions. Differences include branching patterns and precise channel distributions. We modify this framework to accurately reflect Ca^{2+} flux through the N-Methyl-D-aspartic Acid Receptor (NMDAR), an important synaptic channel protein for Ca^{2+}, based on measurements of Ca^{2+} permeability from [189].

All of the channels included in this model are also found in CA1 pyramidal cells though the model was originally developed for pyramidal cells from another brain region. These channels are voltage dependent K^+, Ca^{2+} and Na^+ channels as well as Ca^{2+}-dependent K^+ channels and non-specific channels for Na^+ and K^+ currents (inhibitory synapses and channels are not commonly modelled). At present a built in depletion mechanism is included for Ca^{2+} lost to outward channels and pumps (but not buffers). These help to restore the Ca^{2+} flow to the zero flux equilibrium.

Also included are pre-synaptic and spine mechanisms, defining a rudimentary synapse. At present, these synapses are defined as point processes (see Section 6.5.1) with negligible volume or surface area. pre-syaptic activation always results in the same glutamate volume released to NMDA and AMPA receptor-channels on the postsynaptic (spine) side. We augment this slightly to include silent synapses with no functional AMPA receptors and limited NMDAR activity.

F.4.3 Ion Channels

Flow of ions through a channel (ionic current) depend on several key factors including its specificity (both in ion-type and direction) as well as external factors such as the intra and extracellular differences in concentration. Where possible then, it is useful to take ion-specific potentials into account. By Ohm's law, the current into a compartment j from a particular type of external

channel is

$$i_{channel_j}(t) = g_{channel_j}(v_j, t)[v_j - E_{channel}] \tag{F.3}$$

where $g_{channel_j}$ is the conductance (reciprocal of resistance, $g_{channel_j} = r^{-1}_{channel_j}$) through the channel, which may vary with voltage and $E_{channel}$ is the reversal potential (the membrane voltage at which the net ion flux through the channel is zero). Reversal potentials for ion-specific channels use that ions reversal potential, E_{ion}, while those that accept a mixture depend on the channel's fractional permeability to each ion and so have been derived by experiment. For ion specific channels, E_{na} and E_{ca} are both positive (reflecting the higher extracellular concentrations), while for E_k this value is negative (reflecting the opposite). Here, we use lowercase element symbols rather than the full ion symbol (i.e. Ca^{2+}, K^+ etc.) to remain consistent with NEURON documentation.

Note that more detailed models calculate the potential via the Nernst equation which requires detailed accounting of external and internal ion concentrations [370]. Our modified model uses fixed, equilibrium reversal potentials and internal and external ion concentrations that are less computationally expensive.

Channel States and Conductances

Models of most channel conductances are complex and apart from v_j, can depend on the presence or absence of neurotransmitters as well as other ions or small molecules that may block the channel. The conductance through a single channel is modelled by a set of states. These state mechanism can be as simple as switching between open and closed in a voltage dependent manner (a two state model), to models that include 8 states or by combining open/closed, neurotransmitter bound/unbound and blocker bound/unbound states or even more intermediary states. In all, usually only one of these states allows significant current flow.

AMPAR Channels

The most simple channel mechanism in the model is found specifically at spines (the input side of the synapse). These are the primary channels that induce the massive local increase in membrane potential that allows ion specific channels to become active. The total conductance of the AMPAR channels on the l^{th} synapse in compartment j, $g_{AMPAR_{jl}}$, depends only on the

pre-synaptic release of the neurotransmitter glutamate, which is considered an instantaneous event. The conductance is

$$g_{AMPAR_{jl}}(t) = \begin{cases} 0 & t < t_a \\ \bar{g}_{AMPAR}\left[e^{-\frac{t-t_a}{\tau_2^{AMPAR}}} - e^{-\frac{t-t_a}{\tau_1^{AMPAR}}}\right] & t \geq t_a \end{cases}, \tag{F.4}$$

where t_a is the activation (neurotransmitter release) time, $\tau_1^{AMPAR} = 0.5$ and $\tau_2^{AMPAR} = 1$ and $\bar{g}_{AMPAR} = 1$ is the maximum conductance through the entire population of channels on the spine. This last value, $\bar{g}_{AMPAR}$ is treated as a constant across synapses and represents the maximum channel conductance through the spine. This is a simplifying assumption as the conductivity of synapses varies according to their dimensions and specific channel content.

AMPARs are non-specific channels that are permeable to both K^+ and Na^+ ions. The result is a near zero reversal potential. The equation for the current is then (by Ohm's Law of electrostatics) $i_{AMPAR_{jl}} = v_j \times g_{AMPAR_{jl}}(t)$. The AMPAR model can be considered a two state model in which channels move from the off to the on state deterministically depending on an external signal.

NMDAR Channels

Similar to AMPAR channels, NMDAR channels exist almost exclusively on postsynaptic spines. These channels are dependent on the co-activation of AMPAR channels at the same synapse as well as voltage dependent release of channel blocking magnesium ions (Mg^{2+}). For simplicity, we model NMDAR channel activity using a two state model with voltage dependent levels of activity. The total conductance of the NMDAR channels on the l^{th} synapse in compartment j, depends on the pre-synaptic AMPAR activation. As in the AMPAR case this is considered an instantaneous event. The conductance is

$$g_{NMDAR_{jl}}(t) = \begin{cases} 0 & t < t_a \\ \bar{g}_{NMDAR} m(v_j)\left[e^{-\frac{t-t_a}{\tau_2^{NMDAR}}} - e^{-\frac{t-t_a}{\tau_1^{NMDAR}}}\right] & t \geq t_a \end{cases}, \tag{F.5}$$

where t_a is the activation (neurotransmitter release) time, $\tau_1^{NMDAR} = 2$ and $\tau_2^{NMDAR} = 20$ and $\bar{g}_{NMDAR} = 0.5$ is the maximum conductance through the entire population of channels on the spine. The value for silent synapses is $\bar{g}_{NMDAR,silent} = \bar{g}_{NMDAR}0.1$. The factor $m(v_j)$ represents the degree of Mg^{2+} mediated NMDAR blockage at the spine. This factor depends on

the voltage, v, as follows

$$m(v) = 1/\left(1 + \exp\left[\frac{-0.062v}{3.57}\right]\right).$$

The resulting contribution of NMDAR channels to the overall current is $i_{NMDAR_{jl}} = v_j \times g_{NMDAR_{jl}}(t)$

NMDAR channels are also non-specific channels that are permeable to Ca^{2+}. The approximate fractional contribution of Ca^{2+} to this current is $f = 1/11$ [189]. The Ca^{2+} specific current due to NMDAR activity at l is then given by $fg_{NMDAR_{jl}}(t)$ and $fg_{NMDAR_{jl}}(t)/10$ for l silent.

Ca^{2+} Specific Channels

There are two channels in the model (other than NMDAR channels) that permit a significant Ca^{2+} current, CA and CAT channels. The CA channel model is actually meant to account for a heterogeneous population of channels with similar behaviours.

Most channels are protein complexes composed of multiple subunits. These subunits each must be activated in order for a channel to open. In the case of CA channels, these are composed of a three protein complex with one homodimer and one other bound protein. This constitutes an 8 state model in which all three proteins must be active for current to flow. The activation probabilities for each protein are denoted by $m(v_j, t)$ (for the homodimer) and $h(v_j, t)$. The resulting mean-field conductance formula for a population of CA channels is

$$g_{CA_j}(v_j, t) = \bar{g}_{CA} m(v_j, t)^2 h(v_j, t)$$

where $\bar{g}_{CA}$ is the maximum conductance of all CA channels across compartment j and v_j is the membrane potential at compartment j as defined previously. Each protein activation probability (m or h) is described by a first order differential equation

$$\frac{dm(v_j, t)}{dt} = \frac{a_m(v_j)}{(a_m(v_j) + b_m(v_j))^2} - m(v_j, t)/(a_m(v_j) + b_m(v_j)) \tag{F.6}$$

$$\frac{dh(v_j, t)}{dt} = \frac{a_h(v_j)}{(a_h(v_j) + b_h(v_j))^2} - m(v_j, t)/(a_h(v_j) + b_h(v_j)) \tag{F.7}$$

where

$$a_m(v_j) = 0.055\frac{-27 - v_j}{\exp((-27 - v_j)/3.8) - 1} \qquad b_m(v_j) = 0.94\exp((-75 - v_j)/17)$$
$$a_h(v_j) = 0.000457\exp((-13 - v_j)/50) \qquad b_h(v_j) = \frac{0.0065}{(\exp((-v_j - 15)/28) + 1)}$$

with some rate adjustments for temperature changes between lab and in vivo conditions. These channels are most active for brief periods at high voltages. This stochastically, voltage gated channel model follows a general framework used across computational neuroscience, known as the Hodgekin-Huxley model framework.

Channels for Other Ions

There are eight other channel models for other ions and general non-specific currents. These follow similar formulations to those used above and allow for Na^+ spiking and K^+ to generate the spiking behaviour that is partially responsible for bAP activity as well as the firing of the axon (which lacks calcium channels all together). Na^+ spiking has also been shown to have a positive feedback effect on Ca^{2+} spiking in vivo. For further details, see [361].

F.4.4 Ca^{2+} Depletion Mechanism

There is a built in Ca^{2+} depletion mechanism that limits the Ca^{2+} current. It is given by

$$i_{ca-dep_j} = \left[-10^6\frac{i_{ca-other_j}}{F}\right]_+ C,$$

where $i_{ca-other_j}$ is the calcium current from other sources into j and F is Faraday's constant ($F = 1.602 \times 10^{-19}C$). The Ca^{2+} pumps, protein pores in the cell membrane similar to unidirectional active channels, are outward rectified since they are only able to pump Ca^{2+} out of the cell.

F.5 Kronecker Operators

The Kronecker sum, $\otimes$ is an operator on matrices. For two matrices A of dimension $a \times d$ and B of dimension $b \times c$, the Kronecker sum can be written in block matrix form. The constituent blocks are

$$[A \otimes B]_{i,j} = A_{i,j}B,$$

for blocks $(i, j) \in \mathbb{N}_a \times \mathbb{N}_b$. The resulting matrix has dimension $ab \times cd$. We are interested in the Kronecker sum, which is defined for square matrices A and B ($a = d$ and $b = c$), which is given by

$$A \oplus B = A \otimes I_b + B \otimes I_a$$

where I_a and I_b are the identity matrices of order a and b respectively.