UNIVERSITY OF CALIFORNIA
RIVERSIDE

Reverse Engineering Synchronization of Brain Network Dynamics:
Controllability Properties and Functional Patterns

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Mechanical Engineering

by

Tommaso Menara

December 2021

Dissertation Committee:

Dr. Fabio Pasqualetti, Chairperson
Dr. Konstantinos Karydis
Dr. Erfan Nozari
Dr. Danielle S. Bassett

The Dissertation of Tommaso Menara is approved:

_____

_____

_____

<div align="right">Committee Chairperson</div>

University of California, Riverside

Acknowledgments

I would like to express my deepest gratitude to my advisor, Dr. Pasqualetti, whose guidance, patience, and assistance have led me where I am today. During the time we have worked together, our excellent connection yielded many professional achievements that I never thought I could aspire to. More importantly, he taught me to recognize and pursue the questions that *matter*.

This thesis would not have been possible without the supervision of Dr. Bassett, whose boundless neuroscience expertise shed light on the the path whenever I got lost in the woods, and who also kindly allowed me to collaborate with her lab members on exciting projects.

I am extremely grateful to all my friends who accompanied me during this long journey. Thank you Siddharth, Gianluca, Rajasekhar, Federico, Giacomo, Maïko, Lorenzo, Yin-Chen, Vaibhav, Yuzhen (and many more) for sharing memorable times with me.

A special mention goes to Alessandro ⋆. Your exceptional friendship and support have been the foundations of my career since our very first (*awkward*) engineering test.

I also acknowledge the invaluable help, advocacy, and feedback that affectionate mentors such as Dr. Pequito and Dr. Cortese have provided me. I owe you.

To my wife LJ. This achievement is yours as much as it is mine.

I also dedicate this thesis to my parents. Despite not fully comprehending what I

have been working on for the past 5 years, they have always and unconditionally

supported me.

Same holds for all my friends in Italy, who inexplicably think I still build robots.

I made it, guys!

⋆ *waves* ⋆

ABSTRACT OF THE DISSERTATION

Reverse Engineering Synchronization of Brain Network Dynamics:
Controllability Properties and Functional Patterns

by

Tommaso Menara

Doctor of Philosophy, Graduate Program in Mechanical Engineering
University of California, Riverside, December 2021
Dr. Fabio Pasqualetti, Chairperson

Reverse engineering the brain holds the promise to overhaul the quality of life of human beings and vastly benefit mankind. Further advances towards this goal will lead to the reversal of cognitive decline, the creation of pioneering neural prostheses, the establishments of novel treatments for neurological disorders, and the development of human augmentation methods. This dissertation presents a cross-disciplinary approach to the study of the structure-function relationship in neural systems. Specifically, we study the brain as a dynamical system that obeys network-wide principles, and address three foundational challenges by using mathematically grounded methods that lie at the intersection of control theory, network science, and neuroscience.

First, through the application of control and graph-theoretic paradigms, we investigate how the spatial organization of anatomical brain network components governs and constrains its complex dynamical behaviors. The first chapters are dedicated to the study of structural brain networks – that is, empirically reconstructed large-scale networks that describe the interconnection scheme between different brain regions. We rigorously reveal that brain state transitions can be controlled by a single region, and that structural brain networks possess distinct controllability profiles with respect to random networks of the same size.

Second, we address the modeling and analysis of neural activity synchronization across different brain areas – which can be described by functional brain networks. To do so, we juxtapose

a bottom-up approach and a top-down approach. In the former, we utilize data-driven dynamical models to reveal that the synchronization of brain network dynamics is resilient to data heterogeneity, thus supporting the utilization of large heterogeneous repositories of brain recordings. In the latter, we abstract rhythmic activity of a neural system as the output of a network of diffusively coupled oscillators, and derive prescriptive conditions for the emergence of cluster synchronization. Such a phenomenon emerges when different groups of synchronized components coexist in a network, and regulates the functional interactions among network components.

Third and final, we build upon our previous findings and take aim at the tantalizing idea of controlling the synchronization of brain dynamics through minimally invasive local interventions. We derive a method to optimally intervene on the structural network parameters to achieve desired cluster-synchronized trajectories and, thus, prescribed functional interactions. Additionally, we show that our synchronization-based framework is robust to mismatches in network parameters, and validate it using a realistic neurovascular model to simulate neural activity and functional connectivity in the human brain.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"We have always two universes

of discourse [...], one dealing

with questions of quantitative

and formal structure,

the other with those qualities

that constitute a 'world'."

O. SACKS [1]

Reverse engineering the human brain is the quintessential challenge of modern science. This seemingly insurmountable task holds the promise to not only reveal what establishes the *persona*, the self, but also vastly benefit mankind. Importantly, while our understanding of the brain working principles may still be far from complete, the joint efforts of the scientific community towards this goal will lead to breakthrough clinical and technological achievements. Examples include the development of novel treatments for neurological disorders, the reversal of cognitive decline, and the engineering of human augmentation devices. Furthermore, such advances in brain science will

alleviate the global economic burden of neurological and psychiatric disorders, which is projected to reach $6T by 2030 [2].

A key aspect in neuroscience research is the characterization of the link between structure and function in the nervous system and the brain. The study of this complex structure-function relationship has sparked the interest of physicians and scientists since the early days of medicine – the earliest known documents reporting cases of brain injuries along with various treatment recommendations date back a staggering five thousand years [3]. Three thousand years later, Hippocrates of Kos, a forefather of neurology, boldly hypothesized that intelligence resides in the brain [4]. Fast forward to modern times, outstanding technological advances have enabled tremendous discoveries in neuroscience. An exquisite example – easily digestible to any interested reader – of how modern brain probing techniques have further enhanced our collective knowledge is the collection of clinical stories by the neurologist O. Sacks [1]. Yet, despite recent formidable advances, the investigation into the brain structure-function relationship is by no means concluded. The main thrust of this dissertation is to shed light onto this critical topic through a cross-disciplinary approach.

Here, we leverage the emerging intersection between control theory, network science, and neuroscience, to investigate how the spatial organization of neural circuits at different scales (the *structural domain*) supports and governs the large-scale complex coordinated behavior of neural activity (the *functional domain*) that ultimately maps into cognitive processes (see Fig.1.1). The main objective of this thesis is to present a series of results wherein, by applying rigorous engineering methods, we (i) elucidate the structural brain properties that allow for endogenous and exogenous control of neural activity, (ii) unveil the mechanisms underpinning large-scale neural synchronization, and (iii) gaze upon the etiology of cognition as a result of dynamical process that evolve at multiple spatio-temporal scales.
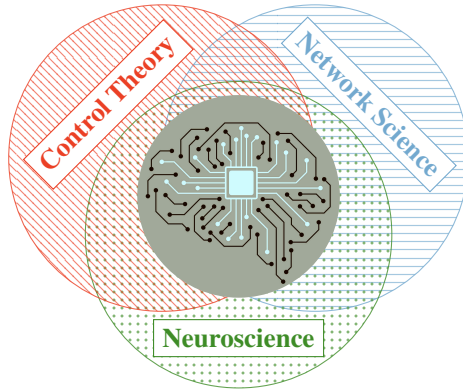
Figure 1.1: The synergistic interaction of control theory, network science and neuroscience promotes the discovery of the brain working principles and the design of systems-level interventions to correct abnormal brain dynamics.

1.1 Control and Graph-Theoretic Paradigms Meet Neuroscience Challenges

At the core of this thesis is the study of the brain as a dynamical system whose relevant dynamics are governed by network-wide rules. In this setup, the communication between single neurons or, at a coarser scale, between neuronal populations is constrained by axons or axonal bundles, and is hierarchically regulated across many scales of spatial and temporal organization. That is, the interconnection scheme between neural components constrains the evolution of brain dynamics, and the integrative nature of brain function can be addressed from a complex network perspective. Such a framework lends itself to be investigated through control-theoretic tools. In fact, the intersection between control theory and the study of the brain as a network system has the potential to overhaul our understanding of the brain dynamical wandering between different states and motivates the first scientific questions addressed in this thesis.

What is the role of the brain anatomical (structural) interconnection scheme in enabling cognitive control? How do external stimuli propagate through the web of neural interconnections to steer the state of the system? The capability of driving the state of a network system through endogenous or exogenous stimuli – controllability – is inherently determined by the network organi-

zation. By employing control-theoretic tools, we provide prescriptive conditions for structural brain networks to be controllable and demonstrate that such brain networks possess distinct controllability profiles. Notably, control theory offers two primary advantages over traditional approaches to the analysis of brain network function. First, the theoretical framework enables the study of the intimate relationship between structure and function, in contrast to approaches that characterize each separately. Second, control theory often goes beyond the purely descriptive approach of network science by defining generative models that are parameterized by both a network's spatial and temporal features. These methods and tools hold the promise to inform the development of novel treatments and stimulation schemes to treat multiple neurological disorders.

1.2 Reverse Engineering Synchronization of Brain Network Dynamics

Synchronization phenomena are ubiquitous in the brain. During a cognitively demanding task or at rest, the brain exhibits a rich repertoire of large-scale synchronization patterns supported by its static interconnection scheme. These patterns are a measure of the coherence among the neural activities in different brain areas, and enable a multitude of functions in the brain. Examples include motor coordination, sleep spindles, circadian rhythms, and large scale integration of sensori stimuli. Furthermore, synchronization can be used as a biomarker in multiple psychiatric and neurological disorders, and abnormal or decreased synchronized activity has been linked to neurological damage or cognitive decline due to aging. In this part, we leverage the view of the brain as a dynamical network system, and propose two complementary approaches to address the following scientific challenge.

What are the mechanisms underpinning different synchrony configurations of brain activity? We will present a bottom-up and a top-down explorations of the relationship between data, models, and paradigms of neural synchronization. In the bottom-up approach, we infer a data-driven dynamical model, where distinct synchronization patterns define different states. We use

such a model of large-scale trajectories to identify brain dynamics fingerprints and assess the quality of multi-site repositories of brain activity recordings. Next, in the top-down approach, we make use of oscillator networks and mathematical abstractions to reverse engineer the synchronization of oscillatory brain rhythms. By abstracting oscillatory neural activity as the result of interconnected nonlinear oscillators whose dynamics are constrained by the brain's structure, we reveal the mechanisms that enable a variety of synchronization phenomena. Specifically, we study the phenomenon of cluster synchronization in networks of oscillator – where distinct groups of synchronized oscillators coexist in a network.

1.3   Control of Synchronization in Brain Networks

The development of an overarching theory to bridge the intertwined nature of the neural activity synchronization with our view of the brain as a complex network system enables the development of principled methods to effectively control synchronization phenomena. By developing the aforementioned framework, where neural activity evolution can be approximated by the dynamics of interconnected nonlinear oscillators, the following question arises naturally.

Which control parameters can we leverage to govern the level of synchrony between different brain regions? We explore the tantalizing idea of controlling brain-wide functional relations – defined as *functional connectivity* – by selectively acting on the brain's structure and parameters (see Fig. 1.2). Functional connectivity, which measures the degree of correlation between neural activities in different brain regions, can be used to distinguish between healthy and certain diseased brain dynamics and, possibly, as a control parameter to restore healthy functions. In this part of the thesis, we show that functional connectivity is essentially regulated by the degree of synchronization between different clusters of oscillators. Then, we propose a minimally invasive method to correct the oscillators' interconnections and frequencies to enforce arbitrary and stable synchronization patterns among the oscillators and, consequently, a desired pattern of functional connectivity.
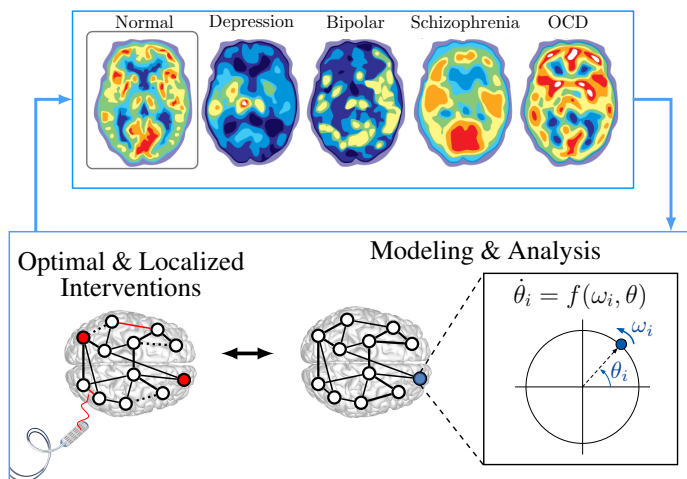
Figure 1.2: Networks of nonlinear oscillators can be used to inform the design of optimal and localized interventions.

1.4 Organization and Summary of Contributions

The contents of this dissertation are organized into nine main chapters, followed by a shared conclusion. The main contributions of each chapter are listed below.

Chapter 2 – Preliminaries in Graph Theory, Control Theory, and Brain Networks. In this chapter, we introduce some preliminary notions and results that will be used throughout the thesis. Specifically, we present graph theory concepts used to describe network systems, we define controllability of network systems, and we describe how the brain and its activity can be studied as complex networks.

Chapter 3 – Structural Controllability of Empirically Reconstructed Brain Networks. The contribution of this chapter is twofold. First, we extend the theory of structural controllability. Second, we apply such a theory to prove that empirically-reconstructed structural brain networks are controllable from any node. The theory of structural controllability allows us to assess controllability of a network as a function of its interconnection graph and independently of the edge weights. Yet, existing structural controllability results require the weights to be selected arbitrarily and independently from one another and provide no guarantees when these conditions are not

satisfied. Here, we develop a new theory for structural controllability of networks with symmetric, thus constrained, weights. In this chapter, we show that network controllability remains a generic property even when the weights are symmetric, and we demonstrate that a symmetric network is structurally controllable if and only if it is structurally controllable without weight constraints. Our results enables to assess the controllability of symmetric structural brain networks reconstructed via diffusion imaging methods.

Chapter 4 – The Structured Controllability Radius of Symmetric (Brain) Networks. In this chapter, we further elaborate the study of controllability of symmetric brain networks and propose and analyze a novel notion of controllability of network systems with linear dynamics and symmetric weights. Namely, we quantify the controllability degree of a network with its distance from the set of uncontrollable networks with the same structure, that is, with the minimum Frobenius norm of a structured perturbation rendering the network uncontrollable (structured controllability radius). We derive analytical conditions to compute the structured controllability radius of a network with symmetric weights, and illustrate our results through a number of examples. In particular, we use our theoretical results to study the controllability properties of a set of brain networks reconstructed from diffusion MRI data, and compare them with the controllability properties of a class of random networks. Our results show that brain networks feature a controllability radius that is consistently smaller than the one of random networks with similar weights, indicating that the considered brain networks may not be optimized to favor controllability.

Chapter 5 – Brain Network Dynamics Fingerprints are Resilient to Data Heterogeneity In this chapter, we investigate how large-scale synchronization of neural activity defines reliable brain dynamics fingerprints. Our objective is to validate the estimation of individual brain network dynamics fingerprints and to appraise sources of variability in large resting-state functional magnetic resonance imaging datasets by providing a novel point of view based on data-driven dynamical models. We utilize hidden Markov models to examine how diverse scanning factors in multi-site fMRI recordings affect our ability to infer the brain's spatiotemporal wandering between large-scale

networks of activity. Specifically, we leverage a stable hidden Markov model trained on the Human Connectome Project (homogeneous) dataset, which we then apply to an heterogeneous dataset of traveling subjects scanned under a multitude of conditions. Building upon this premise, we first replicate previous work on the emergence of non-random sequences of brain states. We next highlight how these time-varying brain activity patterns are robust subject-specific fingerprints. Finally, we suggest these fingerprints may be used to assess which scanning factors induce high variability in the data. These results demonstrate that we can use large scale dataset to train models that can be then used to interrogate subject-specific data, recover the unique trajectories of brain activity changes in each individual, but also urge caution as our ability to infer such patterns is affected by how, where and when we do so.

Chapter 6 – Cluster Synchronization in Networks of Phase Oscillators. Building upon the premises of the previous chapter, we model neural activity as the output of a network of interconnected oscillators. In this framework, each oscillator represents a brain region, and the anatomical structure of the brain defines the network over which such oscillators interact. During rest or cognitive tasks, neural activity moves between specific brain-wide synchronization patterns. To enable a mathematical investigation of the mechanisms underlying this behavior, we focus on cluster synchronization of oscillators, which emerges when the oscillators can be partitioned in a way that their phases remain identical over time within each group. This phenomenon is critically important for normal and abnormal behaviors in technological and biological systems ranging from the power grid to the human brain. Yet, despite its importance, cluster synchronization has received limited attention, so that the fundamental mechanisms regulating cluster synchronization in important classes of oscillatory networks are still unknown. In this chapter, we provide the first conditions for the stability of the cluster synchronization manifold for general weighted networks of heterogeneous oscillators with Kuramoto dynamics. In particular, we discuss how existing results are inapplicable or insufficient to characterize the stability of cluster synchronization for oscillators with Kuramoto dynamics, provide rigorous quantitative conditions that reveal how the network

weights and oscillators' natural frequencies regulate cluster synchronization, and offer examples to quantify the tightness of our conditions. Further, we develop approximate conditions that, despite their heuristic nature, are numerically shown to tightly capture the transition to stability of the cluster synchronization manifold.

Chapter 7 – Relay Interactions Enable Remote Synchronization of Phase Oscillators. This chapter concludes the analysis of synchronization phenomena by focusing on remote synchronization. This fascinating phenomenon occurs when oscillators that are not directly connected via physical links evolve synchronously. Remore synchronization is thought to be critical for distributed information processing in the mammalian brain, where long-range synchronization is empirically observed between neural populations belonging to spatially distant brain regions. Inspired by the growing belief that this phenomenon may be prompted by intermediate mediating brain regions, such as the thalamus, in this chapter we derive a novel mechanism to achieve remote synchronization. This mechanism prescribes remotely synchronized oscillators to be stably connected to a cohesive relay in the network – a group of tightly connected oscillators mediating the distant ones. Remote synchronization unfolds whenever the stability of the subnetwork formed by relays and remotely synchronized oscillators is not affected by the rest of the oscillators. In accordance with our results, we find that remotely-synchronized cortico-thalamo-cortical circuits in the brain posses strong interconnection profiles. Finally, we demonstrate that the absence of cohesive relays prevents stable remote synchronization in a large class of cases, further validating our results.

Chapter 8 – Feedback Linearization of Nonlinear Network Systems. This chapter introduces the section on the control of neural synchronization. We utilize the framework of interconnected oscillators introduced in the previous chapters and provide novel conditions to test whether a nonlinear network is feedback-linearizable. Feedback linearization allows for the local transformation of a nonlinear system to an equivalent linear one by means of a coordinate transformation and a feedback law. However, feedback linearization of large-scale non-linear network systems is typically difficult, since existing conditions become harder to check as the network size

becomes larger. In this chapter, given some dedicated control inputs injected to a set of network nodes, we derive an easy-to-check algebraic condition that can be tested on the Jacobian matrix of the network dynamics evaluated at some desired working point. Furthermore, our requirements are sufficient for (local) controllability, and thus provide a testable condition for controllability of large-scale nonlinear networks. Finally, we validate our findings by enforcing the formation of desired synchronization patterns in networks of coupled oscillators.

Chapter 9 – A Framework to Control Functional Connectivity. In this chapter, we propose a framework to control brain-wide functional connectivity by selectively acting on the brain's structure and parameters. Functional connectivity, which measures the degree of correlation between neural activities in different brain regions, can be used to distinguish between healthy and certain diseased brain dynamics and, possibly, as a control parameter to restore healthy functions. Here, we use a collection of interconnected Kuramoto oscillators to model oscillatory neural activity, and show that functional connectivity is essentially regulated by the degree of synchronization between different clusters of oscillators. Then, we propose a minimally invasive method to correct the oscillators' interconnections and frequencies to enforce arbitrary and stable synchronization patterns among the oscillators and, consequently, a desired pattern of functional connectivity. Additionally, we show that our synchronization-based framework is robust to parameter mismatches and numerical inaccuracies, and validate it using a realistic neurovascular model to simulate neural activity and functional connectivity in the human brain.

Chapter 10 – Conclusions. This chapter concludes the dissertation. We summarize the main contributions of this manuscript, and present directions for future research.

Chapter 2

Preliminaries in Graph Theory, Control Theory, and Brain Networks

> *"Far from being able to accept the idea of the individuality and independence*
>
> *of each nerve element, I have never had reason, up to now,*
>
> *to give up the concept which I have always stressed, that nerve cells,*
>
> *instead of working individually, act together [. . .].*
>
> *However opposed it may seem to the popular tendency to individualize the elements,*
>
> *I cannot abandon the idea of a unitary action of the nervous system [. . .]"*
>
> C. GOLGI [5]

What Camillo Golgi hypothesized in the late 1800s is now commonly accepted by numerous research communities: brain functions are enabled and supported by the concurrent activation of ensembles of neurons in distinct cortical and subcortical regions. To allow for a rigorous study of such phenomenon, this chapter presents a primer on concepts from Graph Theory, Control Theory, and Network Neuroscience that will be used throughout this dissertation.

2.1 Preliminaries on Graph Theory

Networks are used to describe many natural and engineering systems where the spatial or temporal interactions between components are governed by distinct architectural features [6]. Examples include power distribution networks, online interactions between Facebook friends, complex ecological systems, communicating satellites in orbit, swarms of drones, and the brain. The interaction scheme defining a network can be described by a graph. We let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ denote a **graph**, where $\mathcal{V} = \{1, \ldots, n\}$ is the vertex (or node) set, and $\mathcal{E} = \{(i,j) \; : \; i \text{ and } j \text{ are connected}\} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge (or link) set. The connections between the nodes of a graph can be weighted or unweighted. In both cases, a graph of cardinality n (i.e., comprising n nodes) can be conveniently described by a square adjacency matrix $A \in \mathbb{R}^{n \times n}$:

$$A = [a_{ij}] = \begin{cases} a_{ij} \in \mathbb{R} & \text{if } (i,j) \in \mathcal{E}, \\ 0 & \text{otherwise.} \end{cases}$$

Loosely speaking, an adjacency matrix encodes the network's sparsity (whether $a_{ij} \neq 0$) and the interconnection strength between its components (the scalar value of a_{ij}). In the case of unweighted networks, either $a_{ij} = 1$ or $a_{ij} = 0$. Unless specified differently, in this dissertation we use weighted graphs, as we are interested in characterizing how the interconnection scheme *and* the interconnection strength between a (brain) network components constrain its dynamics.

2.2 Control Theory Notions

We study the brain as a dynamical system whose components interact based on a network structure. Ultimately, our goal is the application of engineering techniques and principles to the analysis and control of neural systems. To formalize these concepts, we make use of control theory notions. Control theory offers two primary advantages over traditional approaches to the analysis of brain network function. First, the theoretical framework enables the study of the intimate relationship between structure and function, in contrast to approaches that characterize each sepa-

rately. Second, control theory often goes beyond the purely descriptive approach of network science by defining generative models that are parameterized by both a network's spatial and temporal features.

Before framing research problems in the context of networks representing brain structure and function, we introduce the key concepts of dynamical systems and their control properties. A continuous-time time-invariant nonlinear system is described by the equation

$$\dot{x}(t) = f(x(t), u(t)) \tag{2.1}$$

where $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$ denote the state vector and the input vector, respectively, $f(\cdot)$ is a vector field that describes the dynamics of the system. Engineers are usually interested in controlling the system (2.1) around a working point – an equilibrium point. We define an equilibrium point of the system (2.1) as any point $\bar{x} \in \mathbb{R}^n$ that satisfies $f(\bar{x}) = 0$.

Since nonlinear dynamics are typically challenging to handle, we can obtain a good approximation of the nonlinear system (2.1) in a small neighborhood of $\bar{x}$ through linearization of the system around the equilibrium point. By taking the derivatives $\partial f / \partial x = A$ and $\partial f / \partial u = B$, we obtain a continuous-time linear time-invariant system, which is defined by the equation

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{2.2}$$

where $A \in \mathbb{R}^{n \times n}$ the dynamics matrix, and $B \in \mathbb{R}^{n \times m}$ the input matrix.

We are interested in the control of the dynamical system (2.2). That is, we want to answer the question: *Does there exists a control input $u(t)$ that can steer the state from an initial state $x(0) = x_0$ to a final desired state $x(T) = x_f$ in a finite time T?*. To answer this question, we let the controllability matrix of (2.2) be

$$\mathcal{C}(A, B) = [B \ AB \ A^2B \ \cdots \ A^{n-1}B]$$

and recall that the network (2.2) is controllable if and only if its controllability matrix $\mathcal{C}(A, B)$ is invertible [7]. Fig. 2.1 depicts a 3-dimensional system's trajectory going from an initial state $x_0 = 0$
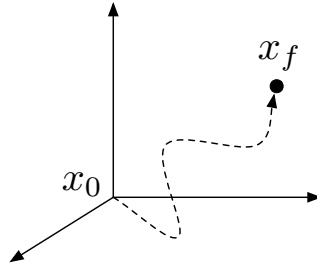
Figure 2.1: A controllable system's state can be steered from its initial state x_0 to a desired final state x_f in finite time.
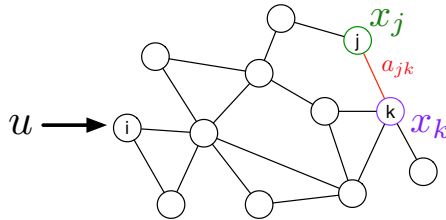
Figure 2.2: A control input u is injected into the network at node i. The network weight between nodes j and k, denoted by a_{jk}, determines how the states x_j and x_k influence each other.

to a final state x_f.

Dynamical Network Systems. Utilizing an adjacency matrix in (2.2) yields a network systems that obeys linear dynamics – a linear network system. Fig. 2.2 illustrates a network where a scalar control input is injected at node i through the input matrix $B = e_i$, where e_i is the i-th canonical vector.

Clearly, the network structure encoded by the adjacency matrix A can also underlie a nonlinear network system, where the interconnection scheme is described by A, but the interactions between the network components are nonlinear. We explore these types of networks in the second part of this dissertation.
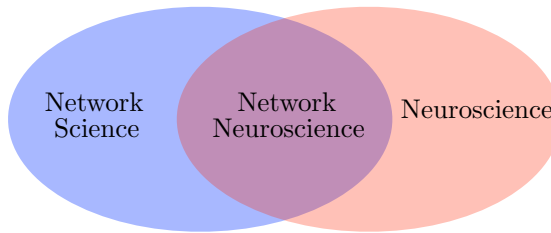
Figure 2.3: The field of network neuroscience lies at the intersection of network theory (from which many tools and metric to study graphs are borrowed) and neuroscience.

2.3   Brain Networks

In this dissertation, we focus on the modeling, analysis and control of brain networks. Brain networks are network representation of the brain's structure or function [8]. Brain networks span multiple spatial scales, from the microscale of individual cells and synapses to the macroscale of cognitive systems and embodied organisms. Thanks to enormous advances in the methods and technologies used to record and trace the brain anatomy and its signals, the field of network neuroscience (see Fig. 2.3) has matured and ultimately exploded in the last decade. This field defined two main types of brain networks: structural brain networks and functional brain networks.

Structural brain networks. These networks represent the anatomical organization of the brain – a set of physical or structural (anatomical) connections linking neural elements. The spatial scale of structural brain networks range from axonal interconnections between single neurons to large fiber bundles between brain regions. Depending on how such connections are recorded or traced, the estimated network can have weighted or unweighted links.

Functional brain networks. These networks represent functional relations (such as correlation or statistical deviation from independence) between distributed neural units. Data may be extracted from cellular recordings, EEG, MEG, fMRI, or other techniques. Notice that the interconnection values in this type of network are inherently time-dependent. In many cases, functional brain networks change on a scale of hundreds of milliseconds, and are modulated by external task demands and sensory stimulation, as well as the internal state of the organism.
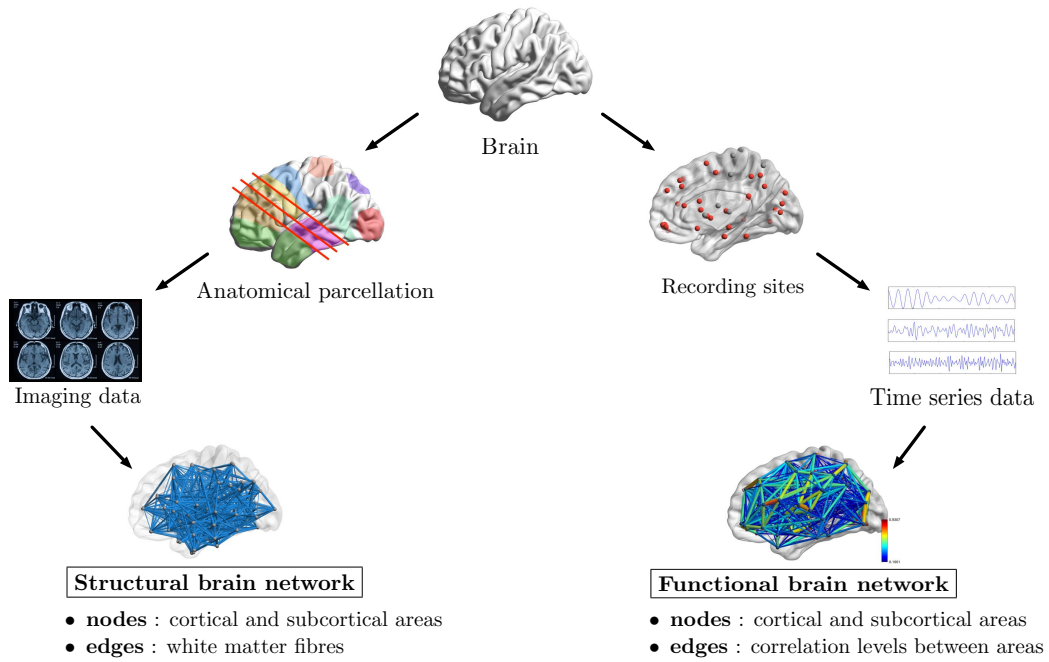
Figure 2.4: The left side of this figure shows a pipeline to construct a structural brain network, where nodes represent neural units such as brain regions and edges represent the estimated thickness of the fiber bundles connecting such regions. The right side of this figure illustrates a pipeline to obtain a functional brain network from brain activity recordings. The nodes represent the same neural units as in the structural network, but the edges are proportional to the correlation between the neural activity recorded at the nodes.

Fig. 2.4 illustrates the construction of structural and functional brain networks.

Part I

Application of Control and Graph-Theoretic Methods To Structural Brain Networks

Chapter 3

Structural Controllability of Empirically Reconstructed Brain Networks

This chapter presents seminal results on the controllability from any region of empirically reconstructed brain networks. We demonstrate that structural controllability depends only on the network structure and can be assessed reliably and efficiently, even when the network weights are constrained to be symmetric – as in the case of many empirically reconstructed brain networks. As an additional contribution, we settle a disagreement between the authors in [9] and the ones in [10], proving that the conclusions in [10] are numerically accurate and theoretically correct. Briefly, the topics of debate relate to the question of whether a linear dynamical network is controllable from a single node, where the network structure is akin to those reconstructed from tract tracing data in non-human animals, or from diffusion tensor, diffusion spectrum, and other diffusion weighted imaging scans in humans. By developing a mathematically correct argument that structural controllability of network systems from one node depends only on the network structure, we demonstrate that brain

networks numerically controllable for any generic choice of network weights. We refer the reader to [9, 11–13] for a comprehensive discussion on this topic.

3.1 Introduction

The question of controllability of complex network systems arising in engineering, social, and biological domains has been the subject of intensive study in the last few years [14–16]. One key question motivating the investigation is to characterize relationships and tradeoffs between the interconnection structure of a network and its controllability [17–19]. To this end, graphical tools from structural systems theory [20–23] are typically preferred over algebraic controllability tests, which suffer from numerical instabilities when the network cardinality grows, require exact knowledge of the network weights, and are agnostic to the graph supporting the dynamics.

While the theory of structured systems and generic properties of linear systems is well-developed and understood [24], all results assume the possibility of assigning the network weights arbitrarily and independently from one another. In fact, when this condition is violated, the conclusions drawn from structural analysis may lead to incorrect results [20, 25]. Unfortunately, it is often the case that this assumption is violated in real networks due to physical, technological, or biological reasons. For instance, the small-signal network-preserving model of a power network contains a Laplacian submatrix, whose entries are symmetric and satisfy linear constraints (row sums equal to zero) [26, 27]. Similar constraints appear also when studying synchronization in networks of Kuramoto oscillators [28] and general systems with consensus dynamics [29]. Novel theories and tools are needed to study controllability of networks with constrained weights.

In this chapter, we focus on networks with symmetric weights and derive graph-theoretic conditions for their structural controllability from dedicated control inputs. While (group) symmetry has previously been found to be responsible for network uncontrollability [30, 31], the question of how symmetric edge weights affect structural controllability has not been investigated, with the

exception of [32]. In [32], however, the proposed conditions for structural controllability of undirected (symmetric) networks are implicit and based on the generalized zero forcing sets to estimate the dimension of the controllable subspace. Similarly, although the recent paper [33] studies structural controllability for a class of networks with constrained parameters, this class of network matrices does not contain the set of symmetric matrices considered in this chapter. Thus, the necessary and sufficient conditions derived here are the first graph-theoretic conditions for structural controllability of networks with symmetric weights.

The contribution of this chapter is two-fold. First, we show that a network with symmetric weights is structurally controllable if and only if it is spanned by a (symmetric) cactus rooted at the control node. By comparing our result with those in [34], our analysis shows that a network is structurally controllable with symmetric weights if and only if it is structurally controllable with unconstrained weights. Second, we use our results to show that a class of (symmetric) brain networks reconstructed from diffusion MRI data is structurally controllable from a single dedicated control region. Finally, we note that, due to duality between controllability and observability, the results of this chapter extend directly to the study of structural observability of networks with symmetric weights and a dedicated sensor.

The rest of the chapter is organized as follows. Section 3.2 contains our network model and preliminary notions. Section 3.3 contains our analysis and conditions for structural controllability of networks with symmetric weights, and some examples. Finally, Section 3.4 contains an illustrative example featuring brain networks, and Section 9.6 concludes the chapter.

3.2 Problem setup and preliminary notions

We study controllability of symmetric network systems, which are described by a weighted directed graph (digraph) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \dots, n\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ are the vertex and edge sets, respectively, and a symmetric weighted adjacency matrix $A = [a_{ij}]$ with $a_{ij} = 0$ if $(i, j) \notin \mathcal{E}$ and $a_{ij} \in \mathbb{R}$ if $(i, j) \in \mathcal{E}$. Let $x \in \mathbb{R}^n$ be the vector containing the state of the network nodes over time, and let $i \in \mathcal{V}$ be the control node. We let x evolve according to linear time-invariant dynamics:

$$\delta(x) = Ax + b^{\mathrm{i}} u, \tag{3.1}$$

where $\delta(x)$ denotes the time derivative (resp. time shift) operator for continuous-time (resp. discrete-time) dynamics, and $b^{\mathrm{i}} = e_i$, with e_i the i-th canonical vector of dimension n. Finally, let the controllability matrix of (4.1) be

$$\mathcal{C}(A, b^{\mathrm{i}}) = \begin{bmatrix} b^{\mathrm{i}} & Ab^{\mathrm{i}} & \cdots & A^{n-1}b^{\mathrm{i}} \end{bmatrix}, \tag{3.2}$$

and recall that the network (4.1) is controllable if and only if its controllability matrix $\mathcal{C}(A, b^{\mathrm{i}})$ is invertible [7].

Assessing controllability of network systems is numerically difficult because the controllability matrix typically becomes ill-conditioned as the network cardinality increases; e.g., see [17,35]. Because different controllability tests suffer similar numerical difficulties, a convenient tool to study controllability of networks is to resort to the theory of structural systems. To formalize this discussion, let $a_{\mathcal{E}} = \{a_{ij} \,:\, (i, j) \in \mathcal{E}\}_{\mathrm{ordered}}$ denote the set of nonzero entries of A in lexicographic order, and notice that the determinant $\det(\mathcal{C}(A, b^{\mathrm{i}})) = \phi(a_{\mathcal{E}})$ is a polynomial function with variables $a_{\mathcal{E}}$. From the above reasoning, the network (4.1) is uncontrollable when the weights $a_{\mathcal{E}}$ are chosen so that $\phi(a_{\mathcal{E}}) = 0$. Let $\mathcal{S}$ contain the choices of weights that render the network (4.1) uncontrollable; that is,

$$\mathcal{S} = \{z \in \mathbb{R}^d \,:\, \phi(z_1, \dots, z_d) = 0\}, \tag{3.3}$$

where $d = |\mathcal{E}| = |a_\mathcal{E}|$. Notice that $\mathcal{S}$ describes an algebraic variety of $\mathbb{R}^d$ [36]. This implies that controllability of (4.1) is a *generic property*, as it fails to hold on an algebraic variety of the parameter space [36–38]. Consequently, when assessing controllability of the network (4.1) as a function of the weights, only two mutually exclusive cases are possible:

1. either there is *no* choice of weights a_{ij}, with $a_{ij} = 0$ if $(i,j) \notin \mathcal{E}$, rendering the network (4.1) controllable; or

2. the network (4.1) is controllable for all choices of weights a_{ij}, with $a_{ij} = 0$ if $(i,j) \notin \mathcal{E}$, except, possibly, those belonging to the proper algebraic variety $\mathcal{S} \subset \mathbb{R}^d$.[1]

Loosely speaking, if one can find a choice of weights such that the network (4.1) is controllable, then *almost all* choices of weights yield a controllable network. In this case, the network is said to be *structurally controllable* [20, 34, 39].

Classical results on structural controllability cannot be directly applied to networks where the weights are constrained [20, 25]. In fact, these results assume that the network weights can be selected arbitrarily and independently from one another, a condition that cannot be satisfied, for instance, when the weights need to be symmetric. In this note we overcome this limitation, and extend the results on structural controllability to symmetric networks. In particular we show that a network is structurally controllable with symmetric weights if and only if it is structurally controllable with unconstrained weights.

---

[1]The variety $\mathcal{S}$ of $\mathbb{R}^d$ is proper when $\mathcal{S} \neq \mathbb{R}^d$ [36].

3.3 Structural controllability of symmetric networks

In this section we derive necessary and sufficient graph-theoretic conditions for structural controllability of networks with symmetric weights. We proceed as follows. First, we show that network controllability remains a generic property when the weights are symmetric. Second, we provide conditions to construct controllable networks with symmetric weights. Finally, combining these results yields conditions for structural controllability of networks with symmetric weights.

Theorem 1 *(Symmetry and genericity)* *Controllability of the network* (4.1) *with symmetric matrix A is a generic property.*

Proof. Let $d = |\mathcal{E}|$ and $d_s = |\{i \ : \ (i,i) \in \mathcal{E}\}|$. Notice that a network with symmetric weights is uniquely specified by $(d + d_s)/2$ parameters, for instance, by the set $a'_{\mathcal{E}} = \{a_{ij} \ : \ (i,j) \in \mathcal{E}, \ i \leq j\}_{\text{ordered}}$ in lexicographic order. Further, because of the symmetry constraint, the determinant of the controllability matrix of (4.1) is a polynomial function $\phi'(a'_{\mathcal{E}})$, which can be obtained, for instance, from the determinant $\det(\mathcal{C}(A, b^{\text{i}}))$ by substituting a_{ij} with a_{ji} whenever $i > j$. Thus, even for symmetric networks, the determinant of the controllability matrix is a polynomial function of the network weights, and the weights that render the network uncontrollable define the algebraic variety $\mathcal{P} = \{z \in \mathbb{R}^{(d+d_s)/2} \ : \ \phi'(z_1, \ldots, z_{(d+d_s)/2}) = 0\}$. To conclude, either $\mathcal{P} = \mathbb{R}^{(d+d_s)/2}$, and the network is uncontrollable for all choices of symmetric weights, or $\mathcal{P}$ is a proper algebraic variety of $\mathbb{R}^{(d+d_s)/2}$, and the network is controllable for all choices of symmetric weights except, if any, those belonging to the set $\mathcal{P}$ of zero Lebesgue measure [36]. ∎

Theorem 1 shows that controllability remains a generic property even when the weights are constrained to be symmetric. This result will be key in the derivation of our conditions for structural controllability of networks with symmetric weights. In fact, because controllability remains a generic property, it will be sufficient to show that a network is controllable for a specific choice of symmetric weights to guarantee that controllability holds for almost all choices of weights. In the next example we illustrate that the set of symmetric weights preventing controllability forms an algebraic variety.
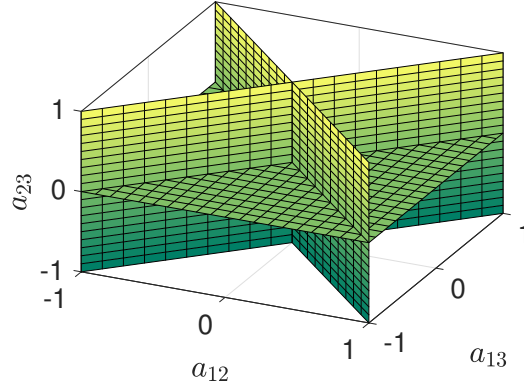
Figure 3.1: Algebraic variety defined by $a_{23}a_{12}^2 - a_{23}a_{13}^2 = 0$, which determines the weights for which the network in Example 2 is not controllable. The network is controllable for all weights outside of this algebraic variety.

Example 2 *(Structural controllability with symmetric weights) Consider a network with symmetric adjacency matrix*

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{12} & 0 & a_{23} \\ a_{13} & a_{23} & 0 \end{bmatrix}, \tag{3.4}$$

and input vector $b^1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{\mathsf{T}}$. From (3.2), the controllability matrix of the pair (A, b^1) is

$$\mathcal{C}(A, b^1) = \begin{bmatrix} 1 & 0 & a_{12}^2 + a_{13}^2 \\ 0 & a_{12} & a_{13}a_{23} \\ 0 & a_{13} & a_{12}a_{23} \end{bmatrix}, \tag{3.5}$$

with determinant $\det(\mathcal{C}(A, b^1)) = a_{23}a_{12}^2 - a_{23}a_{13}^2$. Thus, the network is controllable (i.e., $\det(\mathcal{C}(A, b^1)) \neq 0$) for all symmetric choices of weights a_{12}, a_{13}, and a_{23}, except those lying on the proper algebraic variety shown in Fig. 3.1 and defined by the equation $a_{23}a_{12}^2 - a_{23}a_{13}^2 = 0$. □

Remark 3 *(Structural controllability of consensus systems) A multi-agent consensus network with leader nodes is described by a linear time-invariant dynamical system, where the nonzero entries of A have a specified sign and the sums along the rows of A equal to a constant (1 for*

24

discrete-time networks, and 0 *in the case of continuous-time networks) [40, 41]. Theorem 1 can be*

easily extended to include constraints on the sign of the entries of A *and on their sums. In fact, if*

$\sum_{j=1}^{n} a_{ij} = c$, *for some constant* $c \in \mathbb{R}$, *then*[2] $a_{i1} = c - \sum_{j=2}^{n} a_{ij}$ *can be substituted in the polyno-*

mial $\det(\mathcal{C}(A, b^{i}))$, *showing that the set of parameters preventing controllability forms an algebraic*

variety of the free parameter space, and that controllability remains a generic property despite the

constraints. Similarly, when some entries have a specified sign or need to assume identical values,

the set of parameters preventing controllability can be shown to be a subset of an algebraic variety,

which either equals the set of feasible parameters, or remains of zero Lebesgue measure. $\square$

Example 4 *(Structural controllability of consensus systems) Consider a linear discrete-time*

consensus system with node 1 *as a leader and adjacency matrix*

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{bmatrix}.$$

Because the rows of A *need to sum to* 1, *it is possible to rewrite* 3 *parameters as a function of the*

others. For instance, rewrite $a_{13} = 1 - a_{12}$, $a_{23} = 1 - a_{21}$, *and* $a_{32} = 1 - a_{31}$. *By doing so, the*

determinant $\det(\mathcal{C}(A, b^{1})) = -a_{21}^{2}a_{31} + a_{21}^{2} + a_{21}a_{31}^{2} - a_{31}^{2}$, *and the set of weights that make such*

determinant vanish defines a proper algebraic variety of $\mathbb{R}^{3}$. $\square$

We next introduce some graph-theoretic notions [20, 42]. Given a digraph, a path is an ordered sequence of nodes such that any pair of consecutive nodes in the sequence is a directed edge of the digraph. A digraph is strongly connected if there exists a directed path from any node to any other node. Furthermore, given the digraphs $\mathcal{G}_1, \ldots, \mathcal{G}_m$, let $\mathcal{G} = \overline{\bigcup}_{i=1}^{m} \mathcal{G}_i$ be the connected digraph $(\mathcal{V}, \mathcal{E})$ defined as follows: $\mathcal{V} = \bigcup_{i=1}^{m} \mathcal{V}_i$ and $\mathcal{E} = \bigcup_{i=1}^{m} \mathcal{E}_i \cup \bar{\mathcal{E}}$, where $|\bar{\mathcal{E}}| = 2(m-1)$ and, for all $i \in \{2, \ldots, m\}$, there is a unique pair of edges $(p_i, q_i) \in \bar{\mathcal{E}}$ and $(q_i, p_i) \in \bar{\mathcal{E}}$ with $p_i \in \mathcal{V}_i$ and $q_i \in \bigcup_{j=1}^{i-1} \mathcal{V}_j$. Finally, we present some definitions that are inspired by [34] and will be used to derive our structural controllability conditions for networks with symmetric weights.

---

[2]If $a_{i1} = 0$, then select a different nonzero entry.
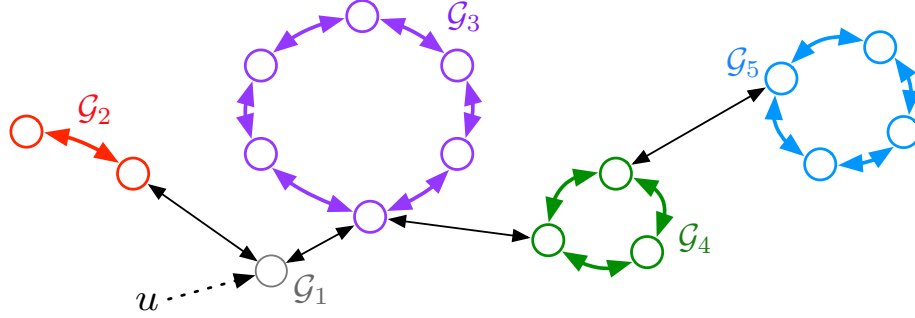
Figure 3.2: A sym-cactus $\mathcal{G} = \overline{\bigcup}_{i=1}^{5} \mathcal{G}_i$ rooted at the control node. Sym-cycles $\mathcal{G}_2, \ldots, \mathcal{G}_5$ are highlighted with different colors. Notice that $\mathcal{G}_1$ is not a sym-cycle because it comprises 1 node without a self-loop. See Definition 6.

Definition 5 *(Sym-cycle)* *A sym-cycle is a strongly connected digraph with $n \geq 1$ nodes, edge set*

$\{(i,j) \,:\, |i - j| = 1\} \cup \{(1,n), (n,1)\}\}$, *and symmetric weights $a_{ij} = a_{ji}$.* ☐

From Definition 5, the adjacency matrix of a sym-cycle is

$$A = \begin{cases} a_{ij} \neq 0 & \text{if } |i - j| = 1 \text{ or } (i,j) \in \{(1,n), (n,1)\}, \\ \\ a_{ij} = 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

Definition 6 *(Sym-cactus)* *A sym-cactus is a strongly connected digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ defined as $\mathcal{G} = \overline{\bigcup}_{i=1}^{m} \mathcal{G}_i$ and satisfying the following properties:*

1. $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ is a sym-cycle if $|\mathcal{V}_1| > 1$ (if $|\mathcal{V}_1| = 1$, we allow $\mathcal{G}_1$ to contain no edges, that is,

   $\mathcal{E}_1 = \emptyset$),

2. $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i)$ is a sym-cycle for every $i \in \{2, \ldots, m\}$,

3. the node sets satisfy $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset$, whenever $i \neq j$. ☐

Notice that, if $|\mathcal{V}_1| = 1$, the graph $\mathcal{G}_1$ in Definition 6 can either be a sym-cycle (thus having a self-loop), or a node without self-loop.

Remark 7 *(Stem, buds, cactus, and sym-cactus)* *Our definitions of sym-cycle and sym-cactus are compatible with the classic notions of stem, bud, and cactus as defined in [34]. In particular,*

because we focus on networks with symmetric weights, stems, buds, and cacti [34] become equivalent to interconnected sym-cycles. □

We say that a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is spanned by $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ if $\mathcal{V}' = \mathcal{V}$ and $\mathcal{E}' \subseteq \mathcal{E}$. Further, the sym-cactus $\mathcal{G} = \overline{\bigcup}_{i=1}^{m} \mathcal{G}_i$ is rooted at the node i if i is a node of $\mathcal{G}_1$. Fig. 3.2 illustrates the definitions of sym-cycle and sym-cactus rooted at node i.

The following lemma shows that every sym-cycle is structurally controllable from any node. That is, for almost all symmetric choices of network weights, every cycle network is controllable independently of the location of the control node.

Lemma 8 (Every sym-cycle is structurally controllable) *Let $A \in \mathbb{R}^{n \times n}$ be the adjacency matrix of a sym-cycle. The pair (A, b^i) is structurally controllable for all $i \in \{1, \ldots, n\}$.*

Proof. Owing to Theorem 1 we need to show that, for every sym-cycle and control node, there exists a choice of weights rendering the network controllable. Without affecting generality, we assume that the control node is $i = 1$ (if $i \neq 1$, simply apply a similarity transformation PAP^{T} via a permutation matrix P to reorder the nodes as desired).

If $n \leq 2$, the network is clearly controllable. For $n > 2$, partition the matrix A as

$$A = \begin{bmatrix} 0 & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{12} \in \mathbb{R}^{1 \times (n-1)}$, $A_{21} \in \mathbb{R}^{(n-1) \times 1}$ and $A_{22} \in \mathbb{R}^{(n-1) \times (n-1)}$. Notice that A_{22} is a tridiagonal matrix.

Suppose that the pair (A, b^1) is not structurally controllable. Then, for all choices of weights, there exists an eigenvector[3] v of A such that $v^{\mathsf{T}} b^1 = 0$ [7]. Thus, $v = [v_1, v_2, \ldots, v_n]^{\mathsf{T}} = [v_1, \bar{v}]^{\mathsf{T}} = [0, \bar{v}]^{\mathsf{T}}$, and the eigenproblem $Av = \lambda v$ becomes

$$\begin{bmatrix} 0 & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} 0 \\ \bar{v} \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda \bar{v} \end{bmatrix}. \tag{3.7}$$

---

[3]Since $A = A^{\mathsf{T}}$, we do not distinguish between left and right eigenvectors.

From (3.7), the pair (A, b^1) is uncontrollable if and only if A_{22} has an eigenvector $\bar{v}$ that lies in the null space of A_{12}. Equivalently, $A_{22}\bar{v} = \lambda\bar{v}$ and $A_{12}\bar{v} = a_{12}\bar{v}_1 + a_{1n}\bar{v}_{n-1} = 0$.

Assign all the weights of A_{22} as 1 (or any other constant), and notice that A_{22} is a Toeplitz tridiagonal matrix with eigenvectors $\bar{v}^i = [\bar{v}^i_j] = \left[\sin\left(\frac{ji\pi}{n}\right)\right]$, for $i, j \in \{1, \ldots, n-1\}$ [43, Example 7.2.5]. Finally, to ensure controllability of (A, b^1), select a_{12} and a_{1n} such that, for all $i \in \{1, \ldots, n-1\}$,

$$a_{12}\sin\left(\frac{i\pi}{n}\right) + a_{1n}\sin\left(\frac{(n-1)i\pi}{n}\right) \neq 0. \tag{3.8}$$

Notice that (3.8) can be ensured by $|a_{12}| \neq |a_{1n}|$. In fact,

1. for i odd, $\sin\left(\frac{i\pi}{n}\right) = \sin\left(\frac{(n-1)i\pi}{n}\right)$ because

$$\pi - \frac{i\pi}{n} + 2k\pi - \frac{(n-1)i\pi}{n} = (2k + 1 - i)\pi = 0,$$

   by selecting $k = (i-1)/2$;

2. for i even, $\sin\left(\frac{i\pi}{n}\right) = -\sin\left(\frac{(n-1)i\pi}{n}\right)$ because

$$\frac{(n-1)i\pi}{n} - 2k\pi + \frac{i\pi}{n} = (i - 2k)\pi = 0,$$

   by selecting $k = i/2$.

This concludes the proof. ∎

Lemma 8 implies that every sym-cycle is controllable for almost every symmetric choice of weights. In particular, the following choice of weights yields a controllable sym-cycle from node i (see the proof of Lemma 8 and Eq. (3.8)):

$$A = \begin{cases} a_{ij} = a, & \text{if } |i - j| = 1, \text{ and} \\ a_{1n} = a_{n1} = b, & \text{with } |a| \neq |b|, \end{cases} \tag{3.9}$$

for some nonzero constants a and b. We next show that sym-cacti are also a fundamentally controllable structure contained in every structurally controllable symmetric network.

Theorem 9 *(Structural controllability of symmetric networks with one control node)*

The network $\mathcal{G}$ with control node i is structurally controllable with symmetric weights if and only if it is spanned by a sym-cactus rooted at i.

The proof of Theorem 9 is postponed to the Appendix. In Theorem 9 we show that a necessary and sufficient condition for structural controllability of networks with symmetric weights is the existence of a spanning sym-cactus rooted at the control node. This result implies that the symmetry constraint on the network weights does not prevent controllability if the same un-constrained network is structurally controllable. It should be noticed that a sym-cactus is not, in general, strongly structurally controllable [44]. That is, there exist choices of weights that render a sym-cactus uncontrollable. We next illustrate a systematic procedure to construct an uncontrollable sym-cactus composed of controllable sym-cycles.

Example 10 *(Uncontrollable sym-cactus)* *Consider the sym-cactus $\mathcal{G} = \mathcal{G}_1 \bar{\cup} \mathcal{G}_2$ with control node 1 and adjacency matrix*

$$
A = \left[\begin{array}{ccc|ccc}
0 & 1 & 2 & 0 & 0 & 0 \\
1 & 0 & 1 & c_2 & 0 & 0 \\
2 & 1 & 0 & 0 & 0 & 0 \\
\hline
0 & c_2 & 0 & 0 & 3 & -4 \\
0 & 0 & 0 & 3 & 0 & 3 \\
0 & 0 & 0 & -4 & 3 & 0
\end{array} \right],
$$

where the diagonal blocks are the adjacency matrices of the sym-cycles $\mathcal{G}_1$ and $\mathcal{G}_2$, and the remaining blocks denote the interconnection between $\mathcal{G}_1$ and $\mathcal{G}_2$ with weight c_2. It can be verified that $\lambda_0 = -2$ is a transmission zero of the system [7]

$$
\delta(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix},
$$

$$
y = \begin{bmatrix} 1 & 2 \end{bmatrix} x,
$$

29

and that the pair (A, b^1) is uncontrollable when c_2 satisfies

$$\left(\begin{bmatrix} 2 & 3 & -4 \\ 3 & 2 & 3 \\ -4 & 3 & 2 \end{bmatrix}_{1,1}^{-1} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}_{2,2}^{-1} \right)^{-\frac{1}{2}} = 6.292853089,$$

where $[M]_{i,i}^{-1}$ denotes the i-th diagonal entry of the matrix M^{-1}. The reader is referred to the proof

of Theorem 9, Case 2.a, for a detailed derivation of this result. □

Following the above discussion and the derivation in the proof of Theorem 9, we next

describe an algorithm to assign the weights of a sym-cactus to guarantee controllability. To this

aim, let $\text{spec}(M)$ denote the spectrum of the matrix M, and notice that the adjacency matrix A of

the sym-cactus mbox $\mathcal{G} = \overline{\bigcup}_{i=1}^{m} \mathcal{G}_i$ can be written recursively as $(k = 2, \ldots, m)$

$$A_k = \begin{bmatrix} A_{k-1} & c_k e_{q_k} e_1^{\mathsf{T}} \\ c_k e_1 e_{q_k}^{\mathsf{T}} & H_k \end{bmatrix}, \tag{3.10}$$

where H_k is the adjacency matrix of $\mathcal{G}_k$, $A_1 = H_1$, $A_m = A$, and $c_k \neq 0$, for some index $q_k \in$

$\{1, \ldots, \sum_{j=1}^{k-1} |\mathcal{V}_j|\}$.[4] Let

$$A_{k-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{3.11}$$

where A_{11} is a scalar, and let $\mathcal{Z}_k = \{\lambda : A_{12}(A_{22} - \lambda I)^{-1} e_{q_k} = 0\}$ be the zeros of the single-input

single-output system $(A_{22}, e_{q_k}, A_{12})$. Then, the pair (A, b^1) can be made controllable by selecting

the weights in $\mathcal{G}_k$ to recursively satisfy the following conditions:

1. (H_k, b^1) is controllable (see (3.9) for a choice of weights),

2. $\text{spec}(H_k) \cap \text{spec}(A_{22}) = \emptyset$, and

3. for all $\lambda \in \mathcal{Z}_k$, $c_k^{-2} \neq \left[H_k - \lambda I \right]_{1,1}^{-1} \left[A_{22} - \lambda I \right]_{q_k, q_k}^{-1}$,

---
Algorithm 1 *Design of controllable sym-cactus*

Input   : $\{H_k \ : \ k = 1, \ldots, m, \ H_k \text{ satisfying } (3.6)\}$

Output   : Controllable pair (A, b^1), with A adjacency matrix of the sym-cactus $\mathcal{G} = \overline{\bigcup}_{i=1}^{m} \mathcal{G}_i$ rooted at 1

1 Select the weights of H_1 as in (3.9)

2 Set $A_1 = H_1$

   for $k = 2 : m$ **do**

3   Partition A_{k-1} according to (3.11)

4   Select $q_k \in \{1, \ldots, \sum_{j=1}^{k-1} |\mathcal{V}_j|\}$

5   Select the weights of H_k as in (3.9) and so that $\mathrm{spec}(H_k) \cap \mathrm{spec}(A_{22}) = \emptyset$

6   Compute $\mathcal{Z} = \{\lambda \ : \ A_{12}(A_{22} - \lambda I)^{-1} e_{q_k} = 0\}$

7   Select $c_k \neq c_\lambda$ for every $\lambda \in \mathcal{Z}$, where $c_\lambda = \left(\left[H_k - \lambda I \right]_{1,1}^{-1} \left[A_{22} - \lambda I \right]_{q_k, q_k}^{-1} \right)^{-\frac{1}{2}}$

8   Generate A_k as in (3.10)

9 **return** $A = A_m$. The pair (A, b^1) is controllable

A procedure to construct a controllable sym-cactus is summarized in Algorithm 1, whose complexity is linear in the number of sym-cycles and cubic in their dimension.

Remark 11 *(**Structural controllability of symmetric networks with multiple dedicated control nodes**) Theorem 9 can be extended to the case of multiple dedicated control nodes; that is, when the input matrix in (4.1) satisfies $B = \begin{bmatrix} e_{c_1} & \cdots & e_{c_m} \end{bmatrix}$ and $\{c_1, \ldots, c_m\} \subseteq \mathcal{V}$ is the set of control nodes. In particular, a network $\mathcal{G}$ with m control nodes $\{c_1, \ldots, c_m\}$ is structurally controllable with symmetric weights if and only if it is spanned by a disjoint union of sym-cacti rooted at the nodes $\{c_1, \ldots, c_m\}$. The necessity of this result follows directly from [45, Theorem 1], while its sufficiency is obtained by applying the same steps as in the proof of Theorem 9 to each disjoint cactus.* $\qquad\square$

We conclude this section with an example of structural controllability in the case of multiple dedicated control nodes.

---

[4]This recursive construction follows directly from Definition 6.

Example 12 *(Structural controllability with symmetric weights and multiple dedicated control nodes)* Consider the network in Fig. 3.3(a) with adjacency matrix

$$A = \left[\begin{array}{cccc|ccc} 0 & 0 & a_{13} & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{23} & a_{24} & 0 & 0 & 0 \\ a_{13} & a_{23} & 0 & a_{34} & a_{35} & 0 & 0 \\ 0 & a_{24} & a_{34} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & a_{35} & 0 & 0 & a_{56} & a_{57} \\ 0 & 0 & 0 & 0 & a_{56} & 0 & a_{67} \\ 0 & 0 & 0 & 0 & a_{57} & a_{67} & 0 \end{array} \right]$$

and control vector $b^1 = e_1$. The pair (A, b^1) is structurally controllable because of Theorem 9. In fact, there exists a sym-cactus $\mathcal{G} = \overline{\bigcup}_{i=1}^{3} \mathcal{G}_i$ that spans the network and is rooted at 1. Consider now the network in Fig. 3.3(c) with adjacency matrix $\tilde{A} = A$ and disconnect nodes 6 and 7; that is, $a_{67} = 0$. The pair $(\tilde{A}, b^1)$ is not structurally controllable because there is no sym-cactus that spans the network and is rooted at 1. However, by connecting an additional input at node 7, it is possible to span the network with a disjoint union of sym-cacti. That is, there exist distinct sym-cacti $\mathcal{G}_1 \overline{\cup} \mathcal{G}_2$ and $\mathcal{G}_4 \overline{\cup} \mathcal{G}_5$ that span the network and are rooted at 1 and 7, respectively. Therefore, by setting $b^2 = e_7$, the pair $(\tilde{A}, [b^1 \ b^2])$ is structurally controllable with symmetric weights. □

3.4 Application to structural brain networks

We apply our analysis to a class of structural brain networks reconstructed from diffusion magnetic resonance imaging (MRI) data, where nodes correspond to well-known brain regions and edges correspond to white matter connections between them [46]. Diffusion magnetic resonance images were acquired for a total of eight subjects in triplicate (mean age 27 ± 5 years, two female, two left handed), and at each scanning session a $T1$-weighted anatomical scan was acquired. For each subject, $n = 234$ regions were registered as areas of interest [10]
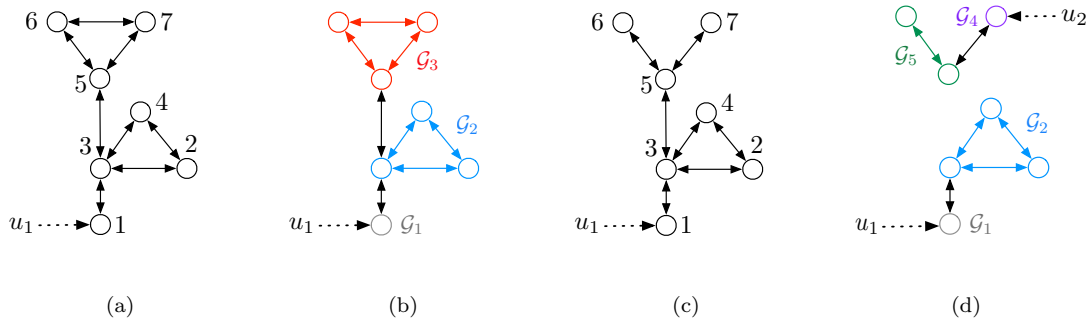
Figure 3.3: The networks considered in Example 12. (a) The network of the pair (A, b^1). (b) The network is structurally controllable because it is spanned by a sym-cactus rooted at 1. (c) The network of the pair $(\tilde{A}, b^1)$ is not structurally controllable with only one control input at node 1. (d) By adding a control input at node 7, the network recovers structural controllability because it is spanned by a disjoint union of sym-cacti rooted at nodes 1 and 7, respectively.

The network dynamics can be derived from the linearization of a general noise-free Wilson-Cowan system [47] and read as $\delta(x) = Ax + b^i u$, where A is a symmetric matrix that represents the anatomical connectivity of the brain. Further, $u : \mathbb{N} \to \mathbb{R}$ is the control input applied to the i-th brain region, and $x : \mathbb{N} \to \mathbb{R}^n$ is the vector containing the state of the brain regions over time. Examples of state values range from the magnitude of electrical activity [48] to the quantity of oxyhemoglobin and deoxyhemoglobin in the hemodynamic response [49]. Although brain dynamics may be nonlinear at the micro-scale, the study of linear network models for macro-scale neural dynamics has been validated in several studies (see e.g. [50]), and has given access to theoretical and practical tools that are particularly useful around an operating point [10, 51]. Controllability of this class of networks has been examined in different studies, including [10], via numerical controllability tests. Yet, because of the large cardinality of these networks, most controllability tests suffer from numerical instabilities, sometimes leading to competing conclusions [9, 10]. Further, because typical diffusion MRI techniques produce symmetric adjacency matrices, the graphical investigation of structural controllability for this type of networks was, up to now, not possible.

As illustrated in Fig. 3.4, the brain networks in our dataset are spanned by a Hamiltonian path,[5] which is a special case of a sym-cactus. Theorem 9 implies that, despite having symmetric

---

[5]A path in a graph is Hamiltonian if it visits all the vertices exactly once.
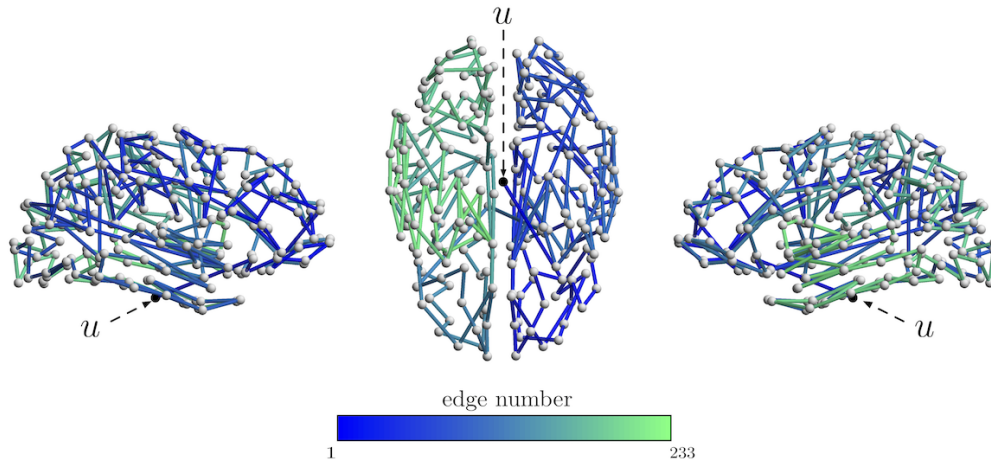
Figure 3.4: Sagittal and axial view of the structural brain network and a spanning Hamiltonian path. Each node represents a brain region of the anatomical scans. The Hamiltonian path starts from the region representing the control node (*brain stem*). Regions are plotted according to the mean location of voxels in each of the 234 parcels in the Lausanne atlas [46] and averaged over the cohort of healthy adult subjects. This figure was obtained with *BrainNet Viewer* [52].

weights, networks reconstructed from diffusion MRI data are structurally controllable from *any* single brain region, thus controllable for almost every symmetric choice of weights.

3.5 Conclusion

In this chapter we derive necessary and sufficient graph-theoretic conditions for structural controllability of networks with symmetric weights and one control node. Because weights need to be symmetric, classic results from structural systems theory cannot be directly applied. Surprisingly, we show that network controllability remains a generic property even when the weights are symmetric, and that a network with symmetric weights is structurally controllable if and only if its unconstrained equivalent network is structurally controllable; that is, if and only if it is spanned by a (symmetric) cactus. While our analysis focuses on symmetric weights and a single control node, as discussed in Remark 1 and 3, our results extend directly to other classes of parameter constraints and to the case of multiple dedicated control nodes.

3.6 Appendix

We now prove some instrumental results and Theorem 9.

Lemma 13 *(**Controllability of subsystems**) Consider the network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with control nodes $\mathcal{K} \subset \mathcal{V}$, input matrix $B_{\mathcal{K}} = \begin{bmatrix} e_1 & \dots & e_m \end{bmatrix}$, and adjacency matrix A partitioned as*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} \in \mathbb{R}^{m \times m}$ and $A_{22} \in \mathbb{R}^{(n-m) \times (n-m)}$. If the pair $(A, B_{\mathcal{K}})$ is controllable, then (A_{22}, A_{21}) is also controllable.

Proof. If (A_{22}, A_{21}) is not controllable, then there exists an eigenvector v_2 associated with $\lambda \in$ spec(A_{22}) satisfying [7]

$$v_2^{\mathsf{T}} \begin{bmatrix} A_{21} & A_{22} - \lambda I \end{bmatrix} = 0.$$

Let $v^{\mathsf{T}} = \begin{bmatrix} 0^{\mathsf{T}} & v_2^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$, and notice that

$$\begin{bmatrix} 0^{\mathsf{T}} & v_2^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} A_{11} - \lambda I & A_{12} \\ A_{21} & A_{22} - \lambda I \end{bmatrix} = 0.$$

Then, v is a left eigenvector of A associated with the eigenvalue $\lambda \in$ spec(A), and it satisfies $v^{\mathsf{T}} B_{\mathcal{K}} = 0$. This implies that $(A, B_{\mathcal{K}})$ is not controllable, and concludes the proof. ∎

Lemma 14 *(**Eigenspace of perturbed matrix**) Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let $\Delta = e_i e_i^{\mathsf{T}}$, with $i \in \{1, \dots, n\}$. Then, $\lambda \in$ spec$(A + c\Delta)$ for all $c \in \mathbb{R}$ if and only if there exists $v \neq 0$ satisfying $(A - \lambda I)v = 0$ and $\Delta v = 0$.*

Proof. *(If)* The sufficiency of the statement follows by noting that $(A - \lambda I + c\Delta)v = (A - \lambda I)v = 0$. *(Only if)* Let the vectors $v_c \neq 0$ and $v_0 \neq 0$ satisfy $v_c^{\mathsf{T}}(A - \lambda I + c\Delta) = 0$ and $(A - \lambda I)v_0 = 0$, respectively. Then, for all $c \in \mathbb{R}$, $v_c^{\mathsf{T}}(A - \lambda I + c\Delta)v_0 = c v_c^{\mathsf{T}} \Delta v_0 = 0$. Let $v_{\bar{c}}$ denote the vector v_c with

$c = \bar{c} \neq 0$. Notice that, because $\Delta = e_i e_i^\mathsf{T}$, $v_{\bar{c},i} v_{0,i} = 0$, where $v_{\bar{c},i}$ and $v_{0,i}$ denote the i-th element of $v_{\bar{c}}$ and v_0, respectively. Let $v = v_0$ if $v_{0,i} = 0$, and $v = v_{\bar{c}}$ otherwise. To conclude, notice that $v \neq 0$, $\Delta v = 0$, and $(A - \lambda I)v = 0$. ∎

We are now ready to prove Theorem 9.

Proof of Theorem 9: *(Only if)* Assume that $\mathcal{G}$ is structurally controllable from the node i. From [34], there must exist a directed cactus $\mathcal{D}$ rooted at i that spans $\mathcal{G}$. Because $\mathcal{G}$ has symmetric weights, this also implies the existence of a sym-cactus, which is obtained by adding edges to $\mathcal{D}$ to make it symmetric. See Remark 7 for a discussion of directed and symmetric cacti.

(If) Let the network be spanned by the sym-cactus $\mathcal{G} = \overline{\bigcup}_{i=1}^m \mathcal{G}_i$ rooted at the control node. Let A_k be the adjacency matrix of $\overline{\bigcup}_{i=1}^k \mathcal{G}_i$, $k \leq m$. Without loss of generality, we assume $b^\mathsf{i} = b^1$ (if $b^\mathsf{i} \neq b^1$, reorder the nodes). We will construct a controllable realization (A_m, b^1), thus proving that the original network admits a controllable realization. The claimed statement then follows from Theorem 1.

We proceed by induction. In the base step, Lemma 8 concludes on the controllability of the pair (A_1, b^1). In the inductive step, we assume that (A_{k-1}, b^1) is controllable, and show that (A_k, b^1) is controllable. Let A_k be partitioned as

$$A_k = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \tag{3.12}$$

where $A_{11} \in \mathbb{R}$, $A_{22} \in \mathbb{R}^{(n_1-1) \times (n_1-1)}$, and $A_{33} \in \mathbb{R}^{n_2 \times n_2}$, with n_1 and n_2 being the dimension of A_{k-1} and the difference between the dimension of A_k and A_{k-1}, respectively. Notice that A_{33} corresponds to H_k in decomposition (3.10). We show that A_k has no eigenvector v of the form

$$v = \begin{bmatrix} 0 & v_1^\mathsf{T} & v_2^\mathsf{T} \end{bmatrix}^\mathsf{T}, \tag{3.13}$$

which, by the eigenvector test, implies that (A_k, b^1) is controllable. Due to the definition of the operator $\overline{\bigcup}$ (a single connection between adjacent sym-cycles) and by exploiting the decomposition

of A_k in (3.10), we have that either

(1) $A_{32} = A_{23}^\mathsf{T} = 0$ and $A_{31} = A_{13}^\mathsf{T} = c_k e_1 \neq 0$, or

(2) $A_{31} = A_{13}^\mathsf{T} = 0$ and $A_{32} = A_{23}^\mathsf{T} = c_k e_1 e_{q_k}^\mathsf{T} \neq 0$,

where e_1, e_{q_k} are canonical vectors of appropriate dimensions.

Case (1) Consider the eigenproblem $A_k v = \lambda v$. For v to be of the form (3.13), λ must be an eigenvalue of both A_{22} and A_{33}. Therefore, by choosing the weights in A_{33} such that $\mathrm{spec}(A_{33}) \cap \mathrm{spec}(A_{22}) = \emptyset$, we obtain a controllable (A_k, b^1). Notice that such a choice of weights always exists because A_{33} has generically full rank.[6] For instance, given a full rank realization of A_{33}, we can multiply A_{33} by a suitable constant $c \in \mathbb{R}$ to guarantee that $\mathrm{spec}(cA_{33}) \cap \mathrm{spec}(A_{22}) = \emptyset$.

Case (2) Define the matrix $P(\lambda)$ as

$$
P(\lambda) = \begin{bmatrix} A_{12} & 0 \\ A_{22} - \lambda I & A_{23} \\ A_{32} & A_{33} - \lambda I \end{bmatrix}.
$$

Due to (3.13), the eigenproblem $A_k v = \lambda v$ reduces to

$$
P(\lambda) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0. \tag{3.14}
$$

We will show that $P(\lambda)$ is full rank for all λ, thus ensuring that an eigenvector as (3.13) cannot exist. As in *Case (1)*, we choose weights in A_{33} such that $\mathrm{spec}(A_{33}) \cap \mathrm{spec}(A_{22}) = \emptyset$. Thus, we consider 3 cases:

(2.a) $\lambda \notin \mathrm{spec}(A_{22}) \cup \mathrm{spec}(A_{33})$,

(2.b) $\lambda \in \mathrm{spec}(A_{22})$, and

(2.c) $\lambda \in \mathrm{spec}(A_{33})$.

---

[6]The graph with adjacency matrix A_{33} contains a set of n_2 edges, for instance $\mathcal{M} = \{(1,2), (2,3), \ldots, (n_2 - 1, n_2), (n_2, 1)\}$), where no two edges point to the same node. Such set of edges is called a matching of size n_2, and its existence guarantees that A_{33} is generically full rank [53, §1.1.2].

Case (2.a) Because $A_{22} - \lambda I$ and $A_{33} - \lambda I$ are invertible,

$$\text{Rank}\,(P(\lambda)) = \text{Rank}\left(P(\lambda) \begin{bmatrix} (A_{22} - \lambda I)^{-1} & 0 \\ \\ 0 & (A_{33} - \lambda I)^{-1} \end{bmatrix} \right)$$

$$= \text{Rank}\left(\begin{bmatrix} A_{12}(A_{22} - \lambda I)^{-1} & 0 \\ \\ I & c_k T_3 \\ \\ c_k T_2 & I \end{bmatrix} \right)$$

where $T_2 = e_1 e_{q_k}^{\mathsf{T}}(A_{22} - \lambda I)^{-1}$, and $T_3 = e_{q_k} e_1^{\mathsf{T}}(A_{33} - \lambda I)^{-1}$.

Notice that, for any vector v_3 of appropriate dimension we have $T_3 v_3 = \alpha e_{q_k}$, for some value α dependent on λ and A_{33}. Similarly, $T_2 v_2 = \beta e_1$, for some value β dependent on λ and A_{22}. Further, for any fixed λ, there exists a value c_k such that

$$\begin{bmatrix} I & c_k T_3 \\ \\ c_k T_2 & I \end{bmatrix} \tag{3.15}$$

is invertible. In fact, elementary column operations reveal that

$$\text{Rank}\left(\begin{bmatrix} I & c_k T_3 \\ \\ c_k T_2 & I \end{bmatrix} \right) = \text{Rank}\left(\begin{bmatrix} I - c_k^2 T_3 T_2 & c_k T_3 \\ \\ 0 & I \end{bmatrix} \right).$$

Notice that $T_3 T_2$ is a rank-1 matrix and that $\text{spec}(I - T_3 T_2) = \{1, \ldots, 1, 1 - c_k^2 \tilde{\lambda}\}$, where $\tilde{\lambda}$ is the only nonzero eigenvalue of $T_3 T_2$. Thus, (3.15) is invertible whenever $c_k^2 \neq \tilde{\lambda}^{-1}$.

Let $\mathcal{Z} = \{\lambda \; : \; A_{12}(A_{22} - \lambda I)^{-1} e_{q_k} = 0\}$, and let c_k be such that (3.15) is invertible for all $\lambda \in \mathcal{Z}$. Then, $P(\lambda)$ is also full rank for all $\lambda \in \mathcal{Z}$. Next, assume by contradiction that $P(\lambda)$ loses rank for some value $\bar{\lambda} \notin \mathcal{Z}$. Then, there exist nonzero w_2 and w_3 such that

$$\begin{bmatrix} A_{12}(A_{22} - \bar{\lambda} I)^{-1} & 0 \\ \\ I & c_k T_3 \\ \\ c_k T_2 & I \end{bmatrix} \begin{bmatrix} w_2 \\ \\ w_3 \end{bmatrix} = 0.$$

We have $w_2 = -c_k T_3 w_3 = -c_k \alpha e_{q_k}$, and $c_k \alpha A_{12}(A_{22} - \bar{\lambda} I)^{-1} e_{q_k} = 0$. Notice that $\alpha \neq 0$. Otherwise, $T_3 w_3 = 0$ and, consequently, $w_2 = 0$ and $w_3 = 0$. Further, $A_{12}(A_{22} - \bar{\lambda} I)^{-1} e_{q_k} \neq 0$ because $\bar{\lambda} \notin \mathcal{Z}$. We conclude that, when c_k is such that (3.15) is invertible for all $\lambda \in \mathcal{Z}$, $P(\lambda)$ is full rank.

38

Case (2.b) Because $A_{33} - \lambda I$ is invertible,

$$\text{Rank}\,(P(\lambda)) = \text{Rank}\left(P(\lambda)\begin{bmatrix} I & 0 \\ 0 & (A_{33}-\lambda I)^{-1} \end{bmatrix}\right) = \text{Rank}\left(\begin{bmatrix} A_{12} & 0 \\ A_{22}-\lambda I & c_k T_3 \\ A_{32} & I \end{bmatrix}\right),$$

where $T_3 = e_{q_k} e_1^\mathsf{T} (A_{33}-\lambda I)^{-1}$. By means of elementary column operations we obtain

$$\text{Rank}\,(P(\lambda)) = \text{Rank}\left(\begin{bmatrix} A_{12} & 0 \\ A_{22}-\lambda I - c_k T_3 A_{32} & c_k T_3 \\ 0 & I \end{bmatrix}\right).$$

Notice that, if $\lambda \notin \text{spec}(A_{22} - c_k T_3 A_{32})$ for some c_k, then $P(\lambda)$ can be made full rank by a selection of c_k. Instead, if $\lambda \in \text{spec}(A_{22} - \lambda I - c_k T_3 A_{32})$ for all values of c_k, then, due to Lemma 14, $(A_{22}-\lambda I - c_k T_3 A_{32})v = 0$, for some fixed eigenvector v and for all c_k. Because (A_k, b^1) is controllable by the induction hypothesis, so is the pair $(A_{22}, A_{21}) = (A_{22}, A_{12}^\mathsf{T})$ by Lemma 13. We conclude that $A_{12}v \neq 0$. This implies that, for all values of c_k, the submatrix

$$\begin{bmatrix} A_{12} \\ A_{22}-\lambda I - c_k T_3 A_{32} \end{bmatrix}$$

is full rank and, consequently, so is $P(\lambda)$.

Case (2.c) Because $A_{22} - \lambda I$ is invertible,

$$\text{Rank}\,(P(\lambda)) = \text{Rank}\left(P(\lambda)\begin{bmatrix} (A_{22}-\lambda I)^{-1} & 0 \\ 0 & I \end{bmatrix}\right) = \text{Rank}\left(\begin{bmatrix} A_{12}(A_{22}-\lambda I)^{-1} & 0 \\ I & A_{23} \\ c_k T_2 & A_{33}-\lambda I \end{bmatrix}\right),$$

where $T_2 = e_1 e_{q_k}^\mathsf{T}(A_{22}-\lambda I)^{-1}$. By means of elementary row operations we obtain that $\text{Rank}\,(P(\lambda))$ equals

$$\text{Rank}\left(\begin{bmatrix} A_{12}(A_{22}-\lambda I)^{-1} & 0 \\ I & A_{23} \\ 0 & A_{33}-\lambda I - c_k A_{23} T_2 \end{bmatrix}\right).$$

Notice that, if $\lambda \notin \mathrm{spec}(A_{33} - c_k A_{23} T_2)$ for some c_k, then $P(\lambda)$ can be made full rank by a selection of c_k. Instead, if $\lambda \in \mathrm{spec}(A_{33} - \lambda I - c_k A_{23} T_2)$ for all values of c_k, then, due to Lemma 14, $(A_{33} - \lambda I - c_k A_{23} T_2)v = 0$, for some fixed eigenvector v and for all c_k. Because (A_{33}, b^i) can be made controllable for all indices i due to Lemma 8, the submatrix

$$
\begin{bmatrix}
A_{23} \\
A_{33} - \lambda I - c_k A_{23} T_2
\end{bmatrix}
\tag{3.16}
$$

is full rank. To make $P(\lambda)$ full rank, we proceed by contradiction. Suppose there exist nonzero v_1 and v_2 such that

$$
\begin{bmatrix}
A_{12}(A_{22} - \lambda I)^{-1} & 0 \\
I & A_{23} \\
0 & A_{33} - \lambda I - c_k A_{23} T_2
\end{bmatrix}
\begin{bmatrix}
v_1 \\
v_2
\end{bmatrix}
= 0.
$$

Notice that $v_1 = -A_{23}v_2 = -c_k e_{q_k} e_1^\mathsf{T} v_2$ and that v_1 has exactly one nonzero entry (q_k) due to (3.16) being full rank. Finally, $A_{12}(A_{22} - \lambda I)^{-1} v_1 = 0$ implies that λ must be a transmission zero of the single-input single-output system $(A_{22}, e_{q_k}, A_{12})$ [7]. Thus, $P(\lambda)$ can be made full rank by selecting A_{33} such that its eigenvalues are different from the transmission zeros of the system $(A_{22}, e_{q_k}, A_{12})$.

In conclusion, by choosing A_{33} and the interconnection weight c_k as discussed in *Cases (1), (2.a), (2.b)*, and *(2.c)*, we obtain a controllable realization of the sym-cactus $\mathcal{G} = \overline{\bigcup}_{i=1}^m \mathcal{G}_i$, thus concluding the inductive procedure. $\blacksquare$

Chapter 4

The Structured Controllability Radius of Symmetric (Brain) Networks

In this chapter, we study the controllability radius – that is, the distance from the closest uncontrollable realization – of undirected networks, with a particular emphasis on the controllability radius of empirically reconstructed brain networks. Our results shed light on the unique topological features of structural brain networks. We refer the interested reader to the publication [51] for further details.

4.1 Introduction

The question of controllability of natural and man-made network systems has recently received considerable attention. In the context of the human brain, the study of various controllability properties may not only shed light into the organization and function of different neural circuits, but

also inform the design and implementation of minimally invasive yet effective intervention protocols to treat neurological disorders [54]. Although the study of the human brain as a network system is still in its infancy, some recent results, e.g., see [10, 55, 56], have suggested that the complexity of the brain and its underlying principles can be further untangled with tools from control theory and network science [57].

While the dynamics of most brain processes is clearly nonlinear, linearized models with empirically reconstructed network matrices have been proved useful to characterize how the anatomical structure of the brain influences its dynamic functions [58, 59]. In this chapter we follow this line of work, and model the dynamics of a brain network as a linear, discrete-time, time-invariant system, where the network matrix is empirically estimated from diffusion MRI data. A key feature of these empirically reconstructed networks is that the estimated edges are undirected, giving rise to symmetric network matrices [46]. This constraint on the edge weights adds a layer of complexity to the study of network controllability [30, 60]. For instance, it prevents the use of most tools developed within structural control theory [20].

In this chapter we propose and analyze a novel notion of controllability for symmetric networks, namely, the structured controllability radius. Specifically, we quantify the controllability degree of a network with the size of the smallest symmetric perturbation (measured with the Frobenius norm) that has a given sparsity pattern and renders the network uncontrollable. We provide analytical conditions to compute the structured controllability radius of a symmetric network, and use these conditions to compare a set of brain networks with a class of random networks. Our results show that the considered brain networks feature a controllability radius that is consistently smaller when compared to the considered random networks, suggesting that the topological organization of the brain may lead to unique dynamical features different from those of random network models [61].

Related work Different notions of controllability of a system have been proposed over the years. Starting from the binary definition of controllability proposed in [62], Gramian-based metrics have been proposed to provide a quantitative measure of the controllability degree of a system and, more

recently, of a network based on the energetic effort needed to control the state towards a desired value [17,63]. In [64,65] an alternative notion of controllability is introduced, where the controllability degree of a system is quantified by the smallest norm of a perturbation of the system parameters causing uncontrollability. Later, this notion of controllability radius has been extended to account for several types of constrained perturbations (Hermitian, symmetric, and skew-symmetric) [66]. Yet, with the exception of [67], the use of the controllability radius to quantify the controllability degree of a network has not been investigated. In this case, because only existing edges can typically be modified, perturbations need to feature a pre-specified sparsity pattern, a constraint that renders classic results on the controllability radius inapplicable. In this chapter we improve upon existing results, particularly [67], by focusing on symmetric and structured perturbations, by deriving an explicit set of equations for the computation of the structured controllability radius, and by exploiting our results to compare a class of brain networks with random networks with similar weights.

Contribution The contribution of this chapter is three-fold. First, we propose a novel notion of controllability degree for networks with symmetric adjacency matrix, namely, the structured controllability radius, which equals the smallest Frobenius norm of a symmetric perturbation that renders the network uncontrollable and has a pre-specified sparsity pattern. Second, we derive explicit necessary and sufficient conditions for the computation of the structured controllability radius, and illustrate our procedure through various examples. Third and finally, we use our notion of structured controllability radius to compare a class of brain networks reconstructed from diffusion MRI data with a set of random networks with similar weights and topologies. Our results show that the controllability radius of brain networks is consistently smaller than in the case of random networks, suggesting that the anatomical organization of the brain may favor dynamic properties different from controllability.

Chapter organization The remainder part of the chapter is organized as follows. In Section 4.2 we introduce our network model, we define different controllability metrics, and we state the controllability radius optimization problem. In Section 4.3 we derive our conditions for the computation

43

of the structured controllability radius. Section 4.4 contains our numerical study of the structured

controllability radius of brain and random networks. Section 9.6 concludes the chapter.

Mathematical notation $\text{supp}(\cdot)$ denotes the support of a vector and $\text{vec}(\cdot)$ denotes the vector-

ization of a matrix. $\lambda_{\min}(\cdot)$ and $\sigma_{\min}(\cdot)$ denote the minimum eigenvalue and singular value of

a matrix. $\circ$ and $\otimes$ denote the Hadamard (element-wise) and Kronecker products, respectively.

$\mathbb{1}_n \in \mathbb{R}^n$ ($\mathbb{1}_{n \times n} \in \mathbb{R}^{n \times n}$) denotes a vector (matrix) of all ones. $(\cdot)^+$ denotes the Moore-Penrose

pseudo inverse of a matrix. $\|\cdot\|_{\text{F}}$, $\|\cdot\|_2$ and $tr(\cdot)$ denote the Frobenius norm, spectral norm and

trace of a matrix, respectively. e_i denotes the i-th canonical vector. Finally, we denote a positive

definite (positive semi-definite) matrix A with $A > 0$ ($A \geq 0$).

4.2 Model and problem statement

We consider networks represented by a weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, n\}$

and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ are the node and edge sets, respectively. Let $A = [a_{ij}]$ be the weighted adjacency

matrix of $\mathcal{G}$, where $a_{ij} = 0$ if $(i, j) \notin \mathcal{E}$ and $a_{ij} \in \mathbb{R}$ if $(i, j) \in \mathcal{E}$. Because we study brain networks

reconstructed from diffusion MRI images [46], we assume that $A = A^\mathsf{T}$. An example of adjacency

matrix is in Fig. 4.1.

The network dynamics is described by the following discrete-time linear time-invariant

system:

$$x(t+1) = Ax(t) + Bu(t), \tag{4.1}$$

where $x : \mathbb{N} \to \mathbb{R}^n$ is the vector containing the state of the nodes over time, $u : \mathbb{N} \to \mathbb{R}^m$ is the

control input that is applied to the network through the input matrix $B \in \mathbb{R}^{n \times m}$. Without loss of

generality, we assume that B has full rank.

The system (4.1) is controllable if there exists a control input that can steer the system from

a given initial state to any desired final state. Several notions exist to quantify the controllability

degree of a system. One such metric measures the control energy required to control the state
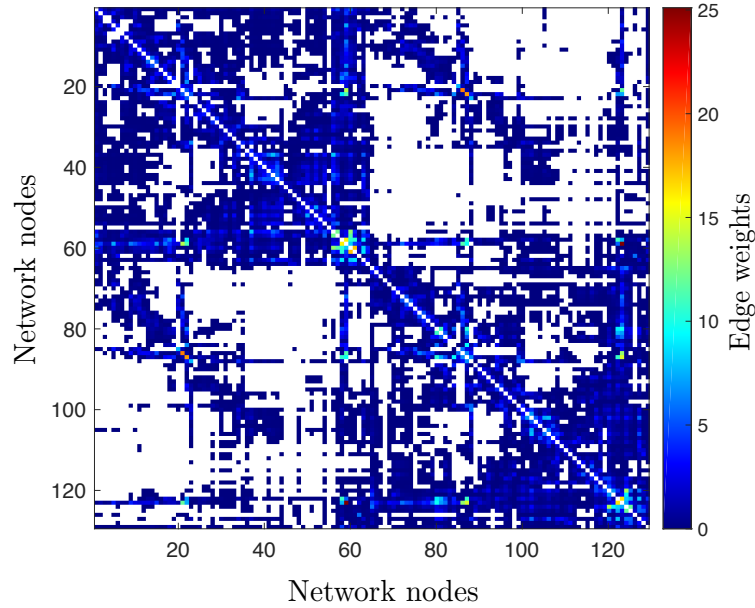
Figure 4.1: Example of an anatomical connectivity matrix derived from diffusion MRI scans of a brain. Notice that entries are symmetric and the diagonal entries are zero.

between two values, and is quantified by the controllability Gramian

$$W = \sum_{\tau=0}^{\infty} A^\tau BB^\mathsf{T}(A^\mathsf{T})^\tau.$$

Notice that $W \geq 0$ and, further, $W > 0$ if and only if the pair (A, B) is controllable [7]. As a known result in system theory [17,63], the upper bound for the control energy required to steer the network to a desired final state is inversely proportional to $\lambda_{\min}(W)$. Thus, if the eigenvalues of W are large, the required control energy is small and hence, the system has a larger controllability degree. In the case of brain networks, a larger controllability degree might enable or guide the use of less invasive treatments.

An alternative characterization of the degree of controllability of a system is provided by the controllability radius, which is defined as the smallest norm of a perturbation that renders the system uncontrollable. Mathematically, such perturbation is obtained by solving the minimization

problem

$$\mu = \min_{\Delta_A, \Delta_B} \quad \| [\Delta_A \ \Delta_B] \|_2$$

$$\text{s.t.} \qquad (A + \Delta_A, B + \Delta_B) \text{ is uncontrollable.}$$

Equivalently [68],

$$\mu = \min_{s \in \mathbb{C}} \sigma_{\min} \left([sI - A, \ B] \right).$$

Clearly, $\mu > 0$ if and only if $W > 0$. Further, when μ is small, a small perturbation of the system weights exist that renders the system uncontrollable.

Although the above metrics $\lambda_{\min}(W)$ and μ provide useful information regarding controllability degree of a system, they may not be directly applicable to networks due to the following reasons. First, the definition of μ allows for only *unstructured* perturbations Δ_A, Δ_B. In contrast, feasible network perturbations may be subject to constraints on their sparsity patterns and edge weights, as in the case of the networks considered in this chapter. Second, the above metrics are in terms of smallest eigenvalue/singular value and, consequently, they do not provide any insight about the magnitude and the distribution of the individual entries of the perturbation. For instance, they do not help in determining which edges of a network are more (or less) sensitive to perturbations with respect to making the system uncontrollable.

To overcome these limitations and to make the metric μ meaningful for structured networks, we reformulate the optimization problem to include both symmetry and sparsity constraints explicitly in the problem. We use the Frobenius norm to measure the size of the perturbation. Further, we consider perturbations only on the network weights, that is, $\Delta_B = 0$. For simplicity, in the remainder of the chapter we denote Δ_A by Δ. We introduce the sparsity constraints on Δ via a constraint graph $\mathcal{H} = (\mathcal{V}, \mathcal{E}_{\mathcal{H}})$, where $\mathcal{E}_{\mathcal{H}}$ denotes the set of edges that can be perturbed.[1] Let H be the 0-1 adjacency matrix associated with $\mathcal{H}$, and let $H^c = \mathbb{1}_{n \times n} - H$ be the unweighted

---

[1] The graph $\mathcal{H}$ has the same nodes as $\mathcal{G}$, but possibly different edge set.

complimentary adjacency matrix. Then, the sparsity constraints can be written as $H^c \circ \Delta = 0$. Since Δ is symmetric, we also assume H to be symmetric.[2]

The structured controllability radius of the pair (A, B) is the solution of the following minimization problem:

$$\min_{\Delta, v, \lambda} \ \|\Delta\|_{\mathrm{F}}^2 \tag{4.2}$$

$$\text{s.t.} \qquad \Delta = \Delta^{\mathsf{T}}, \qquad \text{(symmetry constraint)} \tag{4.2a}$$

$$(A + \Delta)v = \lambda v, \qquad \text{(eigenvalue constraint)} \tag{4.2b}$$

$$\|v\|_2^2 = 1, \qquad \text{(eigenvector constraint)} \tag{4.2c}$$

$$v^{\mathsf{T}} B = 0, \qquad \text{(uncontrollability)} \tag{4.2d}$$

$$H^c \circ \Delta = 0, \qquad \text{(structural constraint)} \tag{4.2e}$$

where constraint (4.2d) follows from the PBH uncontrollability test [7], and constraint (4.2c) is for uniqueness of v. Notice that the minimization problem (4.2) is not convex due to the eigenvalue constraint (4.2b). Consequently, multiple local minima may exist. This is a common feature in various minimum distance and eigenvalue assignment problems [69]. We conclude this section with the following remark.

Remark 15 *(Structured vs unstructured controllability radius) The minimization problem (4.2) admits the trivial solution $\Delta = 0$ if and only if the pair (A, B) is uncontrollable. Further, the controllability radius without structural constraints is always finite, that is, a finite perturbation causing uncontrollability always exists. Instead, the sparsity constraints (4.2e) may render the problem unfeasible (trivially, in the case where (A, B) is controllable and $H = 0$).* $\qquad \square$

---

[2]If H is not symmetric, construct a symmetric H' by removing a minimal set of edges from $\mathcal{H}$. It can be shown that the optimization problems with constraints H and H', respectively, admit the same solutions.

4.3 Solution to the optimization problem

In this section we derive a solution to the non-convex optimization problem (4.2). In the theory of equality constrained non-linear programming, the first-order optimality conditions are meaningful only when the optimal points satisfy the regularity condition given by rank $J = n_c$, where J is the Jacobian of the constraints and n_c equals the total number of independent equality constraints. This regularity condition is mild and usually satisfied for most classes of problems [70]. Before presenting the main result, we derive the Jacobian and state the regularity condition for the optimization problem (4.2). Given the constraint graph $\mathcal{H}$, let n_s and $\bar{n}_s$ satisfy

$$n_s = |\{(i,j) \ : \ H^c = [h_{ij}], \ j \geq i, \ h_{ij} = 1\}|,$$

$$\bar{n}_s = |\{(i,j) \ : \ H^c = [h_{ij}], \ h_{ij} = 1\}|,$$

and note that the constraint (4.2e) can be equivalently written as (see the proof of Lemma 16 for a formal definition of Q)

$$Q\text{vec}(\Delta) = 0, \tag{4.3}$$

for some 0-1 matrix Q of dimension $\bar{n}_s \times n^2$.

Lemma 16 *(**Jacobian of the constraints**) The Jacobian of the equality constraints (4.2a)-(4.2e) is given by*

$$J(\Delta, v, \lambda) = \begin{bmatrix} I - T_n & 0 & 0 \\ v^\mathsf{T} \otimes I & A + \Delta - \lambda I & -v \\ 0 & 2v^\mathsf{T} & 0 \\ 0 & B^\mathsf{T} & 0 \\ Q & 0 & 0 \end{bmatrix}, \tag{4.4}$$

where T_n is the n^2-dimensional permutation matrix satisfying $\text{vec}(\Delta^\mathsf{T}) = T_n \text{vec}(\Delta)$, and Q is as in (4.3). Further, the total number of independent scalar constraints in (4.2a)-(4.2e) is

$$n_c = \frac{n^2 + n}{2} + n + 1 + m + n_s. \tag{4.5}$$

Proof. We construct the Jacobian $J(\Delta, v, \lambda)$ by rewriting the constraints (4.2a)-(4.2e) in vectorized form and taking their derivatives with respect to $\delta \triangleq \text{vec}(\Delta)$, v and λ. Vectorization of (4.2a) yields $(I - T_n)\delta = 0$ and its derivatives read as the first block row of J. Among the total n^2 scalar constraints in (4.2a), $\frac{n^2-n}{2}$ are redundant resulting in only $\frac{n^2+n}{2}$ independent constraints. Using the property $\text{vec}(AB) = (B^\mathsf{T} \otimes I)\text{vec}(A)$, re-vectorization of (4.2b) yields $(A - \lambda I)v + (v^\mathsf{T} \otimes I)\delta = 0$, from which we obtain the second block row of J. Notice that (4.2b) consists of n scalar constraints. Differentiation of (4.2c) and (4.2d) is straightforward and it provides 1 and m rows, respectively. Finally, (4.2e) consists of $\bar{n}_s$ non-trivial sparsity constraints, which can be written as $Q\delta = 0$ where $Q = [e_{i_1}\ e_{i_2}\ \dots\ e_{i_{\bar{n}_s}}]^\mathsf{T}$ and $\{i_1, \dots, i_{\bar{n}_s}\} = \text{supp}(\text{vec}(H^c))$ is the set of indices indicating the ones in $\text{vec}(H^c)$. Because H^c is symmetric, only n_s constraints of (4.2e) in the lower-triangular (or upper-triangular) part of H^c are independent when we combine (4.2e) with the independent constraints of (4.2a). Thus, the total number of independent constraints in (4.2a)-(4.2e) is $n_c = \frac{n^2+n}{2} + n + 1 + m + n_s$ and this concludes the proof. ∎

We now solve the minimization problem (4.2).

Theorem 17 (*Structured controllability radius of symmetric networks*) *Let Δ^*, v^* and λ^* satisfy the constraints (4.2a)-(4.2e). Then, Δ^* is a local minimum of the minimization problem (4.2) if and only if, for some $l^* \in \mathbb{R}^n$ and $q^* \in \mathbb{R}^m$,*

$$\Delta^* = -\frac{1}{4}H \circ \left[v^*(l^*)^\mathsf{T} + l^*(v^*)^\mathsf{T}\right], \tag{4.6a}$$

$$\left[A + \Delta^* - \lambda^* I \quad B\right] \begin{bmatrix} l^* \\ q^* \end{bmatrix} = 0, \tag{4.6b}$$

$$(v^*)^\mathsf{T} l^* = 0, \tag{4.6c}$$

$$\text{rank } J(\Delta^*, v^*, \lambda^*) = n_c, \ \ and \tag{4.6d}$$

$$P^* D^* P^* \geq 0, \tag{4.6e}$$

where n_c is as in (4.5), D^* is the Hessian defined as

$$D^* = \begin{bmatrix} 2I & I \otimes l^* & 0 \\ I \otimes (l^*)^\mathsf{T} & 0 & -l^* \\ 0 & -(l^*)^\mathsf{T} & 0 \end{bmatrix},$$

P^* is the projection matrix of $J(\Delta, v, \lambda)$ defined as

$$P^* = I - J^+(\Delta^*, v^*, \lambda^*) J(\Delta^*, v^*, \lambda^*).$$

Proof. We prove the result using the Lagrange theorem for equality constrained minimization [71]. Let $S \in \mathbb{R}^{n \times n}$, $l \in \mathbb{R}^n$, $h \in \mathbb{R}$, $q \in \mathbb{R}^m$, and $M \in \mathbb{R}^{n \times n}$ be the Lagrange multipliers associated with constraints in (4.2), respectively. We make use of the following properties for the proof:

(a) $tr(A) = tr(A^\mathsf{T})$ and $tr(AB) = tr(BA)$,

(b) $\|A\|_\mathrm{F}^2 = tr(A^\mathsf{T}A) = \mathrm{vec}^\mathsf{T}(A)\mathrm{vec}(A)$,

(c) $\mathbb{1}_n^\mathsf{T}(A \circ B)\mathbb{1}_n = tr(A^\mathsf{T}B)$,

(d) $A \circ B = B \circ A$ and $A \circ (B \circ C) = (A \circ B) \circ C$,

(e) $A \circ (B + C) = (A \circ B) + (A \circ C)$ and $(A \circ B)^\mathsf{T} = A^\mathsf{T} \circ B^\mathsf{T}$,

(f) $\frac{d}{dX} tr(X^\mathsf{T}X) = 2X$ and $\frac{d}{dX} tr(AX) = A^\mathsf{T}$,

(g) $a^\mathsf{T}Xy = y^\mathsf{T}(I \otimes a^\mathsf{T})\mathrm{vec}(X) = \mathrm{vec}^\mathsf{T}(X)(I \otimes a)y$,

The Lagrange function for the optimization problem is

$$\mathcal{L}(\Delta, v, \lambda, S, l, h, q, M) = \|\Delta\|_\mathrm{F}^2 + \mathbb{1}_n^\mathsf{T}[S \circ (\Delta - \Delta^\mathsf{T})]\mathbb{1}_n + l^\mathsf{T}(A + \Delta - \lambda I)v$$

$$+ h(\|v\|_2^2 - 1) + q^\mathsf{T}B^\mathsf{T}v + \mathbb{1}_n^\mathsf{T}[M \circ (H^c \circ \Delta)]\mathbb{1}_n =$$

$$\overset{(a)}{=} tr(\Delta^\mathsf{T}\Delta) + tr[(S^\mathsf{T} - S)\Delta] + l^\mathsf{T}(A + \Delta - \lambda I)v + h\left(v^\mathsf{T}v - 1\right) + q^\mathsf{T}B^\mathsf{T}v + tr[(M \circ H^c)^\mathsf{T}\Delta]$$

50

where (a) follows from properties (i)-(iv). Next, we derive the first-order necessary conditions for a local minimum. Differentiating $\mathcal{L}$ w.r.t. Δ and equating to 0, we get

$$\frac{d}{d\Delta}\mathcal{L} \overset{(vi)}{=} 2\Delta + S - S^{\mathsf{T}} + lv^{\mathsf{T}} + M \circ H^c = 0. \tag{4.7}$$

Taking the Hadamard product of (4.7) with H^c and using $H^c \circ \Delta = 0$ and $H^c \circ H^c = H^c$, we get

$$(S - S^{\mathsf{T}}) \circ H^c + (lv^{\mathsf{T}}) \circ H^c + M \circ H^c = 0. \tag{4.8}$$

Replacing $M \circ H^c$ from (4.8) into (4.7), we get

$$\Delta = -\frac{1}{2}H \circ (S - S^{\mathsf{T}} + lv^T). \tag{4.9}$$

Since H is symmetric, the transpose of (4.9) yields

$$\Delta = \Delta^{\mathsf{T}} \overset{(v)}{=} -\frac{1}{2}H \circ (S^{\mathsf{T}} - S + vl^{\mathsf{T}}). \tag{4.10}$$

Adding (4.9) and (4.10) and using (v), we obtain (4.6a).

Next, we differentiate $\mathcal{L}$ w.r.t. v and equate to 0:

$$(A + \Delta - \lambda I)^{\mathsf{T}}l + 2hv + Bq = 0. \tag{4.11}$$

Pre-multiplying (4.11) by v^{T} and using the eigenvalue, eigenvector and uncontrollability constraints, we get $h = 0$. Then, since A and Δ are symmetric, (4.11) yields (4.6b).

Finally, differentiating $\mathcal{L}$ w.r.t. λ and equating to 0, we get the orthogonality constraint (4.6c).

Equation (4.6d) is the necessary regularity condition and follows from Lemma 16 Next, we obtain the second-order sufficient conditions by deriving the Hessian of $\mathcal{L}$. Recall that $\delta = \text{vec}(\Delta)$. Taking the differential of $\mathcal{L}$ twice, we get

$$d^2\mathcal{L} = 2tr((d\Delta)^{\mathsf{T}}d\Delta) + 2h(dv)^{\mathsf{T}}dv + 2l^{\mathsf{T}}(d\Delta - Id\lambda)dv$$

$$\overset{(b)}{=} 2(d\delta)^{\mathsf{T}}d\delta + 2(dv)^{\mathsf{T}}(I \otimes l^{\mathsf{T}})d\delta - 2d\lambda l^{\mathsf{T}}dv = [(d\delta)^{\mathsf{T}}, (dv)^{\mathsf{T}}, d\lambda]\, D\, [(d\delta)^{\mathsf{T}}, (dv)^{\mathsf{T}}, d\lambda]^{\mathsf{T}},$$
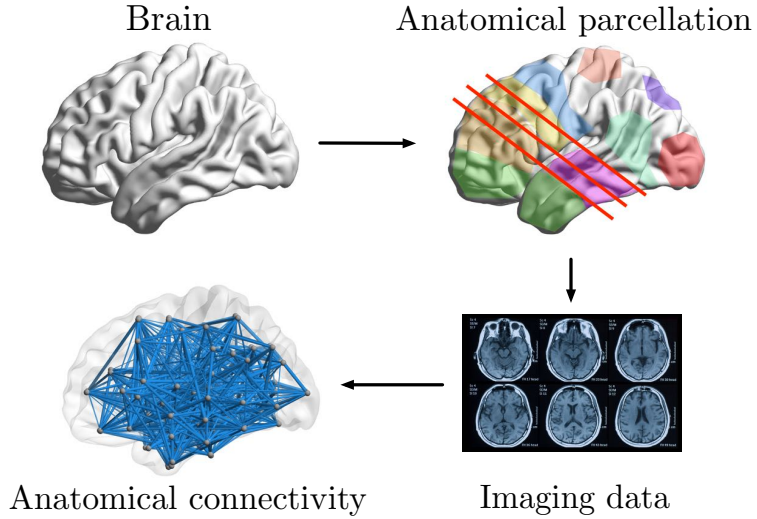
Figure 4.2: Main steps performed to obtain anatomical connectivity matrices.

where (b) follows from properties (ii), (vii) and $h = 0$. The sufficient second-order optimality condition for the optimization problem requires the Hessian matrix to be positive semi-definite in the kernel of the Jacobian at the optimal point [72]. That is, $z^\mathsf{T} D^* z \geq 0$, $\forall z : J(\Delta^*, v^*, \lambda^*)z = 0$. This condition is equivalent to $P^* D^* P^* \geq 0$, since $J(\Delta^*, v^*, \lambda^*)z = 0$ if and only if $z = P^* u$ for any $u \in \mathbb{R}^{n^2+n+1}$ [70]. Since the projection matrix P^* is symmetric, (4.6e) follows and this concludes the proof. ∎

Remark 18 (*Computing an optimal solution*) *Observe that Δ^* in (4.6a) is symmetric (since H is symmetric) and satisfies the structural constraint (4.2e). Thus, to obtain a solution to the minimization problem (4.2), we perform an iterative procedure (starting from some random initial condition) that solves numerically the constraint equations (4.2b)-(4.2d) and the optimality equations (4.6a)-(4.6c). We then verify that these solutions satisfy the regularity and local minima equations (4.6d) and (4.6e), respectively. We repeat this procedure for several initial conditions to improve upon local solutions. However, due to the non-convexity of the minimization problem (4.2), convergence to a global minimum is not guaranteed.* □

Example 19 *(Structured controllability radius of a line network)* *Consider a network with adjacency matrix*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

and input matrix $B = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\mathsf{T}$. *Notice that this is a line network controlled by the first node, which is known to be strongly structurally controllable [73]. We are interested in modifying only the existing edges of the line, that is, $\mathcal{H} = \mathcal{G}$. With this sparsity constraint, the symmetric perturbation with the minimum Frobenius norm (global minimum) that renders the network uncontrollable is the one that removes the edge with the smallest weight [67]:*

$$\Delta_{global}^* = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

To verify our procedure against this result, we solve the constraint and optimality equations in Theorem 17. In addition to the above global minimum (with $v_{global}^* = [0, 0.7071, 0.7071]^\mathsf{T}, l_{global}^* = [5.6569, -1.7038, 1.7038]^\mathsf{T}, \lambda_{global}^* = 2$), we also obtain the following two local minima, where $\lambda_1^* = \lambda_2^* = 0$, and

$$\Delta_1^*, v_1^*, l_1^* = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & -2 & 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 8.1341 \\ 8 \\ 2.5228 \end{bmatrix}, \; and$$

$$\Delta_2^*, v_2^*, l_2^* = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -2 \\ 0 & -2 & 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0.9711 \\ -0.2387 \end{bmatrix}, \begin{bmatrix} 4.1191 \\ 5.1020 \\ 9.4922 \end{bmatrix},$$

which correspond to removing the edges between node 2 and 3, and all the edges of the network, respectively. □

4.4 The controllability radius of symmetric brain and random networks

In the remainder part of the chapter we focus on the numerical analysis and comparison of the controllability radius of brain and random networks. We focus on the case $\mathcal{H} = \mathcal{G}$, and consider the following network models.

Structural brain network (SBN). We use brain networks modeled by (4.1), where the anatomical connectivity matrices represent weighted adjacency matrices. To obtain the connectivity matrices, anatomical scans of 15 healthy subjects were parcellated according to the Lausanne atlas [46], and $n = 129$ regions are chosen as regions of interest. Figure 4.2 shows the main steps performed to obtain the connectivity matrices of structural brain networks. These network dynamics can be derived as a linearization of brain processes [47], and have been used, for instance, in [10, 58, 74]. Among these, [10] has numerically shown that this class of brain networks constructed from diffusion MRI data is controllable from one single node, which will support our assumption of selecting only one control node in our numerical study.

Random network (RN). Starting from a structural brain network, we generate a set of symmetric adjacency matrices by randomly permuting its edges, while maintaining connectivity and controllability from the selected control node.

In our study, we consider 15 SBN's and 10 RN's generated from each SBN, for a total of 165 networks. For each network, the solution to the optimization problem (4.2) is computed numerically. To do so, we run an extensive number of minimizations from random initial conditions. Finally, we compare the controllability radius of each SBN with the mean controllability radius of the 10 corresponding RN's.

We run two sets of numerical studies. First, we compare the structured controllability radius of brain and random networks for a fixed choice of control node, that is, the control node in a brain network and in all its random permutations is the same. Second, we compare brain
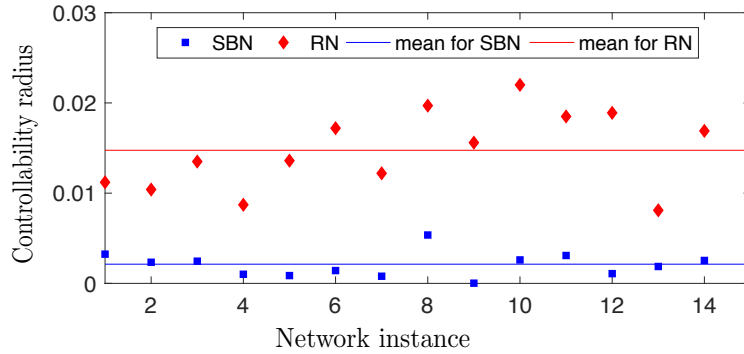
(a)
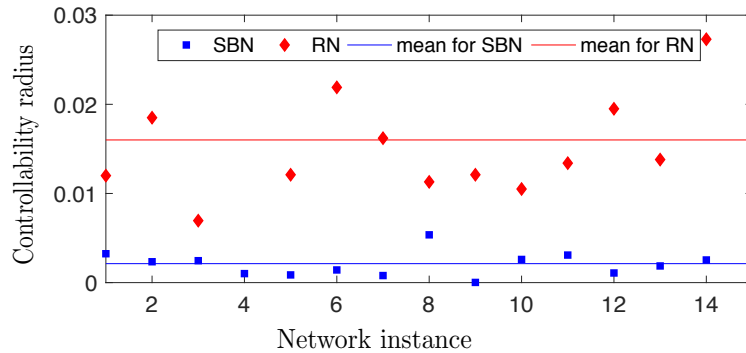
(b)

Figure 4.3: Comparison between the controllability radii of 15 SBN's and 150 RN's. (a) The control node maximizes $\lambda_{\min}(W)$ for the i-th SBN, and the same control node is employed for the 10 randomized brain networks corresponding to the i-th SBN, $i = 1, \ldots, 15$. (b) The control node is selected differently in each network to maximize $\lambda_{\min}(W)$.

and random networks after varying the control node to maximize the smallest eigenvalue of the controllability Gramian. The results of our studies are reported in Fig. 4.3.

In both our sets of numerical studies, the results show that the controllability radius of structural brain networks is on average smaller than the controllability radius of the respective randomized versions. This result suggests that the topology of brain networks may not be accidental. Furthermore, the peculiar architecture of the brain could have evolved to favor dynamic features different from controllability by single regions [75]. This raises several questions, including
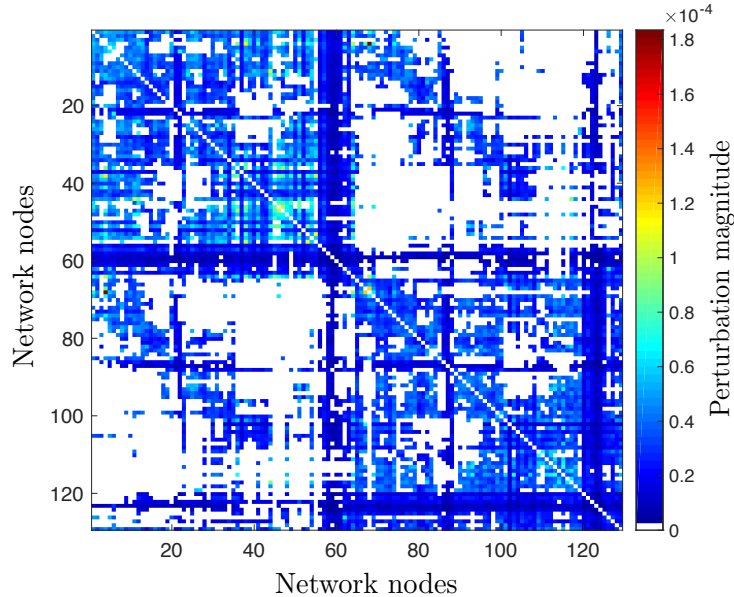
Figure 4.4: Average magnitude of the weight changes among all the perturbations Δ_i for subject 1 when controlled by single brain regions, $i = 1, \ldots, n$.

characterizing the cost functions optimized by the anatomical structure of the brain [76].

The controllability radius of structural brain networks can be further exploited to provide interesting information on the effectiveness of network control from a certain area. In fact, it is possible to understand which areas make a better position for a control node without incurring in numerical artifacts that affect the spectral analysis of the controllability Gramian. For instance, when the brain network of subject 1 is controlled by the *right superiorparietal-3* region[3] (node 30), it displays the smallest controllability radius: $\|\Delta\|_F \approx 10^{-7}$. When the brain network of subject one is controlled by the *left isthmus of cingulate* (node 88), it displays the largest controllability radius: $\|\Delta\|_F \approx 2.5 \cdot 10^{-2}$. This insight could inform the design of novel brain stimulation techniques.

To conclude, we noticed that the most perturbed interconnections correspond to entries with a small magnitude in the connectivity matrix. Furthermore, when comparing the mean perturbation for subject 1 (see Fig. 4.4) with the brain atlas, the areas where the most changes occur among all the perturbation matrices turn out to be the *frontal pole* (node 4 and 68) and the *pars*

---

[3]See [46] for the atlas with labels of the regions of interest in the brain.

orbitalis (node 67). These two areas are indeed important for cortico-cortical control [10]. Assessing which interconnections tend to be more fragile toward uncontrollability of the brain and further enhancing the role that network controllability plays in the correct functioning of this complex organ opens new challenges for future research and may ultimately lead to the development of innovative personalized clinical therapies [54, 75].

4.5   Conclusion

In this chapter we propose and analyze a novel notion of controllability for symmetric networks, namely the structured controllability radius. In particular, we quantify the controllability degree of a network with the smallest norm of a symmetric perturbation that renders the network uncontrollable and satisfies a pre-specified set of sparsity constraints. We derive a set of equations for the computation of the structured controllability radius, and illustrate our results through various examples including networks approximating a class of brain dynamics. Our numerical results show that brain networks feature a controllability radius that is consistently smaller than the one of random networks with similar weights, further highlighting the fundamental role of the organization of the brain for its dynamic functions.

Part II

Reverse Engineering

Synchronization of Brain Network

Dynamics

Chapter 5

Brain Network Dynamics Fingerprints are Resilient to Data Heterogeneity

In this chapter, we focus on validating the estimation of individual brain network dynamics fingerprints and appraise sources of variability in large resting-state functional magnetic resonance imaging (rs-fMRI) datasets by providing a novel point of view based on data-driven dynamical models. We utilize hidden Markov models (HMM) to examine how diverse scanning factors in multi-site fMRI recordings affect our ability to infer the brain's spatiotemporal wandering between large-scale networks of activity. Building upon this premise, we first replicate previous work on the emergence of non-random sequences of brain states. We next highlight how these time-varying brain activity patterns are robust subject-specific fingerprints. We refer the reader to [77] for a comprehensive discussion.

5.1 Introduction

Untangling the brain's dynamics at rest is a central aspect in the quest to reveal the mechanisms underlying the spontaneous wandering of the mind between well-established, large-scale networks of neural activity [78–80]. The characterization of the brain dynamics' spatiotemporal organization into networks has greatly benefited from the creation of very large neuroimaging datasets [81,82], such as the Human Connectome Project (HCP) [83,84], the UK Biobank [85], and, in the context of neurodegenerative diseases, the Alzheimer's Disease Neuroimaging Initiative [86]. Large neuroimaging datasets have, furthermore, played a crucial role in the development of novel biomarkers for psychiatric and neurodegenerative disorders [87–89]. Yet, appraising how differences in physical parameters or scanning protocols affect the quality of these data – especially fMRI recordings – remains an outstanding problem [90–93]. For instance, imaging sequences are considerably affected by site-dependent differences such as scanner drift over time, or maintenance routine [93]. Only few recent works have addressed the problem of data variability in rs-fMRI data across sites [94–97], while some others have proposed techniques to harmonize multi-site data [87,89,91–93,98,99]. Despite growing interest in the intricacies inherent to multi-site data, this line of research is still in its infancy (the first publication appeared in 2013 [100]). Furthermore, although the brain is a complex dynamical system capable of exhibiting rich nonlinear dynamics [101,102], most studies to date have relied on static measures (e.g., functional connectivity); and little to no attempts exist at exploring such issues from the viewpoint of dynamical models.

Data-driven dynamical models are a promising and powerful tool for the analysis and prediction of the spatiotemporal organization of brain activity [103–106]. These models allow us to harness the vast amount of spurious information contained in large datasets [107–109], capture the hierarchical organization of brain activity [110], enhance brain-computer interfaces [111, 112], and may even be employed in clinical settings [87,113–115]. However, how the inference and identification of dynamical models is affected by different factors in multi-site data acquisition has yet to be

investigated. Additionally, dynamical models could provide fine-grained insight into the extent of the effect of these factors on the data.

One limitation of data-driven models is that, generally, large amounts of data are needed to train the model in the first place. Here, we avoid this issue by employing two datasets. We leverage the high number of subjects ($n_{\text{HCP}} > 1000$) with rs-fMRI data available in the HCP dataset [83], to train a stable and reliable Hidden Markov Model (HMM). An HMM infers brain network dynamics from rs-fMRI time series, where networks are probability distributions representing graphs. We then apply the pretrained HMM to the smaller ($n_{\text{TS}} = 9$) Traveling-subject dataset, which consists of a novel, state-of-the-art collection of rs-fMRI measurements of nine healthy subjects who traveled to twelve different sites and were scanned under various conditions (different sites, days, phase encoding, number of channels/coils, manufacturer, scanner; see Materials and Methods and Supplementary Table 5.2 for a full list of scanning factors and attributes) [99]. This way, we were able to infer subject-specific brain states and investigate how the retrieval of brain state time courses is affected by an array of scanning factors. Training the model on the HCP data guarantees that (1) the model is inferred on a large sample, made of carefully collected and homogeneous data and that (2) the model is stable and does not over-fit on a dataset of limited size. We illustrate the methodological approach in Fig. 5.1.

Thus, we first utilize the trained HMM to validate the findings on rs-fMRI fingerprints – robust and reproducible quantitative signatures – reported in previous work [105, 110, 116]. We then generalize these findings by applying the HCP-trained HMM to the Traveling-subject dataset. This important step allows us to exploit the HMM to assess if, and to what extent, mixed scanning factors affect subject-specific fingerprints and, thus, rs-fMRI recordings. We depart from previous work, which has mostly relied on static functional connectivity / correlation measures and smaller datasets, by exploiting dynamical brain network collective states at a finer temporal resolution. Altogether, this chapter juxtapose complementary, yet contrasting, results with respect to rs-fMRI
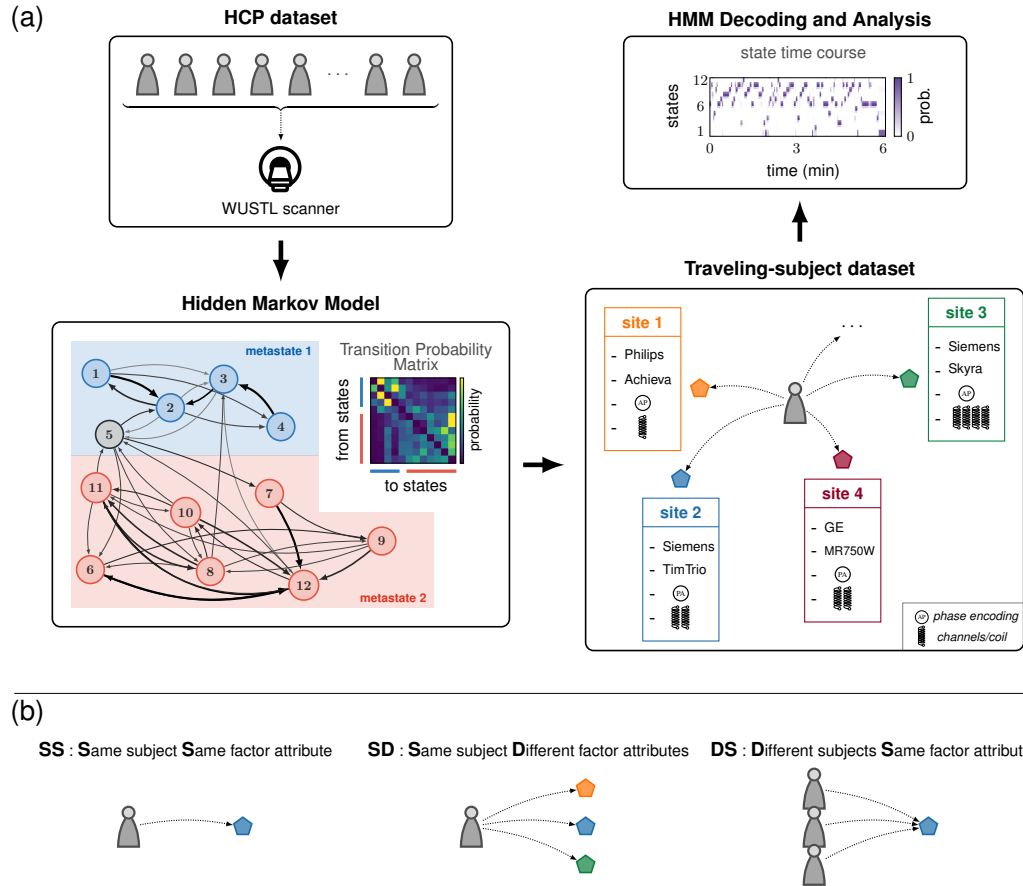
Figure 5.1: Conceptual flow of the analysis and modeling approach. (a) rs-fMRI data from the HCP dataset, collected at the Washington University in St. Louis (WUSTL) Connectome-Skyra scanner, were used to infer a Hidden Markov Model (HMM). This model is described by a transition probability matrix, which encodes the probabilities of jumping from one state to another at each time step. Following [110], 12 states were identified and the graph depicted in the figure illustrates the largest transition probabilities (> 0.1) in our model (see also Supplementary Figure 5.6). The states are color-coded in order to distinguish which set of strongly connected states (*metastate*) they belong to. HMM decoding was then applied to the Traveling-subject dataset, in which rs-fMRI data was collected from subjects travelling to different sites. The state time courses from the Traveling-subject dataset were finally used to (1) validate the subject-specific fingerprints associated with states' dwelling probabilities and the 2-metastate structure put forth previously [110], and (2) analyze the impact of different factors, e.g., site, or scanner model, on fMRI measurements. (b) To gauge how different factors influenced fMRI data collection, the state time courses obtained from the HMM decoding procedure were compared within and across three different groups: SS (Same subject Same factor attribute), SD (Same subject Different factor attributes), and DS (Different subjects Same factor attribute). In this panel, these three categories are illustrated for the factor "site", whose attributes consist of the different geographical locations.

data analysis: we confirm previous findings reporting subject-specific fingerprints, but we also shed light on the presence of factors that induce variability in such fingerprints and, thus, the homogeneity of multi-site fMRI data collections and subsequent inference from the viewpoint of dynamical models.

5.2   Materials and Methods

5.2.1   Datasets

The two dataset used in this study are (1) the HCP 1200-subject distribution (data available at `https://db.humanconnectome.org`) and (2) the Traveling-subject dataset (data available at `https://bicr-resource.atr.jp/srpbsts/` after free registration). The former consists of rs-fMRI data from $N = 1206$ healthy subjects (age 22-35) that were scanned twice (two 15-minute runs) on two different days, one week apart, on a Siemens 3T Connectome-Skyra scanner. For each subject, in total four 15-minute runs of rs-fMRI time series data with a temporal resolution of 0.72 s and a spatial resolution of 2-mm isotropic were available. For our analysis, we used time series from the 1003 subjects with 4 complete scanning sessions. The HCP dataset provides the required ethics and consent needed for study and dissemination, such that no further institutional review board (IRB) approval is required.

The Traveling-subject dataset consists of 9 healthy subjects (all men; age range 24–32; mean age 27 ± 2.6y), who were all scanned at each of the 12 sites, producing a total of 411 10-minute scanning sessions [99]. Each participant underwent three rs-fMRI sessions of 10 min each at nine sites, two sessions of 10 min each at two sites (HKH and HUH), and five cycles (morning, afternoon, following day, following week, and following month) consisting of three 10-min sessions each at a single site (ATT). In the latter situation, one participant underwent four rather than five sessions at the ATT site because of a poor physical condition. Thus, a total of 411 sessions were conducted [8 participants $\times$ ($3 \times 9 + 2 \times 2 + 5 \times 3 \times 1$) + 1 participant $\times$ ($3 \times 9 + 2 \times 2 + 4 \times 3 \times 1$)] (see Supplementary Table 5.2 for all the details on the scanning protocols). In total, there were two

phase-encoding directions (posterior to anterior [P → A] and anterior to posterior [A → P]), three MRI manufacturers (Siemens, GE, and Philips), four numbers of channels per coil (8, 12, 24, and 32), and seven scanner types (TimTrio, Verio, Skyra, Spectra, MR750W, SignaHDxt, and Achieva). All participants in all datasets provided written informed consent. All recruitment procedures and experimental protocols were approved by the institutional review boards of the principal investigators' respective institutions (Advanced Telecommunications Research Institute International [ATR] [approval numbers: 13–133, 14–133, 15–133, 16–133, 17–133, and 18–133], Hiroshima University [E-38], Kyoto Prefectural University of Medicine [KPM] [RBMR-C- 1098], SWA [B-2014-019 and UMIN000016134], the University of Tokyo [UTO] Faculty of Medicine [3150], Kyoto University [C809 and R0027], and Yamaguchi University [H23-153 and H25-85]) and conducted in accordance with the Declaration of Helsinki.

5.2.2   Hidden Markov Model

In this chapter, we utilized Hidden Markov model(s) to capture the dynamical evolution of brain states in subjects scanned at rest. In neuroscience and neuroimaging, HMMs are typically used to represent the stochastic relationship between a finite number of hidden states that underlie the brain's complex dynamics, whose evolution in time is captured by the measured data. That is, Hidden Markov modeling is a powerful technique that enables the description of time series extracted from a system of interest. The underlying assumption of this class of models is that the observed time series of data can be explained by a discrete sequence of hidden states, which must be finite in number. Additionally, to describe a hidden Markov model, an observation model needs to be chosen. We assume multivariate Gaussian observation model, so that, if $\mathbf{x}_t$ denotes the data at time step t, and s_t represents the state at time step t, we can write, whenever state k is active,

$$\mathbf{x}_t | s_t \sim \text{multivariate Gaussian}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

64

where $\boldsymbol{\mu}_k \in \mathbb{R}^c$ is the vector of the mean blood oxygen level-dependent (BOLD) activation for each channel, with c being the number of channels in the data, and $\boldsymbol{\Sigma}_k \in \mathbb{R}^{c \times c}$ is the covariance matrix encoding the variances and covariances between channels. The transitions between different brain states depend on which state is active at the previous time step. Specifically, the probability of a state being active at time t depends on which state is active at time step $t-1$. This is encoded in the Transition Probability Matrix $\boldsymbol{\Theta}$, in which the entry $\boldsymbol{\Theta}_{ij}$ – the transition probability – denotes the probability of state i becoming active at the next time step if state j is currently active. Formally, by denoting a probability with Pr, it holds that

$$\Pr(s_t = i) = \sum_j \boldsymbol{\Theta}_{ij} \Pr(s_{t-1} = j)$$

For large datasets, it is possible to resort to stochastic Variational Bayes inference to estimate the posterior distribution of each state $(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, the probability of each state being active at each time step, and the transition probabilities between each pair of states $\boldsymbol{\Theta}_{ij}$ [108]. Finally, notwithstanding the fact that in this study the model has been inferred by concatenating all the subjects – thus implicitly defining the brain states as the outcome of common brain dynamics – the state time courses are subject-specific. That is, the states are inferred at the group level, but the time instants at which each brain state becomes active is subjective and changes between and across subjects.

5.2.3 Data Preparation and HMM Training

HCP dataset. Following [110], extensively preprocessed HCP ICA time series were used for the model training. The preprocessing followed the steps of [83, 117] and is briefly described here. Spatial preprocessing used the procedure described by [118]. Next, structured artifact removal using ICA was followed by FMRIB's ICA-based X-noisefier (FIX) from the FMRIB Software Library (FSL) [119], which removed more than 99% of the artifactual ICA components in the dataset. Finally, the 50-dimensional extensively preprocessed time series obtained after group spatial ICA are freely

available at `https://www.humanconnectome.org/study/hcp-young-adult/`.

Traveling-subject dataset. The dataset was obtained from `https://bicr-resource.atr.jp/srpbsts/`. Hereafter, we describe the preprocessing procedure that was originally reported in [99]. Raw BOLD signals were preprocessed using SPM8, implemented in MATLAB (R2016b; Mathworks, Natick, MA, USA), The first 10 s of each scan data were discarded to account for T1 calibration. Ensuing preprocessing steps included: slice-timing correction, realignment, coregistration, segmentation of T1-weighted structural images, normalization to Montreal Neurological Institute (MNI) space, and spatial smoothing with an isotropic Gaussian kernel of 6 mm full-width at half-maximum. Thirty-six noise parameters were included in a linear regression model to remove multiple sources of spurious variance (e.g., six motion parameters, average signals over the whole brain, white matter, and cerebrospinal fluid) [120]. Time-series were band-pass filtered using a first-order Butterworth filter [0.01 - 0.08 Hz] to restrict the analysis to low-frequency fluctuations, which are characteristic of rs-fMRI BOLD activity [120]. Finally, to reduce the impact of head motion, scrubbing was performed: framewise displacement (FD) was calculated and volumes with FD > 0.5 mm were removed [121]. Thus, $5.4\% \pm 10.6\%$ volumes (mean [approximately 13 volumes] $\pm$ 1 SD) were removed per 10 min of rs-fMRI scanning (240 volumes). If the number of volumes removed after scrubbing exceeded the average of –3 SD across participants, the sessions were excluded from the analysis. As a result, 14 sessions were removed from the dataset.

Before combining the HCP time series and the Traveling-subject time series for the model inference, we matched the temporal resolution of the two datasets. Specifically, for all results reported in the main text, the Traveling-subject time series were up-sampled in order to match the same repetition time as the HCP data (from TR = 2.5 s to TR = 0.72 s). We also down-sampled the HCP data from TR = 0.72 s to TR = 2.5 s to match the Traveling-subject repetition time. However, the model inferred on HCP down-sampled ICA time-series was not satisfactory (see below). Therefore, we have chosen to re-sample the Traveling-subject data instead of down-sampling the HCP data.

The HMM inference was performed on 50-dimensional standardized ICA time series (0 mean and unitary standard deviation) concatenated along the time direction. To concatenate HCP rs-fMRI data and the ones from the Traveling-subject dataset, we proceeded as follows. First, we matched the voxel coordinates of the Traveling-subject data with the group average spatial maps from the group-ICA decomposition of the HCP time series. These spatial maps were extracted from the group average analysis across all the subjects of the S1200 release and are available on the HCP website: `https://www.humanconnectome.org/`. Because the spatial maps are in a gray-ordinate CIFTI format [118], we extracted the xyz coordinates in a standard stereotaxic space MNI152 by using a *mid-thickness* surface file for the surface vertices and the coordinate transformation matrix included in the CIFTI file. Next, we extracted the time series from the Traveling-subject data corresponding to the same xyz coordinates of the aforementioned spatial map in Matlab by using the ROI Signal Extractor provided by the toolbox DPABI [122]. Finally, the HCP group average spatial map allowed us to obtain the estimated 50-dimensional ICs for the Traveling-subject data from the extracted time series. To note, once the Traveling-subject rs-fMRI time series were reduced to 50 ICs, they matched the spatial dimension of the HCP data used to infer the HMM in [110]. Finally, to train our HMM, we used the publicly available toolbox `HMM-MAR` (`https://github.com/OHBA-analysis/HMM-MAR`) [123]. We inferred $N = 50$ models with 12 states (the number of states was chosen based on previous work [110]) from random initializations, multiple priors, and different combinations of the available datasets. Specifically, we inferred $N_{1200} = 28$ models inferred on time series from the 1200-subject HCP release only with random priors, $N_{820} = 14$ models inferred on the 820-subject HCP release (a subset of the 1200-subject release, which was used in the original work on the HMM-derived hierarchical organization of brain states [110]) with random priors, and $N_{TS} = 8$ models inferred on the 9 subjects of the Traveling-Subject dataset using the best model inferred from HCP data only one as a prior, so that $N_{1200} + N_{820} + N_{\text{TS}} = 50$). The selection of the best model (described below in detail) took into account both classical model evaluation methods and the definition of the metrics used in this study.

5.2.4   FO Correlation Matrix and Fingerprints Computation

By applying (i.e., decoding) an HMM to a dataset with multiple subjects, we obtained the state time courses for each subject, from which it is possible to compute the vector of the Fractional Occupancy (FO) of every state for each subject. Stacking such vectors in a matrix yielded the FO Matrix $\mathbf{R}$, which is a *(no. of subjects) $\times$ (no. of states)* matrix that encodes state dynamics similarities across subjects. Each element $\mathbf{R}_{ij}$ of this matrix denotes the fraction of time spent by subject i in state j. Further, by taking the pairwise correlation of the columns of the FO Matrix $\mathbf{R}$, we obtained the *(no. of states) $\times$ (no. of states)* FO Correlation Matrix

$$\mathbf{C} = \mathrm{corr}(\mathbf{R}_{:,k}, \mathbf{R}_{:,\ell}),$$

where $\mathbf{R}_{:,k}$ denotes the column vector of the FO of all subjects for the k-th state. This matrix captured the overall organization of brain dynamics across states, and its entries quantified the affinity between the FOs of each pair of states across all subjects. In other words, the FO Correlation Matrix highlighted the similarities and dissimilarities between brain states, and encoded the temporal characteristics of brain network dynamics. The organization of the FO Correlation matrix revealed (both by visual inspection and by numerical investigation) the emergence of two groups of states, known as metastates. Metastates are distinct sets of functional network states that the brain has a propensity to cycle within, and have been shown to hierarchically group brain states into a 2-metastate structure [110].

We made use of the information encoded in the FO Correlation Matrix to calculate two different subject-specific metrics in the Traveling-subject data that were key in this study: the Metastate Profile (MP) Differences and the Fractional Occupancy (FO) Correlations. Loosely speaking, the former provided the difference between the time spent in the two distinct metastates that emerged in our model, compatibly with previous findings [110]. The latter provided the pairwise correlation between the FO vectors of different scanning runs. To derive these metrics, we first construct the MP matrix, whose entry (i, k) represents the FO of the second metastate (states 6 to 12) minus the

FO of the first metastate (states 1 to 4) for the subject i during the scanning session k. We excluded state 5 from our analysis as it was uncorrelated from the other states, had the highest variance, and was previously shown to be associated with head motion in the scanner [110]. Formally, given the FO Matrix $\mathbf{R}$ for the run k, $\mathrm{MP}_{i,k}$ is computed as follows:

$$\mathrm{MP}_{i,k} = \sum_{j=6}^{12} \mathbf{R}_{i,j} - \sum_{j=1}^{4} \mathbf{R}_{i,j}.$$

Then, the MP Difference between run k_1 for subject i_1 and run k_2 for subject i_2 reads as

$$\mathrm{MP\ Difference} = |\mathrm{MP}_{i_1,k_1} - \mathrm{MP}_{i_2,k_2}|.$$

Instead, the FO Correlation between run k_1 for subject i_1 and run k_2 for subject i_2 is defined as

$$\mathrm{FO\ Correlation} = \mathrm{corr}(\mathbf{FO}_{i_1,k_1}, \mathbf{FO}_{i_2,k_2}),$$

where $\mathbf{FO}_{i,k}$ denotes the 11-dimensional column vector of the FOs of all 12 states minus state 5 for subject i and run k.

It is worth noting that exploiting and comparing the two metrics defined above gave us a remarkable advantage with respect to utilizing only the model's TPM. Namely, because of the stochastic nature of the model inference, we were able to avoid the non-uniqueness issue of the TPM and, at the same time, to reliably capture the temporal characteristics of the state time courses. In fact, due to the availability of numerous scanning sessions for each subject, both metrics could be computed not only across different subjects, but also at the individual level. We capitalized on the robustness of the HMM model inferred on HCP homogeneous data to reveal, through MP Differences and FO Correlations, temporal information of brain state time series in the heterogeneous Traveling-subject dataset. These metrics allowed us to perform a richer analysis rather than simply limiting ourselves to the study of a model's TPM – in this context it was one single matrix valid for all subjects (Fig. 5.1(a)).

5.2.5 Model Selection

To select the model that best fit the data, we computed the free energy for each of the fifty different models that were inferred. The free energy provides a bound on the log-evidence for any model [124], and can be derived as the sum of the model average log-likelihood, the negative entropy and the Kullback–Leibler divergence [125]. Because the data sets have different sizes (HCP_{1200}, HCP_{820}, and Traveling-subject only), we corrected the free energy according to the size of the dataset used for the model inference in order to compare different models fairly. Next, we ranked the $N = 50$ models inferred in this study based on their free energy, and chose the one minimizing this quantity.

Based on previous findings [110], and because the definition of MP Difference strongly rely on the hierarchical structure of the inferred states, we also verified that the selected model presented a sufficiently marked 2-metastate structure. To take this topological notion into account, we computed for each model the Euclidean distance from the ideal FO Correlation Matrix (Supplementary Figure 5.7), which gauged how well the metastates emerged in the model's FO Correlation Matrix. Mathematically, this distance is defined as:

$$d_i = \left\| \mathbf{C}_i - \begin{bmatrix} \mathbf{1}_{4\times4} & \mathbf{0}_{4\times1} & -\mathbf{1}_{4\times7} \\ \mathbf{0}_{1\times4} & 1 & \mathbf{0}_{1\times7} \\ -\mathbf{1}_{7\times4} & \mathbf{0}_{7\times1} & \mathbf{1}_{7\times7} \end{bmatrix} \right\|$$

for $i = 1, \ldots, 50$, where $\mathbf{C}_i$ is the FO Correlation Matrix of model i, $\mathbf{1}$ is a matrix of all ones, $\mathbf{0}$ is a zero matrix, and $\| \cdot \|$ denotes the Euclidean norm. The model that we have used in this study was not only the one with the lowest free energy, but also the one with and the smallest d_i. Thus, our model fit the data the best and simultaneously embodied a pronounced 2-metastate structure.

Finally, to verify the robustness of our model when applied to time series other than the HCP data, we applied the HCP-trained HMM to autoregressive data (Appendix and Supplementary Figure 5.8). We found that this control analysis yielded state time courses that spend most of the time on state 5. Moreover, state 5 was not only uncorrelated to all other states in our model, but

had previously been found to be associated to motion artifacts in HCP data [110]. This result substantiates the robustness of our results in regards to application of our HCP-trained model to the Traveling-subject time series.

5.2.6   Subject Classification Using Brain Dynamics Fingerprints

To support our findings, and the robustness of the subject-specific fingerprints to data heterogeneity, we used machine learning on these fingerprints to perform subject-level classification. Specifically, individual subjects from the Traveling-subject dataset were classified based on their Metastate Profiles and Fractional Occupancies. We detail the procedure hereafter.

For each scanning factor, we trained a logistic regression classifier – which minimizes the cross-entropy loss – with the `scikit-learn` machine learning package [126] in Python 3 with the following parameters: default L2 penalty, default L-BFGS-B algorithm [127], and 'multi_class' option set to 'multinomial'. The classification task was repeated multiple times by splitting the data into different training and validation sets as follows. We repeated the training and validation of the linear regression classifier for each factor attribute (e.g., for the scanner parameter, we repeated the procedure for each scanner model) by performing a leave-one-attribute-out cross-validation: we chose as validation set all the samples (i.e. fingerprints) belonging to one factor attribute, and we used as training set all the remaining samples. This analysis allowed us to (1) compare the classification based on brain dynamics fingerprints in the presence of different scanning protocols and heterogeneous data with baseline chance level, and (2) investigate which scanning factors tend to affect data collections more than others.
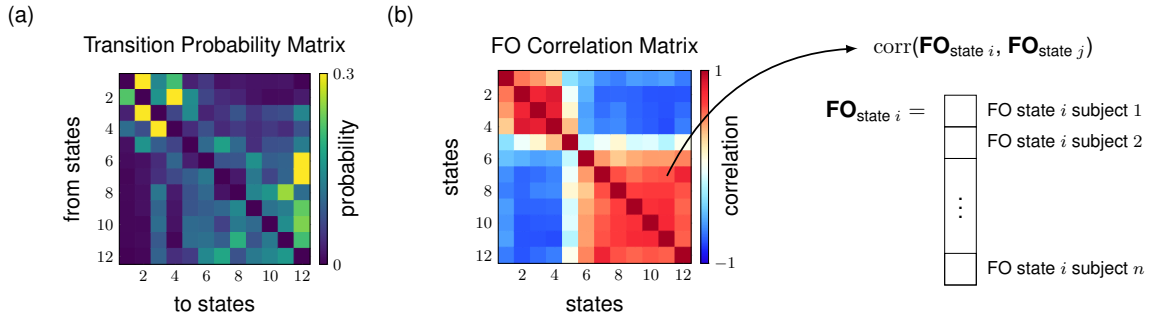
Figure 5.2: HCP-trained Hidden Markov Model. (a) Transition Probability Matrix. The emergence of the two metastates can be recognized by simple visual inspection, and was confirmed by a community-detection algorithm. (b) FO Correlation Matrix. The two metastates are clearly delineated, with state 5 being mostly uncorrelated from all other states [110]. The state FOs are highly correlated (Person correlation > 0.8) within the two metastates across subjects.

5.3 Results

5.3.1 Test-retest reliability of brain dynamics estimation

We first inferred the HMM by leveraging the large amount of rs-fMRI data in the HCP dataset. Due to the stochastic nature of the HMM inference – which is based on the probabilistic process of Bayesian inference – the results might vary at each new model training. Thus, we inferred multiple models and selected for further analyses the one with the best fit. We show in Fig. 5.2 the HMM selected and employed in this chapter, which is the model that ranked best with respect to free energy, displayed the smallest distance from the ideal FO Correlation Matrix, and was trained solely on HCP time series (see also Supplementary Figure 5.9). Further details and matrices of notable models different from the best one can be found in Supplementary Figure 5.10-5.11.

Given the stochastic nature of the Variational Bayes approach used to infer the HMM [108], it was unlikely that one would obtain an exact replica of the model originally reported in [110]. However, as displayed in Fig. 5.2(b) and Supplementary Figure 5.11, all models showed a clear 2-metastate structure, validating the claims that resting-state brain dynamics tend to be hierarchically organized in two larger sets of states (one associated with higher-order cognition, and the other

one with sensorimotor and perceptual states, as originally reported in [110]). Moreover, a visual inspection of the TPM matrix alone suggested the emergence of two groups of states that tended to be more (statistically) connected. We confirmed this hypothesis by employing the generalized Louvain algorithm [128] for the discovery of communities in networks.

To note, we also used the HCP-derived TPM as a prior to train an HMM on the Traveling-subject dataset alone. This choice of prior ensured that the inference started from established initial conditions before dealing with the small size of the Traveling-subject dataset. Surprisingly, although the number of subjects in the Traveling-subject dataset was much smaller than the number of subjects in the HCP dataset, the 2-metastate structure still emerged in the model's matrices (Supplementary Figure 5.11(d)), as also confirmed by the generalized Louvain algorithm. This result highlighted that, notwithstanding mixed scanning protocols and small sample, the metastates could be retrieved and unfold as a robust feature of resting-state data.

5.3.2 Metastate Profiles and Fractional Occupancies Are Robust Subject-Specific Fingerprints

Previous findings reported that brain dynamics is subject-specific and nonrandom. To extend this notion, we applied the best-fitting, HCP-trained HMM, to the Traveling-subject dataset, obtaining the state time courses for each 10-minute scanning session. Next, from each individual's state time courses, we calculated the MP Differences and the FO Correlations. We summarize the derivation of these two measures in Fig. 5.3(a). To note, here we use the notation subject i_1 and i_2 for a general case, but this naturally applies to two different scans belonging to the same subject (i.e., within-subject comparison).

Before delving into the main analyses of the Traveling-subject dataset, we considered the consistency of these two measures of brain activity dynamics both in the HCP and the Traveling-subjects datasets. In both datasets, there were multiple scans per subject ($m_{\text{HCP}} = 4$ and $m_{\text{TS}} > 42$,
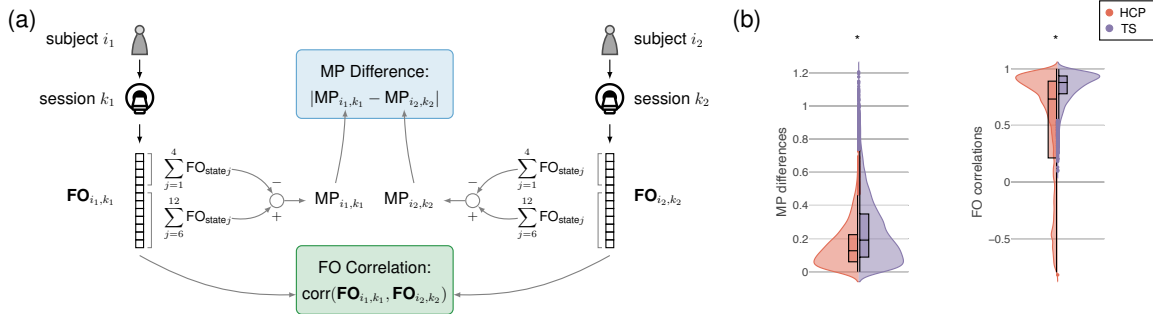
Figure 5.3: Metastate Profile Differences and FO Correlations computation, and within-subject comparison. (a) Schematic illustrating the computation of the MP Differences and FO Correlations. To note, subject i_1 and i_2 can mean both the same subject's data but from different scans, or different subjects. (b) Within-subject MP Differences and FO Correlations in the HCP (in red) and the Traveling-subject (TS, in purple) datasets. We report hereafter the median and interquartile range. For MP Differences (left panel) in the HCP data: median = 0.13 [0.06 0.22]; for MP Differences in the TS data: median = 0.19 [0.09 0.35]. For the FO Correlations (right panel) in the HCP data: median = 0.74 [0.22 0.89]; for FO Correlations in the TS data median = 0.88 [0.78 0.94].

respectively), allowing us to compute MP Differences and FO Correlations within subjects. Given the high homogeneity of the HCP dataset, we expected this to provide a lower bound in terms of dissimilarity between scans belonging to a given subject. Notwithstanding inherent differences (2-sample Kolmogorov-Smirnov test, $k = 0.19$ and $p < 10^{-3}$ for MP Differences, $k = 0.329$ and $p < 10^{-3}$ for FO Correlations), both MP Differences and FO Correlations distributions displayed remarkable similarity in the distributions of MP Differences (peak = 0.06 for HCP data and peak = 0.06 for TS data, Fig. 5.3(b) left plot) and FO correlations (peak = 0.9 for HCP data and peak = 0.93 for TS data, Fig. 5.3(b) right plot). Moreover, the interquartile range also had large overlap, particularly for MP Differences (Fig. 5.3(b) legend). These results provided initial evidence for the presence of – and our ability to infer – subject-specific brain dynamics patterns.

We next interrogated in detail the Traveling-subject dataset. Each scanning factor considered in this study had multiple distinct attributes. For instance, for the factor *scanner manufacturer* there were sessions recorded through scanners produced by three different manufacturers (Siemens, Philips, and General Electric, see also Supplementary Table 5.2). We computed the values of MP

Differences and FO Correlations for all the runs of the same subject and the same factor attribute (SS), the same subject and different factor attributes (SD), and different subjects but the same factor attribute (DS). A 1-way ANOVA on the median MP Differences (Fig. 5.4(a)), and on the median FO Correlations (Fig. 5.4(b)), resulted in a highly significant main effect of comparison group (SS, SD, DS), for both measures (MP differences: $F_{2,15} = 7.64$, $p = 0.005$; FO correlations: $F_{2,15} = 19.76$, $p < 10^{-3}$). Applying post-hoc comparisons, we found that, on average, the median MP Differences for the same subject within the same factor (SS) were significantly lower than the median MP Differences for different subjects within the same factor (DS), ($\sim 38\%$ lower, 2-sided t-test, $t_{10} = -3.59$, $p = 0.005$, Fig. 5.4(a)). Analogously, on average, the median FO Correlations within the same factor for the same subject were higher than across different subjects ($\sim 10\%$ higher, 2-sided t-test, $t_{10} = 8.15$, $p < 10^{-3}$, Fig. 5.4(b)). Additional evidence for how the state time courses of a given subject (within the same factor attributes) tended to be particularly similar was also evident in the FO Correlations median values of the group SS being significantly higher than the median FO Correlations in the groups SD and DS (Fig. 5.4(b)). These findings support the hypothesis that MP Differences and FO Correlations are robust subject-specific measures, as they were resilient to the single effect of all the factors considered in this study.

To further substantiate these results, we used a simple machine learning approach to predict individuals based on their brain dynamics fingerprints. We applied logistic regression to classify the individuals in the Traveling-subject dataset by a leave-one-attribute-out cross-validation procedure (Materials and Methods). In brief, for each factor, we repeated the training and validation of the classifier as many times as the number of factor attributes, using each time the samples of one left-out factor attribute as validation set and the remaining samples from all other attributes as training set. We found the accuracy of the classification to be consistently well above the theoretical chance level (9 subjects: $1/9 \approx 0.11$), scoring on average 0.22 ± 0.02 for the classification based on MPs (a single value for each factor attribute) (t-test against chance level, $t_5 = 16.4$, $p < 10^{-4}$),
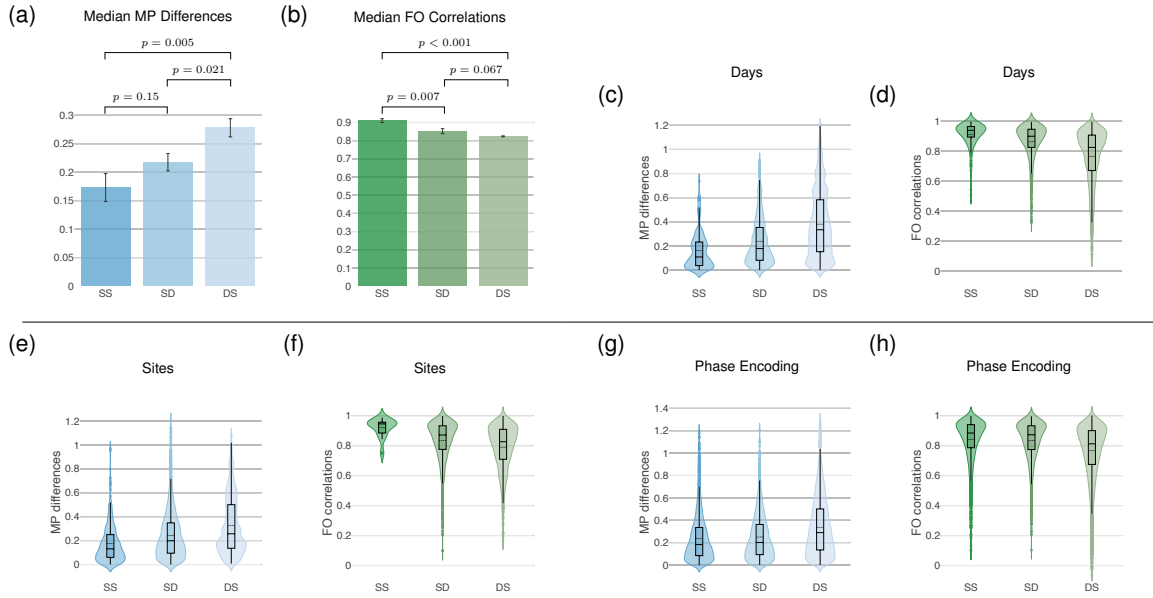
Figure 5.4: Metastate Profile Differences and FO Correlations within vs between subjects, across scanning factors. (a)-(b) The average median of the MP Differences and FO Correlations for the three sets SS (Same subject Same factor attribute), SD (Same subject Different factor attributes), and DS (Different subjects Same factor attribute). MP Differences are the absolute difference between the Metastate Profiles of different runs, while FO Correlations are the pairwise correlation between the Fractional Occupancy vectors of different runs. The set SS consistently displays lower MP Differences and higher FO Correlations than the set DS, confirming the fact that such metrics are subject-specific. The fact that the set SD lies between SS and DS suggests that some scanning parameters influence the aforementioned metrics for resting-state scans of the same subject, but not as much as inter-individual differences. Bars represent the median, error bars the SEM. Statistical comparisons were performed with 2-sided t-tests. (c)-(h) Distributions of values for both metrics and all subjects pooled. The set SS comprises the MP Differences (resp., FO Correlations) computed for each subject within the same factor attribute (e.g., for 'Days', day 1), and the SS distribution displays these values for all subjects; the set SD consists of the MP Differences (resp., FO Correlations) computed for each subject across different attributes of the same factor (e.g., all possible combinations within 'Days'), and the SD distribution displays these values for all subjects; finally, the set DS consists of the MP Differences (resp., FO Correlations) computed across all subjects within the same factor attribute, and the DS distribution displays these values for all attributes of the same factor. For all the distributions, the black dashed lines illustrate the mean. In panels (g) and (h) the difference between SS and SD distributions was not significant (Table 5.1).

0.30 ± 0.03 for the classification based on FOs (a length-11 vector for each factor attribute) (t-test against chance level, $t_5 = 17.62$, $p < 10^{-3}$), and 0.28 ± 0.02 when using the combined measures (t-test against chance level, $t_5 = 21.56$, $p < 10^{-3}$). We report the classification results for each factor in see Supplementary Table 5.4 (see also Supplementary Figure 5.12).

Table 5.1:   2-Sample Kolmogorov-Smirnov test results for MP Differences and FO Correlations. The check-mark indicates that the difference is significant (i.e., the null hypothesis that the samples are drawn from the same underlying continuous population can be rejected at the 5% significance level), and the cross otherwise. All p-values have been FDR-adjusted [129] and they all satisfy $p < 10^{-3}$ when the null hypothesis is rejected. Test statistics are reported in Supplementary Table 5.3. SS: Same subject Same factor attribute. SD: Same subject Different factor attributes. DS: Different subjects Same factor attribute.

	MP Diff.		FO Corr.	
	SS	SD	SS	SD
Factor	vs	vs	vs	vs
	SD	DS	SD	DS
1. Site	✓	✓	✓	✓
2. Day	✓	✓	✓	✓
3. Phase	✗	✓	✗	✓
4. Channels/Coil	✓	✓	✓	✓
5. Manufacturer	✓	✓	✓	✓
6. Scanner	✓	✓	✓	✗

5.3.3   In rs-fMRI Data, Not All Factors are Equal

Given that the Traveling-subject dataset contained a considerable number of factors, we inquired which of these factors, and to what extent, influenced the subject-specific fingerprints defined on the HMM state time courses. Specifically, we asked which factors affected the MP Differences and the FO Correlations most, both within and across subjects. Thus, we compared three different groups (SS, SD, and DS, as illustrated in Fig. 5.1(b)) of MP Differences and FO Correlations, for six different factors, each containing at least two attributes (see Supplementary Table 5.2 for the full list of factors and associated attributes).

Although different runs always carried some variability, some factors seemed to influence the MP Differences and the FO Correlations more than inherent inter-subject differences. We summarize the main results of this comparison in Fig. 5.4 and report the additional ones in Supplementary

Figure 5.13. We also report in Table 5.1 the results of Kolmogorov-Smirnov nonparametric tests between all the distributions of values for the groups of MP Differences and FO Correlations. More in detail, by comparing the distributions of values for both metrics between the sets SS (Same subject and Same factor attribute) and SD (Same subject and Different factor attributes), we found them to be statistically different ($p < 10^{-3}$, see Table 5.1) for all factors except for the phase encoding direction, as also noticeable in Fig. 5.4(g)-(h). It is worth noting that the median MP Difference of any given subject displayed only small changes in the comparison within attribute vs. between attributes for all factors (2-sided t-test, $t_{10} = -1.55$, $p = 0.15$); compatibly, the median FO Correlations were, on average, $\sim 6.5\%$ higher in the group SS than in the group SD (2-sided t-test, $t_{10} = 3.43$, $p = 0.007$).

Additionally, the machine learning classifications of brain dynamics fingerprints described earlier were qualitatively generally in agreement with these findings. Leave-one-attribute-out classification revealed that, for both fingerprints, the accuracy in predicting individual subjects was the lowest when the training and validation sets were based on different days (see Supplementary Table 5.4).

To further evaluate the influence that different scanning variables have on MP Differences and FO Correlations, we directly compared their effects across these fingerprints. We first analyzed the raw medians of the distributions of MP Differences and FO Correlations for each scanning factor in the groups SS (Same subject Same factor attribute), SD (Same subject Different factor attributes), and DS (Different subjects Same factor attribute). We found that, while both fingerprints possessed a shared variance (Fig. 5.5(a), Coefficient of determination $R^2 = 0.375$), they also provided independent information. In fact, as evident not only by their distributions of values in Fig. 5.4, MP Differences displayed consistently larger median differences within the three groups of values (SS, SD, DS) than FO Correlations (Fig. 5.5(a), 2-sided Wilcoxon signed rank test, $z = 3.68$, $p < 10^{-3}$).
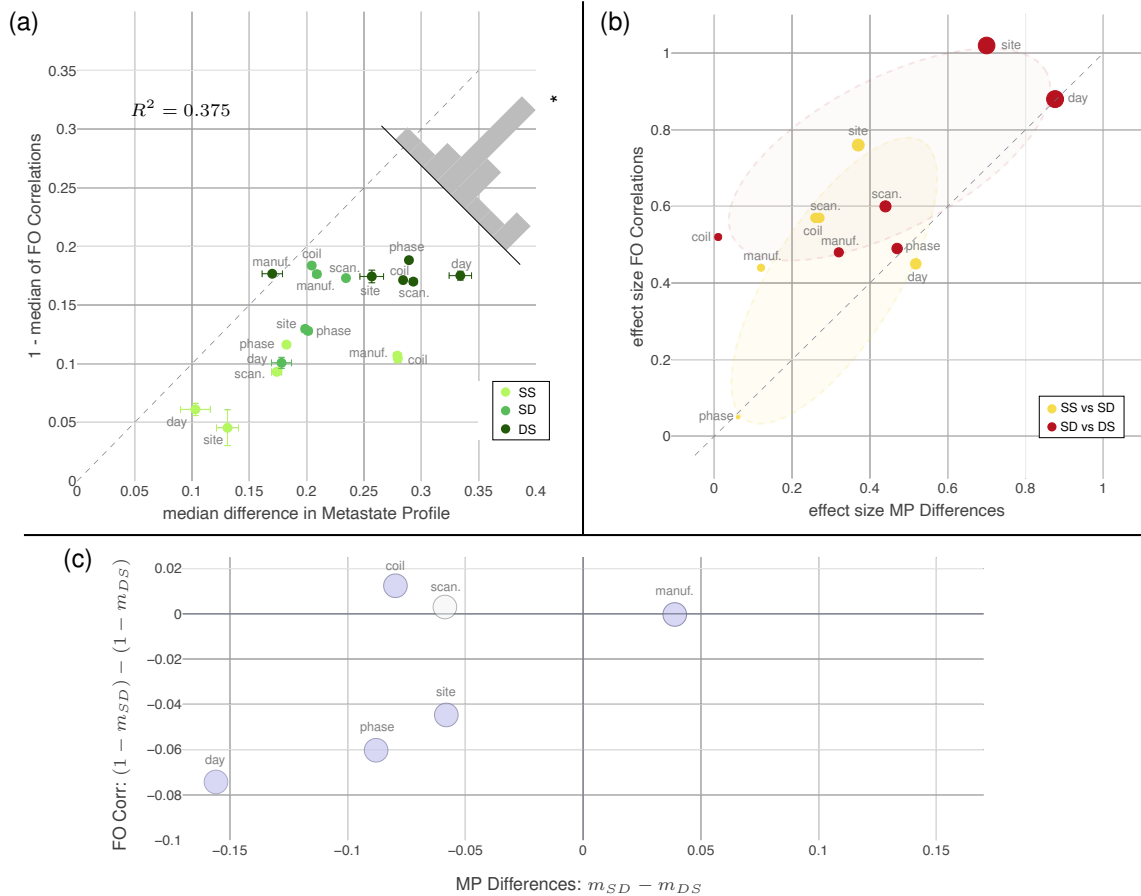
Figure 5.5: Effect of scanning factors within and across MP Differences and FO Correlations distributions. The dashed line represents the diagonal $y=x$. (a) The x and y axes represent the median of MP Differences and 1 - median of FO Correlations, respectively, for different scanning factors in different groups SS, SD, and DS, along with the standard error of mean. The median of MP Differences is more affected by all of the scanning factors (Wilcoxon signed-rank test, $z = 3.68$, $p < 10^{-3}$). (b) The effect size (Cohen's d) values obtained by comparing the log-transformed distributions of the MP Differences and FO Correlations across different scanning session factors and attributes. The ellipses represent the least squares minimization of the distance from the cloud of points for each of the three sets [130]. The largest effect sizes were consistently caused by the factors site and day, for all the comparisons between groups of distributions. (c) Median differences between the sets SD and DS. Positive values suggest that the noise induced in our metrics by different scanning factors is larger than the inherent inter-subject differences. The factor scanner is depicted differently as it was not statistically significant for the FO Correlations (see Table 5.1).

Next, to achieve an unbiased estimate of the effect size of each factor on the distributions

of MP Differences and FO Correlations, we computed the Cohen's d from the log-transformed

distributions of the MP Differences and FO Correlations across all scanning factors, between groups

SS-SD, and SD-DS. Fig. 5.5(b) highlights how the dissimilarity between brain dynamics fingerprints was the greatest when comparing, for the same scanning factor, measures from the same subject and measures from different subjects. Based on this observation, we assessed which scanning factors influenced the median values of the groups SD and DS the most by computing $m_{SD} - m_{DS}$, where m denoted the median, for the MP Differences, and $(1 - m_{SD}) - (1 - m_{DS})$ for the FO Correlations. Notice that a positive value indicates that the noise induced by different factors (group SD) has a larger effect than than inter-subject differences (group DS). In accordance with the analysis above, we found that most of the scanning factors seemed to induce less noise on our metrics than the inter-subject differences. Therefore, while site and day were the co-variate associated with the largest effect in the two groups SS vs SD and SD vs DS (Fig. 5.5(b)), the results in Fig. 5.5(c) suggest that the number of coils and the manufacturer are the only factors for which we can robustly estimate the effect on the data, beyond inter-subject differences.

5.4 Discussion

In this chapter, we addressed the issues of reproducibility and variability of fMRI data from the angle of brain dynamics. We leveraged the large HCP collection of rs-fMRI data to infer a hidden Markov model capable of describing brain state time courses at the subject level. By applying such a model to a dataset of traveling subjects, we found that brain network dynamics displayed signature fingerprints that were robust to different physical and temporal factors affecting the data that populates multi-site collections of neuroimaging data. Precisely, we found that MP Differences and FO Correlations are reliable, stable individual traits, as shown by the SS/DS differences, when taken across all factors (Fig. 5.4(a)-(b)). This study corroborates and complements previous work that found that the emergence of temporal patterns of brain activity tend to repeat more similarly within the same subject and over time [105,110,131]. This result promotes further investigations on the dynamical characteristics of brain states.

Recent years have witnessed a growing interest in the identification and characterization of the factors that tend to introduce spurious effects in multi-site fMRI recordings, endangering the reproducibility and the overall quality of the results that may be inferred from these data. The first warnings came from a study that investigated sources of nuisance variation across multiple sites and their impact on rs-fMRI data [100], followed by a number of studies that reported mostly consistent results [91–93, 96, 97]. Although we used different methods, our finding that different scanning factors introduce noise into brain dynamics fingerprints (as can be seen by the SS/SD comparisons in Fig. 5.5(b)) is in line with prior reports [93, 96]. Furthermore, the present study made use of larger datasets. The Traveling-subject dataset contains the largest number of subjects out of all the aforementioned studies. To date, only [93] has more sites than the Traveling-subject dataset used in this study, but it has the drawback of scanning only a single subject. The Traveling-subject dataset also allowed for the analysis of some scanning factors – such as the numbers of channels per coil or different scanner models within the same vendor – that have not been taken into consideration in previous work, giving more breadth and depth to our findings. Nevertheless, it is important to stress out that the potential variability brought in by scanning factors may not always be a necessarily negative feature. In fact, such variability may even be a powerful test for reproducibility of some findings. For instance, although different scanning factors may be confounds for certain analyses (e.g., comparing participant populations from different sites), they can also be used to test the robustness of a model when generalizing analyses across sites with different scanning parameters. Our results complement, from a dynamical point of view, both seminal and more recent work reporting more dissimilar resting-state networks inter-subject than intra-subject [93, 100, 116, 132].

Functional connectivity – typically computed as the correlation between time series representing the average activity in a brain region – has been the mainstay in the analysis of variability in fMRI data. Previous work has demonstrated that a sizable amount of recordings from the same site enables precise measurements of individual variations in functional connectivity [116], and that individual differences in functional networks are not affected by anatomical misalignment [132]. Here,

we complement such studies by showing that individual signatures can still be (easily) recovered within limited recordings from multiple sites (i.e., in the Traveling-subject dataset). To note, the comparison of functional connectivity between scanning sessions is inherently different from the comparison of state time courses. Differently from functional connectivity computed over a whole scanning session, the HMM captures the local temporal wandering of brain activity across states (networks). Therefore, while comparing functional connectivity between different subjects may be akin to comparing longer-term traits, comparing state time courses between subjects may be more closely aligned to comparing a repetition of sequences of brain states at rest.

While the aforementioned studies on functional connectivity have significantly increased our understanding of the brain as a system that obeys network-wide principles, they are mainly agnostic to temporal dynamics within the scanning sessions. This may prevent the level of precision that could at times be the most clinically relevant [133]. Differently from [93, 97], where time seemed to play a negligible effect, we found that different scanning days greatly affected our estimation of brain network dynamics. As mentioned earlier, an intuitive explanation for this apparent discrepancy is that functional connectivity tends to be associated to more coarsely defined subjective traits, whereas an HMM, being inherently more sensitive to temporal differences, is apt to capture more instantaneous cognitive processes. It is worth noting that our findings do not go against the claim that functional connectivity networks remain a reliable subject-specific fingerprint over long period of times, but rather we suggest that brain state trajectories can differ extensively between days, probably due to different cognitive or mental processes. As such, we suggest that dynamic and static measures in fact carry complementary information, which may provide additional insight when used in combination. However, due the limited number of scanning sessions taken on different days in the Traveling-subject dataset (see Supplementary Table 5.2), and to the inter-subject variability being larger than the variability across different factors (see Fig. 5.5(c)), future studies will need to further validate this fact.

How do HMMs compare with the sliding window approach? The sliding window analysis is typically used to improve the temporal definition of functional connectivity studies [134]. Albeit being intrinsically easier to set up, it has crucial limitations. Namely, the sliding window size is constrained by a trade-off between time resolution and quality of the results, and the conclusions from sliding window studies tend to be affected by sampling variability [135]. Conversely, HMM is as fast as the data modality allows, since it provides instantaneous likelihood of high correlation between brain signals [110].

If the goal of a study is a robust and detailed description of a system's dynamics, the HMM approach requires large amounts of data for training purposes, thus appearing not suitable to analyze small cohorts of subjects. However, in this study we give proof-of-concept that one can use a very large dataset (i.e. HCP) to infer an HMM, which can then be applied to a smaller dataset. Our results indicate that this procedure is robust. Interestingly, if only a relatively small number of subjects is available for the inference process, it is still possible to recover a coarser – and nonrandom – representation of the brain dynamics by using the TPM inferred from a large dataset as a prior (see Supplementary Figure 5.11(d)). Thus, detailed analyses and claims based on hidden Markov modeling should be gauged on the size of the available data. This is a common requirement in neuroimaging studies, as functional connectivity studies also require large amount of data to enable precise measurements [116].

Despite its capabilities, hidden Markov modeling is based on some premises (see also [108, 110] for thorough discussions). It is worth noting that the HMM builds on the Markovian assumption, theorizing that we can predict, based on the state we are at time t, which state is more likely to follow at time $t + 1$. Yet, while the brain may violate this assumption due to established long-range temporal dependencies [81], FO Correlations and MP Differences inherently display information that appears at longer time scales.

There are some limitations to this work. For example, while the decoding approach utilized here - training the HMM on the HCP data, and infer brain states trajectories in the Traveling-subject

83

dataset, is a strength of this study, it is also one of its limitations. The HCP and Traveling-subject datasets harbor some differences relating to the scanning protocols or even the countries in which the data were collected (US and Japan). For instance, the sampling rate of the two datasets were originally different ($\text{TR}_{\text{HCP}} = 0.72$ s and $\text{TR}_{\text{TS}} = 2.5$ s). As such, the up-sampling of the Traveling-subject dataset may have been sub-optimal and thus bias the overall HMM-based brain state dynamics estimation. Yet, the presence of these very differences appear to corroborate the finding that brain dynamics fingerprints are subject-specific. Specifically, we still find that, on average, the MP Differences (resp., FO Correlations) are lower (resp., higher) within subjects than across subjects, even when comparing runs with different scanning parameters. The fact that, at the within-subject level, these two measures had very similar values to those obtained from the HCP dataset (where both model and fingerprints were derived from the same data) provides strong support for this interpretation (i.e., brain dynamics fingerprints are subject-specific), such that it is unlikely that these results are due to inherent bias or noise. A second limitation may arise from the factors that were considered in the Traveling-subject dataset. Although there are several factors, some with many attributes (e.g., there are 12 sites), these factors are sometimes nested within each other. For instance, within the same phase encoding attribute there are scans belonging to different sites. This aspect may have partly influenced (reduced) the effect size of such factors which are heterogeneous with respect to *other* factors, while factors such as day or site would remain unaffected, since these scans were recorded at the same site, with the same protocol. Above all, while the Traveling-subject dataset allowed us to investigate the nuisance effect of multiple variables, it did not offer any insight into other relevant scanning factors such as TR length, the duration of the scanning session, and voxel sizes. To enhance our collective appraisal of the sources of variability in heterogeneous collections of rs-fMRI data, it will be important to generate datasets that include variations along these additional dimensions.

Given the considerable recent advances in inference techniques [108,136,137], and the ever-increasing availability of computational power, our work further suggests that the HMM is, and,

most importantly, will be, a powerful technique to explain and interpret the dynamic aspects of the brain. Furthermore, the possibility of inferring an HMM on a very large dataset to apply it to a much smaller one has important implications for clinical applications. In the future, perhaps with even more data, these general models could be built and then utilized to infer subject-specific fingerprints in other smaller cohorts and be used for a more personalized approach to treatments. In other words, a one-size-fits-all approach could be employed to build the model in its general terms, consequently allowing us to move to a personalized course of action by evaluating the model at the individual level. For instance, closed-loop fMRI neurofeedback [138, 139] could significantly benefit from these models, which will allow for a more holistic approach to the dynamical properties of mental and cognitive processes, particularly from a clinical perspective [114, 133, 140, 141].

5.5   Conclusion

In this chapter, we address the important issues of reproducibility and variability of fMRI data. We leveraged the large, homogeneous HCP collection of resting-state data to reliably infer a hidden Markov model capable of describing the brain state time courses at the subject level. By applying such a model to a dataset of traveling subjects, we show that dynamical states can be estimated reliably. Specifically, we find that brain network dynamics displays fingerprints that are robust to different scanning factors and distinctive for each subject. Further, we explore which scanning factors impact measures of brain dynamics the most, and what is the magnitude of their effect. We find that, amongst the scanning factors available in our dataset, sites and days tend to induce higher variability in the estimation of individual brain state time courses. However, due to the large noise induced by inter-subject variability and the limited sample size, this claim will need further validation by future studies.

These results enable and promote further investigations on the dynamical characteristics of brain states. Once a good model is inferred, it can be applied to a battery of different goals,

85

such as the analysis of task-based datasets, the examination of data collections from subjects with neurological disorders, and the promising use in clinical or rehabilitation settings, for instance by using brain state inference in clinical populations to estimate the best time for providing a given treatment.

5.6   Appendix

5.6.1   Control Analysis with Autoregressive Data

To verify the robustness of our analysis in regards to the application of the HMM model to the Traveling-subject dataset, we applied the HCP-trained HMM to autoregressive data. To assess a model $AR(n)$ of order n fairly and prevent it from trying to disproportionately fit to the first n datapoints, we discarded the first $3n$ datapoints. For this analysis, we have chosen autoregressive model of order 5. With this $AR(5)$ model, we have generated 50 time series as dummy ICA timeseries to be used the model inference. Analogously to the real ICA timeseries, the AR time series have been normalized (zero mean and unitary standard deviation).

The decoding of our model on autoregressive time series yielded state time courses that in which state 5 is predominant in all different scanning sessions. See Fig. 5.11 for a few examples of the state time courses obtained from HMM decoding of the randomly permute Traveling-subject time series. It is worth noting that state 5 is the state that is mostly uncorrelated from the remaining 11 states and it is the state with the largest variance. In [110], it is demonstrated that state 5 is associated with motion artifacts in the scanner. The outcome of the HMM decoding on autoregressive data is in accordance with these observations, therefore supporting the robustness of our findings.
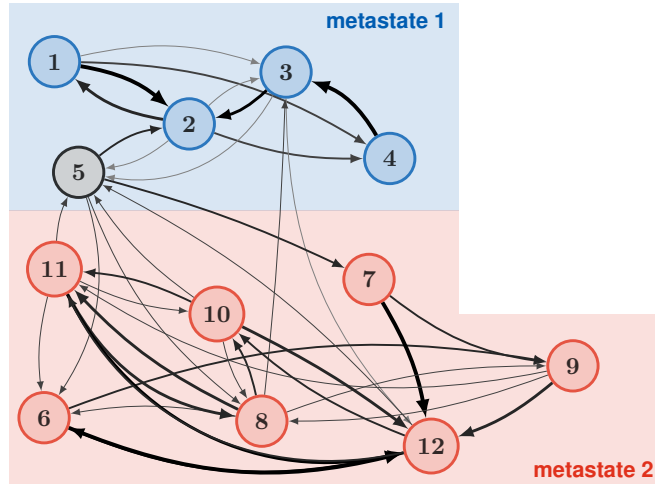
Figure 5.6: Network associated with the Transition Probability Matrix of the HMM used in this study. The first 4 states comprise metastate 1, whereas states 6 to 11 comprise metastate 2. State 5 is mostly uncorrelated to the other states, is associated with head motion, and has the highest variance [110]. The interconnections depicted in this graph represent probabilities higher than 10% (i.e. > 0.1) in the HMM's Transition Probability Matrix, and their color and thickness are proportional to their magnitude.
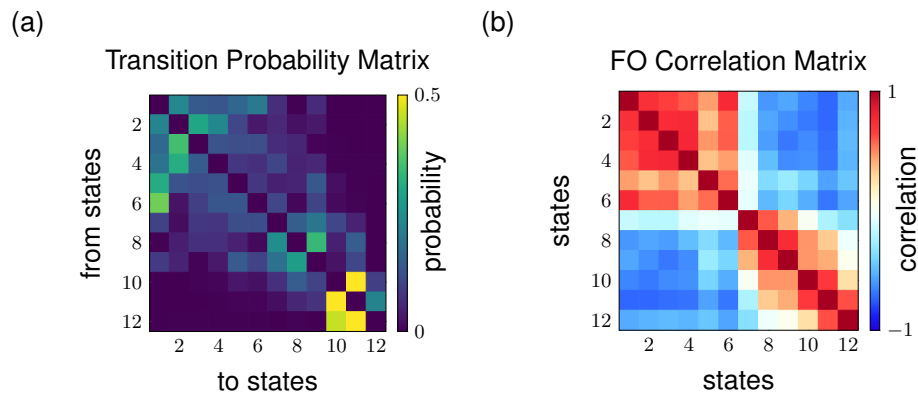
Figure 5.7: Best Model inferred on downsampled HCP time series. (a) Transition probability matrix for the model inferred on downsampled HCP time series. The TPM is very skewed towards only three states: 10,11, 12. (b) FO Correlation Matrix for the model inferred on downsampled HCP time series. The metastate structure (groups of states in which the state time courses tend to remain) is not as clear as on the original time series, but still emergent.
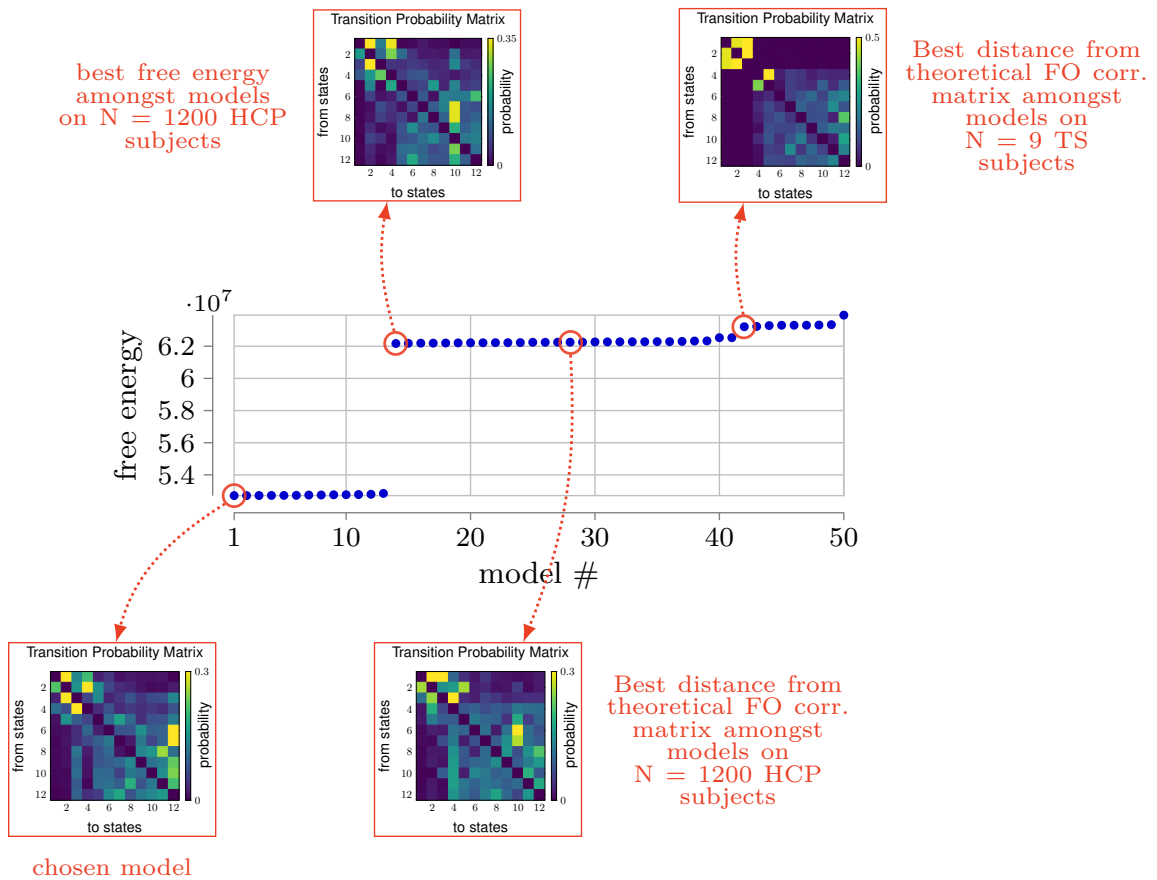
Figure 5.8: Models ranking based on the free energy. We ranked the $N = 50$ models inferred in this study based on their free energy, and chose the one minimizing this quantity. To compare all the models ($N_{1200} = 28$ models inferred on 1200-subject HCP release, $N_{820} = 14$ models inferred on the 820-subject HCP release, and $N_{TS} = 6$ models inferred on the 9 subjects of the Traveling-Subject dataset, so that $N_{1200} + N_{820} + N_{TS} = 50$), each free energy calculation has been adjusted to account for different dataset sizes. The FO Correlation Matrix can be found in Figure S2, and further details on different models from the best one can be found in Figure S4.
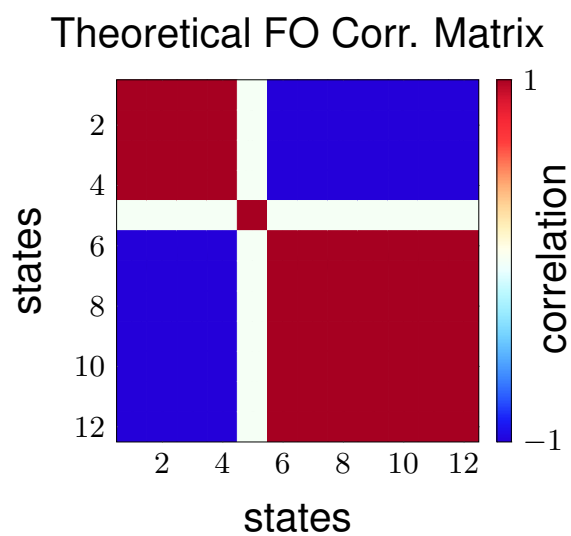
Figure 5.9: Ideal FO Correlation Matrix. For each model inferred from our datasets, we compute the Euclidean distance between the model's FO Correlation Matrix and the ideal one. We rank our models based on this distance with the aim of selecting the model that has the most clear emergence of the 2-metastate structure.
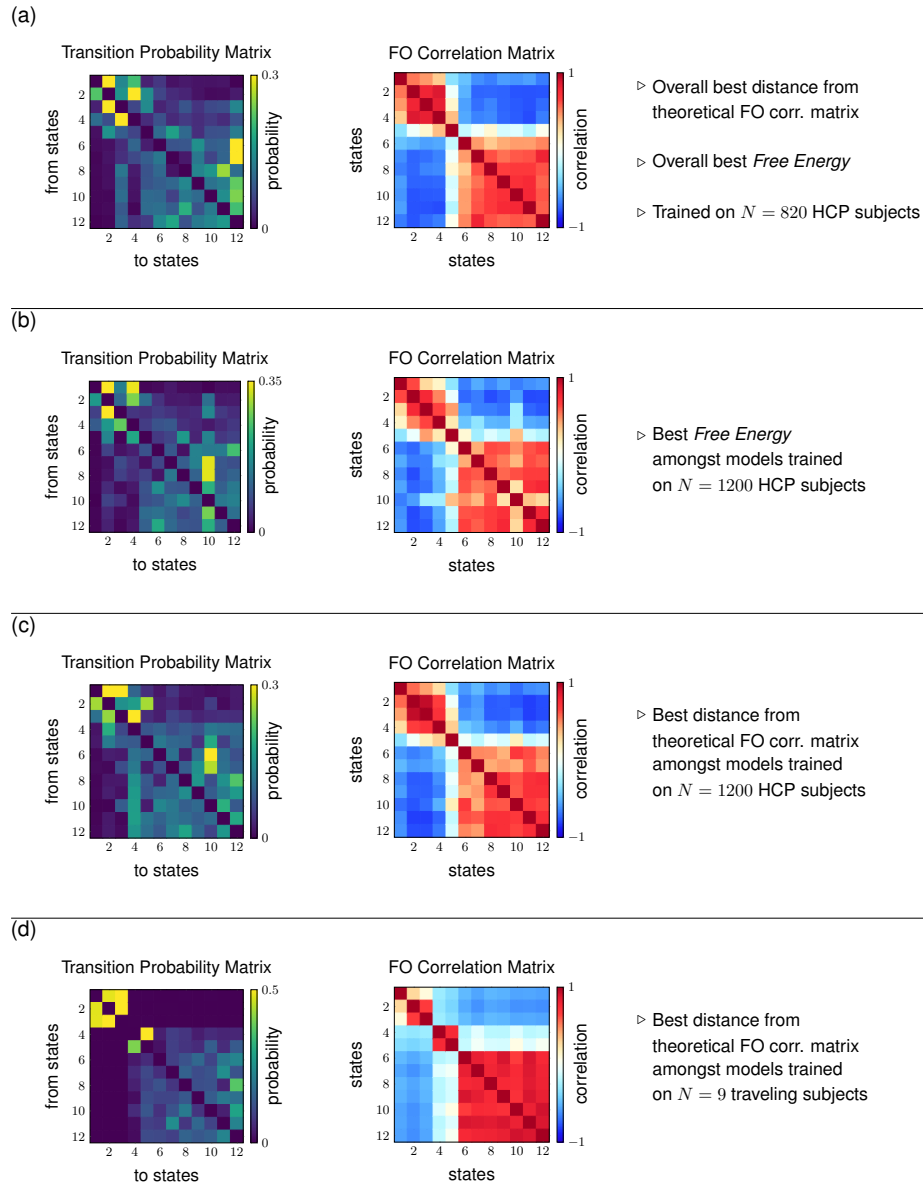
Figure 5.10: Transition Probability Matrix and Fractional Occupancy Correlation Matrix for different HMM models.(a) The best model among all the models trained in terms of free energy and Euclidean distance from the ideal FO Correlation Matrix. (b) The best model in terms of free energy among all models trained on the subjects of the HCP 1200-subject distribution. (c) The best model in terms of Euclidean distance from the ideal FO Correlation Matrix among all models trained on the subjects of the HCP 1200-subject distribution. (d) The best model in terms of Euclidean distance from the ideal FO Correlation Matrix among all models trained on the Traveling-subject dataset by using the model (a) as a prior. Notice that, while the model in this last panel displays very distinct metastate separation in the TPM matrix, such a matrix is not irreducible (i.e. there does not exist a path connecting the two groups of states), making it not suitable to represent any biological system.
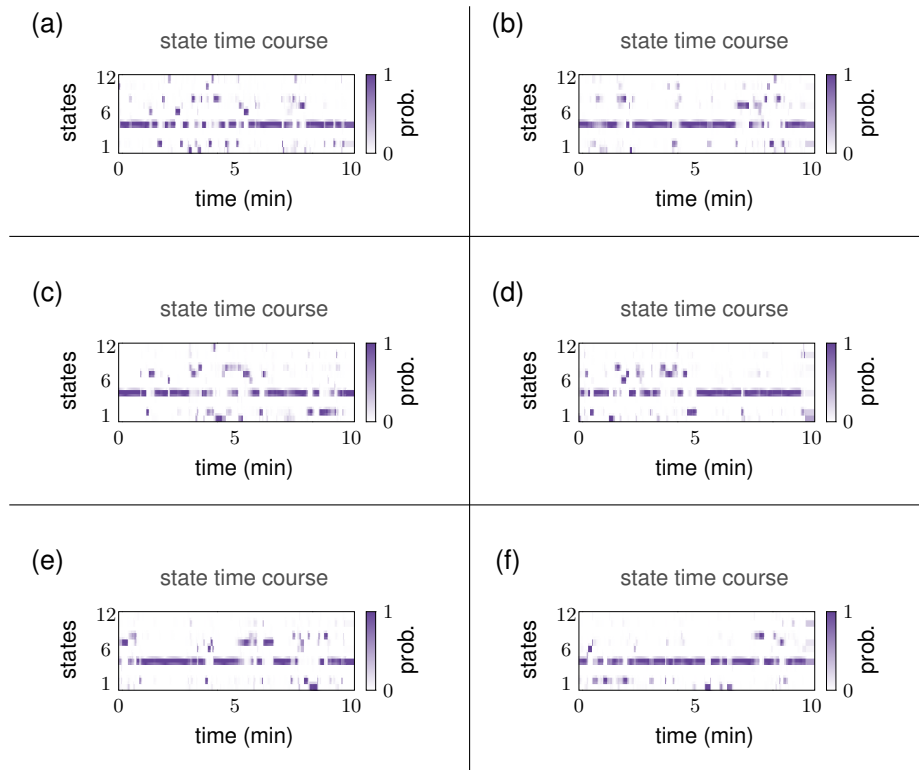
Figure 5.11: Examples of state time courses after HMM decoding on AR(5) time series. (a)-(f) To provide a baseline for our study, we applied the HMM model used in this chapter to autoregressive data with dimension $50 \times T$, where T denotes the number of time points. The HMM decoding yields state time courses that stay most of the time in state 5, which is the state that is highly uncorrelated from the other 11 states and the one with the highest variance. This fact supports the goodness of fit of the inferred model, as randomized time series do not provide meaningful state time courses.
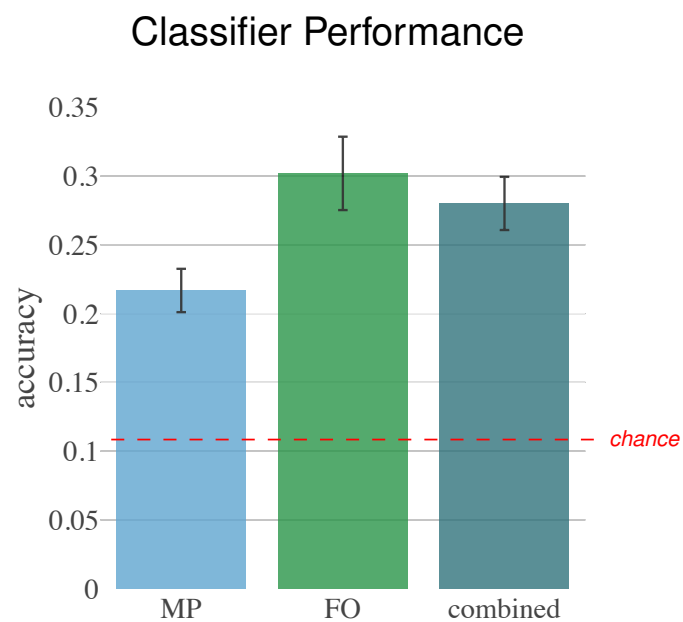
Figure 5.12: Summary of the leave-one-attribute-out cross-validation for all scanning factors. The three bars represent the average classification accuracy across all different scanning factors, along with the standard deviation, for the Metastate Profiles (MP), the Fractional Occupancies (FO), and the combination of the two, respectively. The red dashed line indicates the baseline chance level. These results show that the personal signature of brain dynamics fingerprints emerges even with a simple logistic regression classifier.
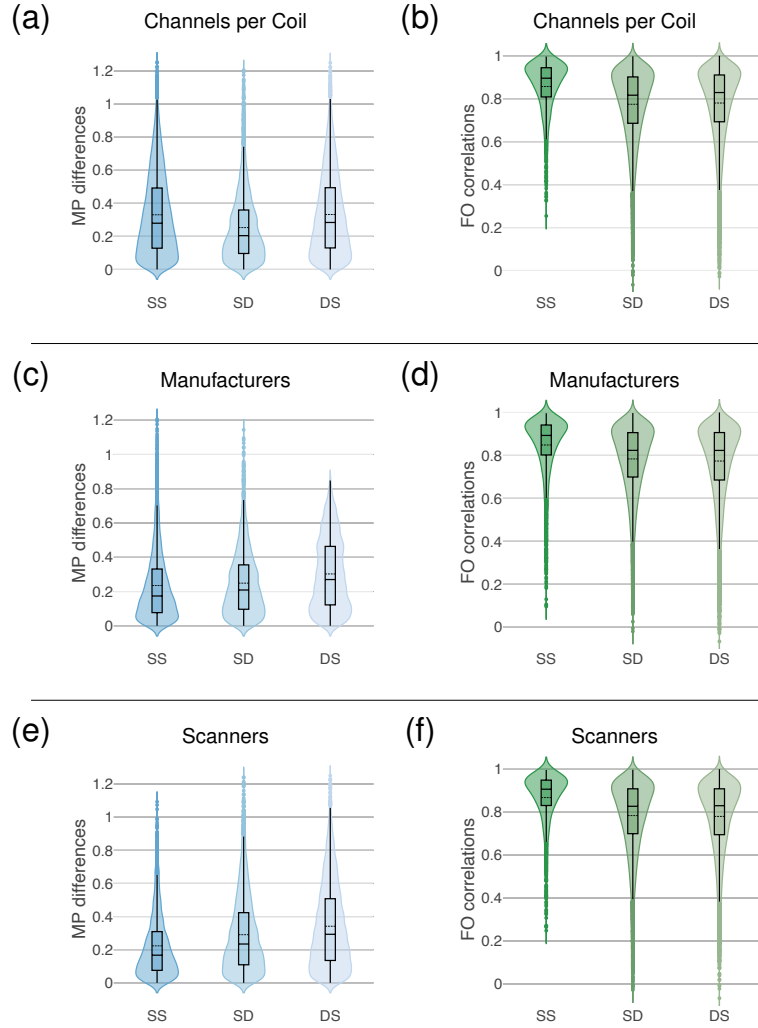
Figure 5.13: Distributions of values for MP Differences and FO Correlations, for the factors: numbers of channels per coil, manufacturers, and scanner model. In panels (a) to (f), the set SS comprises the MP Differences (resp., FO Correlations) computed for each subject within the same factor attribute, and the SS distribution displays these values for all subjects; the set SD consists of the MP Differences (resp., FO Correlations) computed for each subject across different attributes of the same factor, and the SD distribution displays these values for all subjects; finally, the set DS consists of the MP Differences (resp., FO Correlations) computed across all subjects within the same factor attribute, and the DS distribution displays these values for all attributes of the same factor. Further, for all the distributions, the black dashed lines illustrate the mean. The difference between SS and DS distributions in panel (a) and the difference between SD and DS distributions in panel (f) are not statistically significant (see also Table 1 in the main text).

Table 5.2: Imaging protocols for resting-state fMRI in the Traveling-subject dataset.

Site	ATT	ATV	COI	HUH	HKH	KPM	SWA	KUT	KUS	UTO	YC1	YC2
1. Scanner Manufacturer	Siemens	Siemens	Siemens	GE	Siemens	Philips	Siemens	Siemens	Siemens	GE	Philips	Philips
2. Scanner Model	TimTrio	Verio	Verio	Signa HDxt	Spectra	Achieva	Verio	TimTrio	Skyra	MR750W	Achieva	Achieva
3. Magnetic Field Strength	3.0 T	3.0 T	3.0 T	3.0 T	3.0 T	3.0 T	3.0 T	3.0 T	3.0 T	3.0 T	3.0 T	3.0 T
4. No. Channels per Coil	12	12	12	8	12	8	12	32	32	24	8	8
5. Field-of-view (mm)	212 × 212	212 × 212	212 × 212	212 × 212	212 × 212	212 × 212	212 × 212	212 × 212	212 × 212	212 × 212	212 × 212	212 × 212
6. Matrix	64 × 64	64 × 64	64 × 64	64 × 64	64 × 64	64 × 64	64 × 64	64 × 64	64 × 64	64 × 64	64 × 64	64 × 64
7. No. of Slices	40	39	40	35	35	40	40	40	40	40	40	40
8. No. of Volumes	240	240	240	240	240	240	240	240	240	240	240	240
9. In-plane Resolution (mm)	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125	3.3125× 3.3125
10. Slice Thickness (mm)	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2
11. Slice Gap (mm)	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
12. TR (ms)	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
13. TE (mm)	30	30	30	30	30	30	30	30	30	30	30	30
14. Total Scan Time (min:s)	10 : 00	10 : 00	10 : 00	10 : 00	10 : 00	10 : 00	10 : 00	10 : 00	10 : 00	10 : 00	10 : 00	10 : 00
15. Flip Angle (deg)	80	80	80	80	80	80	80	80	80	80	80	80
16. Slice Acquisition Order	Ascending	Ascending	Ascending	Ascending	Ascending	Ascending	Ascending	Ascending	Ascending	Ascending	Ascending	Ascending
17. Phase Encoding	PA	PA	AP	PA	PA	AP	PA	PA	AP	PA	AP	AP
18. Eyes Closed/Fixate	Fixate	Fixate	Fixate	Fixate	Fixate	Fixate	Fixate	Fixate	Fixate	Fixate	Fixate	Fixate
19. Field Map	yes	yes	yes	-	-	yes	yes	yes	yes	yes	yes	-

Table 5.3: Kolmogorov-Smirnov test statistics for MP Differences and FO Correlations

Parameter	MP Differences		FO Correlations	
	SS vs SD	SD vs DS	SS vs SD	SD vs DS
1. Site	0.28	0.16	0.55	0.14
2. Day	0.44	0.25	0.47	0.27
3. Phase	0.04	0.15	0.04	0.17
4. Channels/Coil	0.14	0.15	0.25	0.04
5. Manufacturer	0.07	0.14	0.22	0.03
6. Scanner	0.15	0.09	0.26	0.01

Table 5.4: Logistic regression accuracy results

Parameter	MP	FO	combined
1. Site	0.21	0.29	0.29
2. Day	0.19	0.27	0.25
3. Phase	0.22	0.29	0.27
4. Channels/Coil	0.23	0.30	0.28
5. Manufacturer	0.22	0.34	0.32
6. Scanner	0.22	0.32	0.27
Mean	0.22	0.30	0.28
Standard Deviation	0.02	0.03	0.02

Chapter 6

Cluster Synchronization in Networks of Phase Oscillators

In this chapter, we present a top-down approach to the modeling and analysis of brain dynamics synchronization. That is, we propose and analyze a dynamical network model to replicate the complex structure-function relationship of neural systems. As functional brain networks are characterized by the degree of similarity – *synchronization* – between neural activity of distinct brain regions, here we model neural activity as the output of a network of nonlinear oscillators, and reveal the mechanisms underlying different synchronization patterns.

To do so, we study cluster synchronization in networks of oscillators with heterogenous Kuramoto dynamics, where multiple groups of oscillators with identical phases coexist in a connected network. Cluster synchronization is at the basis of several biological and technological processes; yet the underlying mechanisms to enable cluster synchronization of Kuramoto oscillators have remained elusive. In this chapter we derive quantitative conditions on the network weights, cluster configuration, and oscillators' natural frequency that ensure asymptotic stability of the cluster synchronization manifold; that is, the ability to recover the desired cluster synchronization configuration

following a perturbation of the oscillators' states.

Qualitatively, our results show that cluster synchronization is stable when the intra-cluster coupling is sufficiently stronger than the inter-cluster coupling, the natural frequencies of the oscillators in distinct clusters are sufficiently different, or, in the case of two clusters, when the intra-cluster dynamics is homogeneous. We refer the reader to [142] and [143] for a comprehensive discussion.

6.1 Introduction

Synchronization refers broadly to patterns of coordinated activity that arise spontaneously or by design in several natural and man-made systems [144–146]. Examples include coherent firing of neuronal populations in the brain [147], coordinated flashing of fireflies [148], flocking of birds [149], exchange of signals in wireless networks [150], consensus in multi-agent systems [29], and power generation in the smart grid [151]. Synchronization enables complex functions: while some systems require complete (or full) synchronization among all the components in order to function properly, others rely on cluster (or partial) synchronization, where different groups exhibit different, yet synchronized, internal behaviors [152].

While studies of full synchronization are numerous and have generated a rich literature, e.g., see [153–155], conditions explaining the onset of cluster synchronization and its properties are less well understood. Such conditions are necessary for the analysis and, more importantly, the control of synchronized activity across biological [156–158] and technological [159] systems. For instance, a deeper understanding of the mechanisms enabling cluster synchronization might not only shed light on the nature of the healthy human brain [160], but also enable and guide targeted interventions for patients with neurological disorders, such as epilepsy [161] and Parkinson's disease [162]. In fact, he underlying mechanisms of cluster synchronization are particularly useful to model, analyze, and regulate synchronized neural activity in the human brain (see Fig. 6.1).

We study cluster synchronization in networks of oscillators with Kuramoto dynamics [163],
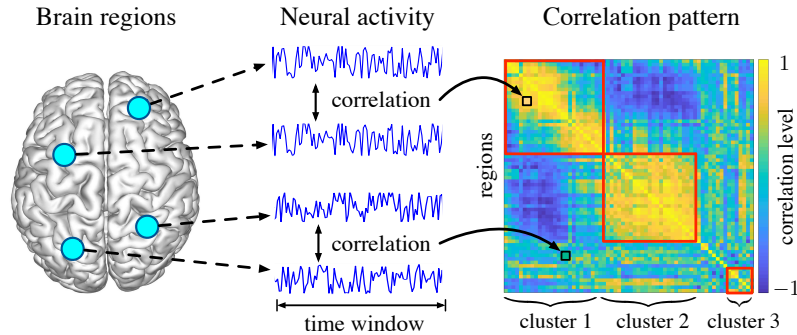
Figure 6.1: As evident from fMRI scans, neural activity across different regions is correlated over time, and this correlation pattern defines clusters of synchronized brain areas corresponding to well-defined neural systems. Distinct correlation patterns are biomarkers of healthy and diseased neural states.

which, despite their apparent simplicity, are particularly suited to represent complex synchronization phenomena in neural systems [164], as well as in many other natural and technological systems [151]. Although our study and modeling choices are guided by the practical need to understand and control patterns of synchronized functional activity in the human brain, as they naturally arise in healthy and diseased populations [165, 166], in this chapter we focus on developing the mathematical foundations of a quantitative approach to the analysis and control of cluster synchronization in a weighted network of Kuramoto oscillators. In particular, we derive conditions on the oscillators' coupling and their natural frequencies that guarantee the stability of an arbitrary cluster configuration.

Related work Cluster synchronization, where multiple synchronized groups of oscillators coexist in a connected network, is an exciting phenomenon that has attracted the attention of the physics, dynamical systems, and controls communities, among others. Existing work on this topic has shown that cluster-synchronized states can be linked to the existence of certain network symmetries [167–171] or symmetries in the nodes' dynamics [172]. More recently, in [173, 174], the stability of cluster states corresponding to network symmetries is addressed with the Master Stability Function approach [175]. In contrast to this previous work, [176] combines network symmetries with contraction analysis to study the stability of synchronized states. Further studies relating contraction properties and cluster synchronization are conducted in [177, 178]. Finally, control algorithms

for cluster synchronization are developed in [179, 180]. To the best of our knowledge, however, the above studies are not applicable to oscillators with Kuramoto dynamics, which we study in this work.

A few papers have studied cluster synchronization of Kuramoto oscillators. Specifically, in [181, 182] the authors provide invariance conditions for an approximate definition of cluster synchronization and for particular types of networks. Invariance of exact cluster synchronization, which is the notion used in this chapter, is also studied in [183, 184]. Stability of exact cluster synchronization is investigated in [185] where, however, only the restrictive case of two clusters for identical Kuramoto oscillators with inertia is considered, and in [186], where only implicit and numerical stability conditions are provided. To the best of our knowledge, our work presents the first explicit analytical conditions for the (local) stability of the cluster synchronization manifold in sparse and weighted networks of heterogeneous Kuramoto oscillators.

Chapter contribution The main contribution of this chapter is to characterize conditions for the stability of cluster synchronization in networks of oscillators with Kuramoto dynamics. We consider a notion of exact cluster synchronization, where the phases of the oscillators within each cluster remain equal to each other over time, and different from the phases of the oscillators in the other clusters. We derive four conditions. First, we show that the cluster synchronization manifold is locally exponentially stable when the intra-cluster coupling is sufficiently stronger than the inter-cluster coupling. We quantify this tradeoff using the theory of perturbation for dynamical systems together with the invariance properties of cluster synchronization. Second, through a Lyapunov argument, we show that the cluster synchronization manifold is locally exponentially stable when the natural frequencies of the oscillators in disjoint clusters are sufficiently different (in their limit to infinity). Third, we focus on the case of two clusters, and provide a quantitative condition on the network weights and oscillators' natural frequency for the stability of the cluster synchronization manifold. Four, we combine the previous conditions through a heuristic approach that provides a tighter, yet approximate, stability condition. Our holistic analysis shows that asymptotic stability

of the cluster synchronization manifold is guaranteed for weak inter-cluster weights, sufficiently different natural frequencies, or even homogeneous intra-cluster configurations.

As minor contributions, we provide examples showing that network symmetries are not necessary for cluster synchronization of Kuramoto oscillators, and a sufficient condition guaranteeing the absence of stable synchronization submanifolds.

Chapter organization The rest of the chapter is organized as follows. Section 9.2 contains our problem setup and some preliminary notions. Section 9.3 contains our main results; that is, our conditions for the stability of the cluster synchronization manifold in Kuramoto networks. Finally, section 9.6 concludes the chapter, and the Appendix contains the proofs of our results.

Mathematical notation The set $\mathbb{R}_{>0}$ (resp. $\mathbb{R}_{<0}$) denotes the positive (resp. negative) real numbers, whereas the sets $\mathbb{S}^1$ and $\mathbb{T}^n$ denote the unit circle and the n-dimensional torus, respectively. The vector of all ones is represented by $\mathbb{1}$. We let $O(f)$ denote the order of the function f. Further, we denote a positive (resp. negative) definite matrix A with $A \succ 0$ (resp. $A \prec 0$). Let $\lambda_i(A)$ and $\sigma_i(A)$ denote the i-th eigenvalue and singular value of $A \in \mathbb{R}^{n \times n}$, respectively. We indicate the smallest (resp. largest) eigenvalue of a symmetric matrix with $\lambda_{\min}(\cdot)$ (resp. $\lambda_{\max}(\cdot)$). A (block-)diagonal matrix is represented by (blk-)diag($\cdot$). Further, let $\overline{\lambda}(A) = \frac{1}{n}\sum_i \lambda_i(A)$ and $\overline{\sigma}(A) = \frac{1}{n}\sum_i \sigma_i(A)$. We let $\| \cdot \|$ denote the ℓ^2-norm, and $\mathsf{i} = \sqrt{-1}$. Finally, $A^\dagger$ represents the Moore-Penrose pseudoinverse of the matrix A.

6.2 Problem setup and preliminary notions

In this work we characterize the stability properties of certain synchronized trajectories arising in networks of oscillators with Kuramoto dynamics. To this aim, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the connected and weighted graph representing the network of oscillators, where $\mathcal{V} = \{1, \ldots, n\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represent the oscillators, or nodes, and their interconnection edges, respectively. Let $A = [a_{ij}]$ be the weighted adjacency matrix of $\mathcal{G}$, where $a_{ij} \in \mathbb{R}_{>0}$ is the weight of the edge

$(i, j) \in \mathcal{E}$, and $a_{ij} = 0$ when $(i, j) \notin \mathcal{E}$. The dynamics of i-th oscillator is

$$\dot{\theta}_i = \omega_i + \sum_{j \neq i} a_{ij} \sin(\theta_j - \theta_i), \qquad (6.1)$$

where $\omega_i \in \mathbb{R}_{>0}$ and $\theta_i \in \mathbb{S}^1$ denote the natural frequency and the phase of the i-th oscillator. Unless specified differently, we assume that the edge weights are symmetric. That is,

(A1) The network adjacency matrix satisfies $A = A^\mathsf{T}$.

Assumption (A1) is typical in the study of (cluster) synchronization in networks of Kuramoto oscillators, e.g., see [28, 187, 188], as it facilitates the derivation of stability results. While relaxing this assumption is beyond the scope of this work, we will discuss how our stability results can also be applied to study cluster synchronization with asymmetric network weights (see Remark 34). Finally, since the diagonal entries of the adjacency matrix A do not contribute to the dynamics in (9.1), we assume that $\mathcal{G}$ does not contain self-loops.

A network exhibits cluster synchronization when the oscillators can be partitioned so that the phases of the oscillators in each cluster evolve identically. To be precise, let $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$, with $m > 1$, be a partition of $\mathcal{V}$, where $\bigcup_{i=1}^m \mathcal{P}_i = \mathcal{V}$ and $\mathcal{P}_i \cap \mathcal{P}_j = \emptyset$ if $i \neq j$. Define the *cluster synchronization manifold* associated with the partition $\mathcal{P}$ as

$$\mathcal{S}_\mathcal{P} = \{\theta \in \mathbb{T}^n \ : \ \theta_i = \theta_j \text{ for all } i, j \in \mathcal{P}_k, \ k = 1, \ldots, m\}.$$

Then, the network is cluster-synchronized with partition $\mathcal{P}$ when the phases of the oscillators belong to $\mathcal{S}_\mathcal{P}$ at all times.

In this chapter we characterize conditions on the network weights and the oscillators' natural frequency that guarantee *local exponential stability* of the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$, for a given partition $\mathcal{P}$.[1] In order to study stability of the cluster synchronization manifold, we assume $\mathcal{S}_\mathcal{P}$ to be invariant [189, Chapter 3].[2] In particular, following [184], invariance of $\mathcal{S}_\mathcal{P}$ is guaranteed by the following conditions:

---

[1]Loosely speaking, the manifold $\mathcal{S}_\mathcal{P}$ is locally exponentially stable if θ converge to $\mathcal{S}_\mathcal{P}$ exponentially fast when $\theta(0)$ is sufficiently close to $\mathcal{S}_\mathcal{P}$.

[2]The manifold $\mathcal{S}_\mathcal{P}$ is invariant if $\theta(0) \in \mathcal{S}_\mathcal{P}$ implies $\theta \in \mathcal{S}_\mathcal{P}$ at all times.

$$A = \begin{bmatrix} 0 & \alpha_1 & 0 & \beta_1 & \beta_2 & 0 \\ \alpha_1 & 0 & \alpha_2 & 0 & \beta_1 & \beta_2 \\ 0 & \alpha_2 & 0 & \beta_2 & 0 & \beta_1 \\ \beta_1 & 0 & \beta_2 & 0 & \alpha_3 & 0 \\ \beta_2 & \beta_1 & 0 & \alpha_3 & 0 & \alpha_4 \\ 0 & \beta_2 & \beta_1 & 0 & \alpha_4 & 0 \end{bmatrix}$$

(a)   (b)
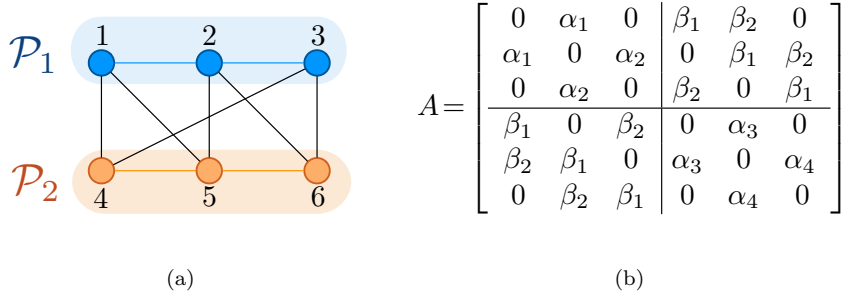
Figure 6.2: Fig. 6.2(a) illustrates a network of 6 oscillators with adjacency matrix as in Fig. 6.2(b). In this network, the partition $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2\}$, which satisfies Assumption (A3), cannot be identified by group symmetries of the network for any choice of the positive weights α_1, α_2, α_3, α_4, β_1 and β_2. The manifold $\mathcal{S}_\mathcal{P}$ is invariant whenever the oscillators' natural frequencies satisfy Assumption (A2). Thus, this example shows that network symmetries are not necessary for cluster synchronization of Kuramoto oscillators.

(A2) Given $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$, the natural frequencies satisfy $\omega_i = \omega_j$ for every $i, j \in \mathcal{P}_k$ and $k \in \{1, \ldots, m\}$,[3] and

(A3) The network weights satisfy $\sum_{k \in \mathcal{P}_\ell} a_{ik} - a_{jk} = 0$ for every $i, j \in \mathcal{P}_z$ and $z, \ell \in \{1, \ldots, m\}$, with $z \neq \ell$.

Thus, in the remainder of the chapter we assume that (A2) and (A3) are satisfied for the network partition being considered.

Remark 20 *(Network symmetries, equitable partitions, and balanced weights) Conditions to ensure the invariance of the cluster synchronization manifold have been linked to network symmetries, which are defined by the group comprising all node permutations that leave the network topology unchanged, e.g., see [173,174,186]. In Fig. 6.2 we propose a network with two clusters, which are not defined by any group symmetry, that satisfies Assumption (A3) and thus admits an invariant cluster synchronization manifold. This example shows that cluster synchronization of Kuramoto oscillators does not require symmetric networks. Our Assumption (A3), and in fact the equivalent notion of external equitable partition [183], is less restrictive than requiring partitions satisfying group sym-*

---

[3]This condition is necessary for $\mathcal{S}_\mathcal{P}$ to be forward invariant, and thus stable [184], and is motivated by observed synchronization phenomena, e.g., see [190].

metries [191–193]. Finally, Assumptions (A2) and (A3) are necessary when the natural frequencies in distinct clusters are sufficiently different (see [184] and Remark 21). ☐

Remark 21 *(**Invariance of submanifolds of** $\mathcal{S}_{\mathcal{P}}$) When the network of oscillators is cluster-synchronized (i.e. $\theta(t) \in \mathcal{S}_{\mathcal{P}}$ for all $t \geq 0$), submanifolds of $\mathcal{S}_{\mathcal{P}}$ may appear whenever the phases belonging to two (or more) disjoint clusters have equal values (see Fig. 6.3). Interestingly, the example in Fig. 6.3 also points out that Assumption (A3) may not be necessary for the invariance of $\mathcal{S}_{\mathcal{P}}$ if the clusters do not evolve with different frequencies (see Assumption (A1) in [184]). In what follows we show that, if the natural frequencies of the oscillators in disjoint clusters are sufficiently different, invariant, and hence stable, submanifolds cannot exist. To see this, assume that the phases of the disjoint clusters $\mathcal{P}_{\ell}$ and $\mathcal{P}_{z}$ remain equal over time. Then, using Assumption (A2) and (A3), the dynamics*

$$\dot{\theta}_{\ell} - \dot{\theta}_{z} = \omega_{\ell} - \omega_{z} + \sum_{k=1}^{m} \left[\left(\sum_{r \in \mathcal{P}_k} a_{\ell r} \right) \sin(\theta_k - \theta_{\ell}) - \left(\sum_{r \in \mathcal{P}_k} a_{zr} \right) \sin(\theta_k - \theta_z) \right], \qquad (6.2)$$

must be identically zero, where θ_i denotes the phase of any oscillator in $\mathcal{P}_i$. Clearly, if the following inequality holds,

$$|\omega_{\ell} - \omega_z| > 2(m-2) \max_{k \neq \ell, z} \left\{ \sum_{r \in \mathcal{P}_k} a_{\ell r} \,,\, \sum_{r \in \mathcal{P}_k} a_{zr} \right\}, \qquad (6.3)$$

Equation (6.2) cannot vanish and, consequently, the clusters $\mathcal{P}_{\ell}$ and $\mathcal{P}_z$ cannot evolve with the same phases when the network is cluster synchronized.[4] More generally, if condition (6.3) is satisfied for all pairs of clusters, then invariant, and hence stable, cluster synchronization submanifolds cannot exist. ☐

We conclude with an example showing that the synchronization manifold $\mathcal{S}_{\mathcal{P}}$ is, in general, not globally asymptotically stable due to the existence of multiple invariant sets.

Example 22 *(**Multiple invariant sets**) Consider a Kuramoto network with $2N$ oscillators ($N \geq 2$) and with an adjacency matrix defined as follows[5] (see Fig. 6.4(a) for the case $N = 5$):*

---

[4]In (6.3), we have $(m-2)$ because for $k = z, \ell$, the sine terms in (6.2) vanish.
[5]This analysis extends directly to arbitrary weights $a_{ij} = a$, $a \in \mathbb{R}_{>0}$.
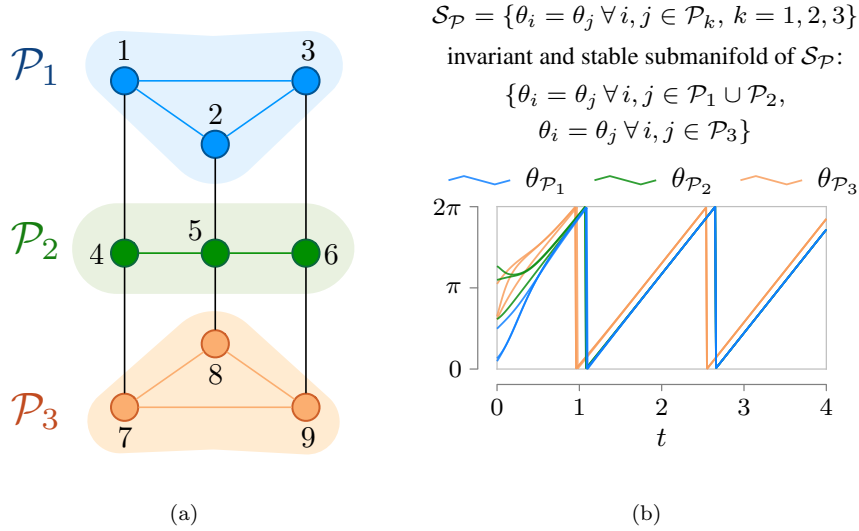
Figure 6.3: Fig. 9.5(a) illustrates a network with partition $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3\}$. As shown in Fig. 6.3(b), the phases of the oscillators in $\mathcal{P}_1$ and $\mathcal{P}_2$ have the same value over time, showing that a submanifold of $\mathcal{S}_\mathcal{P}$ is invariant and stable. For this simulation, we use $\omega_1 = 4$, $\omega_2 = 2$, $\omega_3 = 6$, $a_{14} = 3$, and $a_{47} = 5$.

$$a_{ij} = \begin{cases} 1, & \text{if } |i - j| \leq 2, \\ 0, & \text{otherwise,} \end{cases}$$

with $i, j \in \{1, \ldots, 2N\}$ (and the convention $2N + \ell \triangleq \ell$, $-\ell \triangleq 2N + \ell - 1$, for $\ell \in \{1, 2\}$). Let $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2\}$, with $\mathcal{P}_1 = \{1, 3, \ldots, 2N - 1\}$, $\mathcal{P}_2 = \{2, 4, \ldots, 2N\}$, and define

$$\mathcal{M}_\mathcal{P} = \{\theta \in \mathbb{T}^{2N} : \theta_{i+2} = \theta_i + 2\pi/N, i = 1, \ldots, 2N - 2\}.$$

It can be verified that Assumption (A3) is satisfied, and that the set $\mathcal{S}_\mathcal{P}$ is invariant whenever the natural frequencies satisfy Assumption (A2). Yet, the set $\mathcal{S}_\mathcal{P}$ is not the only invariant set. In fact, $\mathcal{M}_\mathcal{P}$ is also invariant (we prove this by showing that $\dot{\theta}_i = \dot{\theta}_{i+2}$ when $\theta_i, \theta_{i+2} \in \mathcal{M}_\mathcal{P}$):

$$\dot{\theta}_i = \omega_i + \sin(\theta_{i-2} - \theta_i) + \sin(\theta_{i+2} - \theta_i) + \sin(\theta_{i-1} - \theta_i) + \sin(\theta_{i+1} - \theta_i)$$

$$= \omega_i + \sin(\theta_i - \theta_{i+2}) + \sin(\theta_{i+4} - \theta_{i+2}) + \sin(\theta_{i+1} - \theta_{i+2}) + \sin(\theta_{i+3} - \theta_{i+2}) = \dot{\theta}_{i+2},$$

where we have used the fact that $\theta_{i+2} - \theta_i = 2\pi/N$, and $\omega_i = \omega_j$ for all i, j in the same cluster. Further, it can be verified numerically that, depending on the number of oscillators N, the set $\mathcal{M}_\mathcal{P}$

104

(a)

(b)

(c)

(d)
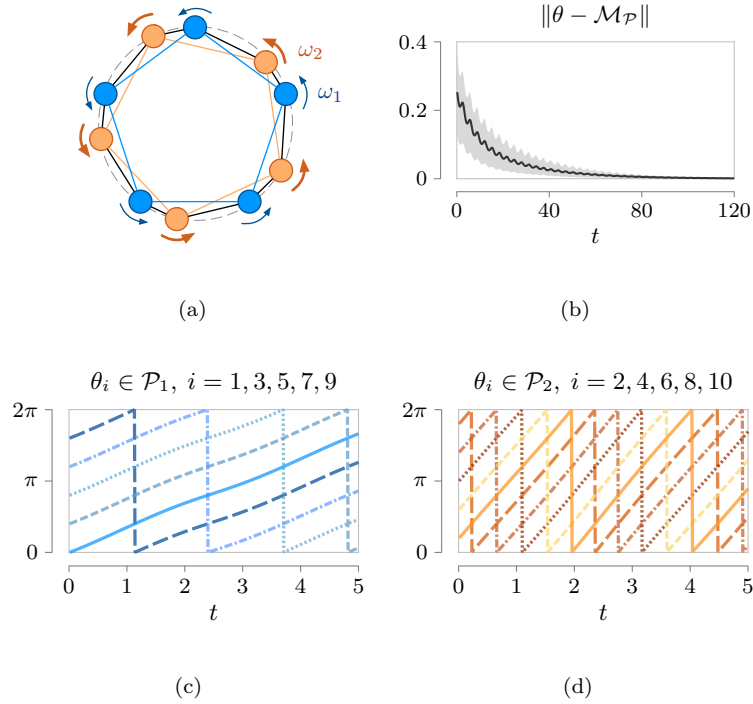
Figure 6.4: Fig. 6.4(a) shows the network in Example 22 for the case $N = 5$. The nodes belonging to partition $\mathcal{P}_1$ are blue and have natural frequency $\omega_1 = 1$, while the nodes belonging to partition $\mathcal{P}_2$ are orange and have $\omega_2 = 3$. Fig. 6.4(b) illustrates the stability of the set $\mathcal{M}_{\mathcal{P}}$ via numerical simulations. We performed 10^3 iterations, each one with $\theta(0)$ chosen randomly within an angle of ± 0.01 [rad] from $\mathcal{M}_{\mathcal{P}}$. The thick line represents the mean value among all simulations of the 2-norm distance between θ and $\mathcal{M}_{\mathcal{P}}$, while the faded area represents the smallest and largest value of the 2-norm distance between θ and $\mathcal{M}_{\mathcal{P}}$. Fig. 6.4(c)-6.4(d) illustrate the invariance of the set $\mathcal{M}_{\mathcal{P}}$ as the phases in the clusters $\mathcal{P}_1$ and $\mathcal{P}_2$ evolve respectively with the same frequencies.

is also locally stable (see Fig. 6.4(b)). We conclude that the cluster synchronization manifold $\mathcal{S}_{\mathcal{P}}$

is not, in general, globally asymptotically stable. In what follows we derive conditions guaranteeing

local stability of $\mathcal{S}_{\mathcal{P}}$. □

6.3 Analytical conditions for the stability of the cluster synchronization manifold

In this section we derive sufficient conditions for the local exponential stability of the cluster synchronization manifold. Define the phase difference $x_{ij} = \theta_j - \theta_i$, and notice that

$$\dot{x}_{ij} = \omega_j - \omega_i + \sum_{z=1}^{n} \left[a_{jz} \sin(x_{jz}) - a_{iz} \sin(x_{iz}) \right]. \tag{6.4}$$

Given a partition $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$ of the set $\mathcal{V}$ in the graph $\mathcal{G}$, we define the following graphs (see also Example 23):

(i) the graph of the k-th cluster, with $k \in \{1, \ldots, m\}$, $\mathcal{G}_k = (\mathcal{P}_k, \mathcal{E}_k)$, where $\mathcal{E}_k = \{(i,j) \; : \; (i,j) \in \mathcal{E}, \; i, j \in \mathcal{P}_k\}$;

(ii) a spanning tree $\mathcal{T}_k = (\mathcal{P}_k, \mathcal{E}_{\text{span},k})$ of $\mathcal{G}_k$;[6]

(iii) a spanning tree $\mathcal{T} = (\mathcal{V}, \mathcal{E}_{\mathcal{T}})$ of $\mathcal{G}$ with $\mathcal{E}_{\mathcal{T}} = \bigcup_{k=1}^{m} \mathcal{E}_{\text{span},k} \cup \mathcal{E}_{\text{inter}}$, where $\mathcal{E}_{\text{inter}}$ satisfies $|\mathcal{E}_{\text{inter}}| = m - 1$.

Further, we define the following vectors of phase differences:

(iv) $x_{\text{intra}}^{(k)} = [x_{ij}]$, for all $(i,j) \in \mathcal{E}_{\text{span},k}$ with $i < j$,

(v) $x_{\text{intra}} = \left[x_{\text{intra}}^{(1)\mathsf{T}}, \ldots, x_{\text{intra}}^{(m)\mathsf{T}} \right]^{\mathsf{T}}$, and

(vi) $x_{\text{inter}} = [x_{ij}]$, for all $(i,j) \in \mathcal{E}_{\text{inter}}$ with $i < j$.

It should be noticed that the vectors $x_{\text{intra}}^{(k)}$, x_{intra} and x_{inter} contain, respectively, $n_{\text{intra},k} = |\mathcal{P}_k| - 1$, $n_{\text{intra}} = n - m$ and $n_{\text{inter}} = m - 1$ entries. Notice that every phase difference can be computed as a linear function of x_{intra} and x_{inter}. To see this, let $i, j \in \mathcal{V}$, and let $\text{p}(i,j) = \{p_1, \ldots, p_\ell\}$ be the unique path on $\mathcal{T}$ from i to j. Define $\text{diff}(\text{p}(i,j)) = \sum_{k=1}^{\ell-1} s_k$, where $s_k = x_{p_k p_{k+1}}$ if $p_k < p_{k+1}$, and $s_k = -x_{p_{k+1} p_k}$ otherwise. Then, $x_{ij} = \text{diff}(\text{p}(i,j))$, and the vectors x_{intra} and x_{inter} contain

---

[6]We assume that $\mathcal{G}$ and its subgraphs $\mathcal{G}_k$ are connected. This guarantees the existence of the (connected) spanning trees defined in (ii) and (iii). A graph is connected if there exists a path between any pair of nodes [42].

a smallest set of phase differences that can be used to quantify synchronization among all of the oscillators in the network.

Let $B = [b_{k\ell}] \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{E}|}$ denote the oriented incidence matrix of the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where ℓ corresponds to the edge $(i,j) \in \mathcal{E}$, $b_{k\ell} = 1$ if node k is the sink of the edge ℓ, $b_{k\ell} = -1$ if k is the source of ℓ, and $b_{k\ell} = 0$ otherwise.[7] Further, let B_k and $B_{\text{span},k}$ denote the incidence matrices of $\mathcal{G}_k$ and $\mathcal{T}_k$, respectively. Notice that $B_{\text{span},k}$ is full rank because it is the incidence matrix of an acyclic graph (tree) [42, Theorem 8.3.1]. Let $T_{\text{intra},k} = B_k^{\mathsf{T}}(B_{\text{span},k}^{\mathsf{T}})^{\dagger}$ be the unique matrix that maps the phase differences contained in $x_{\text{intra}}^{(k)}$ to all intra-cluster phase differences in the k-th cluster. That is,

$$x^{(k)} = T_{\text{intra},k} x_{\text{intra}}^{(k)}, \tag{6.5}$$

where $x^{(k)}$ contains all phase differences in the cluster $\mathcal{P}_k$.

We conclude this part by rewriting the intra-cluster dynamics in a form that will be useful to prove our results. In particular, from the above discussion and for an intra-cluster phase difference x_{ij} of $x_{\text{intra}}^{(k)}$, we rewrite (6.4) as

$$\dot{x}_{ij} = \underbrace{\sum_{z \in \mathcal{P}_k} [a_{jz} \sin(\text{diff}(\text{p}(j,z))) - a_{iz} \sin(\text{diff}(\text{p}(i,z)))]}_{F_{ij}^{(k)}(x_{\text{intra}}^{(k)})}$$

$$+ \underbrace{\sum_{z \notin \mathcal{P}_k} [a_{jz} \sin(\text{diff}(\text{p}(j,z))) - a_{iz} \sin(\text{diff}(\text{p}(i,z)))]}_{G_{ij}^{(k)}(x_{\text{intra}}, x_{\text{inter}})}, \tag{6.6}$$

which leads to

$$\dot{x}_{\text{intra}}^{(k)} = F^{(k)}(x_{\text{intra}}^{(k)}) + G^{(k)}(x_{\text{intra}}, x_{\text{inter}}), \tag{6.7}$$

where $F^{(k)}$ is the vector of $F_{ij}^{(k)}$ and $G^{(k)}$ is the vector of $G_{ij}^{(k)}$, for all $(i,j) \in \mathcal{E}_{\text{span},k}$ with $i < j$. Finally, by concatenating the dynamics (6.7) for all clusters, we obtain

$$\dot{x}_{\text{intra}} = F(x_{\text{intra}}) + G(x_{\text{intra}}, x_{\text{inter}}). \tag{6.8}$$
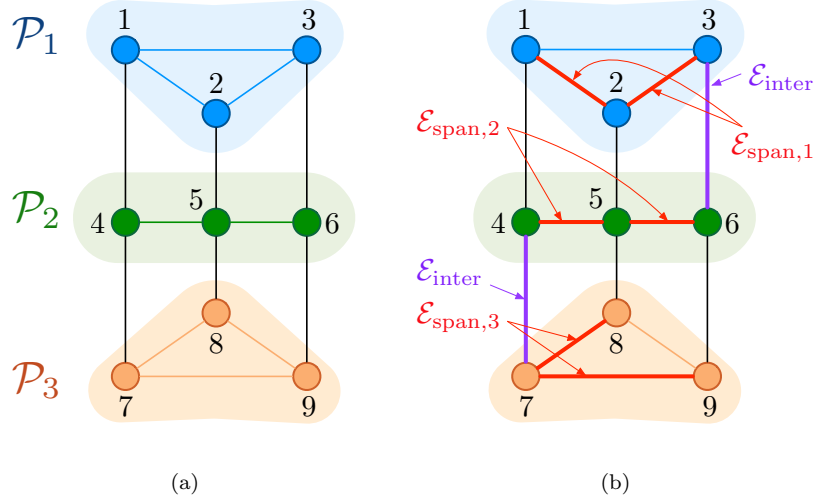
Figure 6.5: This figure illustrates the graph-theoretic definitions introduced in Section 9.3 for a network of 9 Kuramoto oscillators. (see also Example 23). Fig. 6.5(a) shows the partitions $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3\}$. In Fig. 6.5(b), $\mathcal{E}_{\text{span},1}$, $\mathcal{E}_{\text{span},2}$, and $\mathcal{E}_{\text{span},3}$ represent (in red) the edges of the intra-cluster spanning trees $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_3$, while the edges belonging to the set $\mathcal{E}_{\text{inter}}$ are depicted in purple.

Example 23 *(Illustration of the definitions)* *We provide here an illustrative example of the definitions introduced in this section. Consider the network in Fig. 6.5(a) with partition $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3\}$, where $\mathcal{P}_1 = \{1, 2, 3\}$, $\mathcal{P}_2 = \{4, 5, 6\}$ and $\mathcal{P}_3 = \{7, 8, 9\}$. Fig. 6.5(b) illustrates the definitions of spanning trees, together with the edge sets $\mathcal{E}_{span,k}$ $(k = 1, 2, 3)$, and the inter-cluster edges in $\mathcal{E}_{inter} = \{(3, 6), (4, 7)\}$. The vectors of intra-cluster differences read as $x_{intra}^{(1)} = [x_{12}\ x_{23}]^{\mathsf{T}}$, $x_{intra}^{(2)} = [x_{45}\ x_{56}]^{\mathsf{T}}$, and $x_{intra}^{(3)} = [x_{78}\ x_{79}]^{\mathsf{T}}$, whereas the vector of inter-cluster differences reads as $x_{inter} = [x_{36}\ x_{47}]^{\mathsf{T}}$.*

For the partition $\mathcal{P}_1$, order the edges as $\ell_1 = (1, 2)$, $\ell_2 = (1, 3)$, and $\ell_3 = (2, 3)$. Then, a spanning tree is $\mathcal{T}_1 = (\mathcal{P}_1, \mathcal{E}_{span,1})$, with $\mathcal{E}_{span,1} = \{(1, 2), (2, 3)\}$, and the (oriented) incidence matrices B_1 of $\mathcal{G}_1$ and $B_{span,1}$ of $\mathcal{T}_1$ are

$$B_1 = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}, \quad B_{span,1} = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

---

[7]Node i is the source (resp. sink) of (i, j) if $i < j$ (resp. $i > j$).

Finally, the matrix $T_{intra,1} = B_1^\mathsf{T}(B_{span,1}^\mathsf{T})^\dagger$ satisfies

$$T_{intra,1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^\mathsf{T}.$$

$\square$

6.3.1 Asymptotic stability of $\mathcal{S}_\mathcal{P}$ via perturbation theory

In what follows we will make use of perturbation theory of dynamical systems to provide our first stability condition. We first introduce the following instrumental result.

Lemma 24 *(**Properties of intra-cluster dynamics**) The intra-cluster dynamics* (6.8) *satisfies the following properties:*

(i) *the Jacobian matrix J_{intra} of $F(x_{intra})$ computed at the origin is Hurwitz stable and can be written as*

$$J_{intra} = \left.\frac{\partial F(x_{intra})}{\partial x_{\text{intra}}}\right|_{x_{intra}=0} = \text{blkdiag}\left(J_1, \ldots, J_m\right), \tag{6.9}$$

where, for $k \in \{1, \ldots, m\}$, $T_{intra,k}$ is as in (6.5) and

$$J_k = -B_{span,k}^\mathsf{T} B_k \, \text{diag}(\{a_{ij}\}_{(i,j)\in\mathcal{E}_k})T_{intra,k}. \tag{6.10}$$

Thus, the origin is an exponentially stable equilibrium of the system $\dot{x}_{intra} = F(x_{intra})$;

(ii) *There exist constants $\gamma^{(k\ell)} \in \mathbb{R}_{>0}$ such that*

$$\|G^{(k)}(x_{intra}, x_{inter})\| \leq \sum_{\ell=1}^{m} \gamma^{(k\ell)}\|x_{intra}^{(\ell)}\|, \tag{6.11}$$

for all $k, \ell \in \{1, \ldots, m\}$. Specifically,

$$\gamma^{(k\ell)} = 2 \max_r \, n_{intra,r} \, \tilde{\gamma}^{(k\ell)}, \tag{6.12}$$

where, for any $i \in \mathcal{P}_k$,

$$\tilde{\gamma}^{(k\ell)} = \begin{cases} \displaystyle\sum_{j \in \mathcal{P}_\ell} a_{ij}, & \text{if } \ell \neq k, \\ \displaystyle\sum_{\substack{\ell=1 \\ \ell \neq k}}^{m} \sum_{j \in \mathcal{P}_\ell} a_{ij}, & \text{otherwise.} \end{cases} \tag{6.13}$$

As formalized in the next theorem, Lemma 24, together with results on stability of perturbed systems [194, Chapter 9], implies that the origin of (6.8), and thus the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$, is exponentially stable for some choices of the network weights. Recall that an M-matrix is a real nonsingular matrix $A = [a_{ij}]$ such that $a_{ij} \leq 0$ for all $i \neq j$ and all leading principal minors are positive [195, Chapter 2.5].

Theorem 25 *(**Sufficient condition on network weights for the stability of $\mathcal{S}_\mathcal{P}$**) Let $\mathcal{S}_\mathcal{P}$ be the cluster synchronization manifold associated with a partition $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$ of the network $\mathcal{G}$ of Kuramoto oscillators. Let $\gamma^{(k\ell)}$ be the constants defined in (6.12). Define the matrix $S \in \mathbb{R}^{m \times m}$ as*

$$S = [s_{k\ell}] = \begin{cases} \lambda_{max}^{-1}(X_k) - \gamma^{(kk)} & \text{if } k = \ell, \\ -\gamma^{(k\ell)} & \text{if } k \neq \ell, \end{cases} \tag{6.14}$$

where $X_k \succ 0$ is such that $J_k^\mathsf{T} X_k + X_k J_k = -I$, with J_k as in (7.2). If S is an M-matrix, then the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$ is locally exponentially stable.

Remark 26 *(**Family of bounds**) In (7.3), the matrices X_k can be selected as the solutions to the Lyapunov equations $J_k^\mathsf{T} X_k + X_k J_k = -Q_k$, for arbitrary positive definite matrices Q_k. Yet, selecting $Q_k = I$ for all k yields a tighter stability bound. This follows because (i) if S is an M-matrix, then $S + \Delta$ remains an M-matrix whenever Δ is a nonnegative diagonal matrix [195, Theorem 2.5.3], and (ii) the ratio $\lambda_{min}(Q_k)/\lambda_{max}(X_k)$ is maximal whenever $Q_k = I$ [194, Exercise 9.1].* $\square$

Theorem 25 describes a sufficient condition on the network weights for the stability of the cluster synchronization manifold. Loosely speaking, the cluster synchronization manifold is

$$A = \begin{bmatrix} 0 & \alpha_1 & \alpha_1 & \beta & 0 & 0 \\ \alpha_1 & 0 & \alpha_1 & 0 & \frac{\beta}{2} & \frac{\beta}{2} \\ \alpha_1 & \alpha_1 & 0 & 0 & \frac{\beta}{2} & \frac{\beta}{2} \\ \hline \beta & 0 & 0 & 0 & \alpha_2 & \alpha_2 \\ 0 & \frac{\beta}{2} & \frac{\beta}{2} & \alpha_2 & 0 & \alpha_2 \\ 0 & \frac{\beta}{2} & \frac{\beta}{2} & \alpha_2 & \alpha_2 & 0 \end{bmatrix}$$

(a)   (b)

$\sim \ \|x_{\text{intra}}^{(1)}\| \qquad \sim \ \|x_{\text{intra}}^{(2)}\|$

(c)   (d)
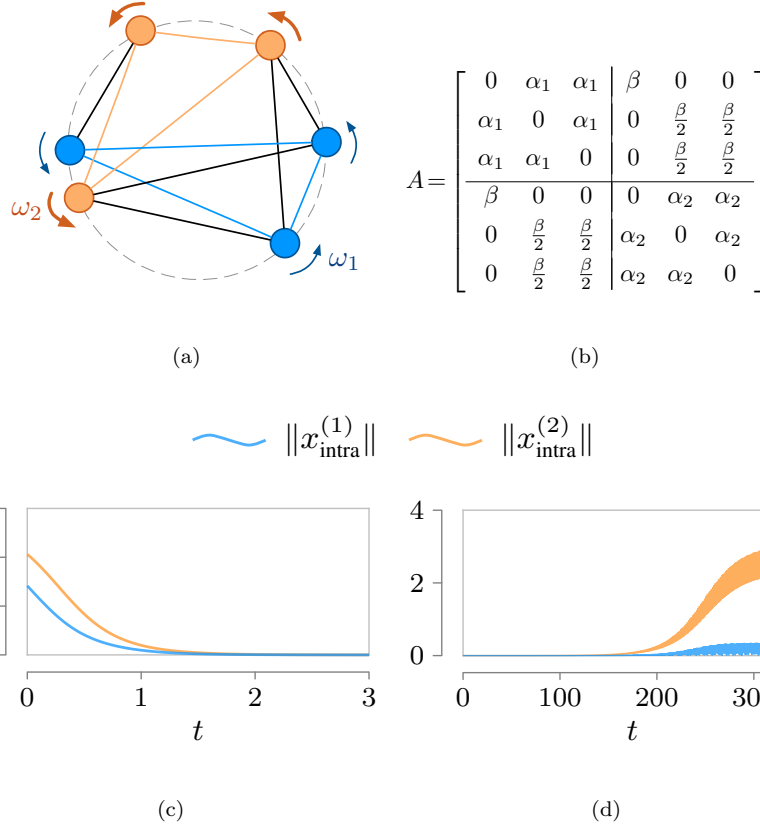
Figure 6.6: Fig. 6.12(a) illustrates the network of 6 Kuramoto oscillators in Example 36. We identify the clusters $\mathcal{P}_1$ and $\mathcal{P}_2$ in blue and orange, respectively. Fig. 6.6(b) contains the adjacency matrix of the network in Fig. 6.12(a). The parameters $\alpha_1, \alpha_2,$ and β represent the intra-cluster and inter-cluster weights, respectively. Fig. 6.6(c) illustrates the stability of the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$ for $\alpha_1 = \alpha_2 = 1$ and $\beta = 0.1$, as predicted by Theorem 25. Fig. 6.6(d) shows that $\mathcal{S}_\mathcal{P}$ is unstable when $\alpha_1 = \beta = 1$ and $\alpha_2 = 0.001$.

exponentially stable when the intra-cluster coupling (measured by $\lambda_{\max}^{-1}(X_k) - \gamma^{(kk)}$) is sufficiently stronger than the perturbation induced by the inter-cluster connections (measured by $\gamma^{(k\ell)}$). In particular, the term $\lambda_{\max}^{-1}(X_k)$ is proportional to the intra-cluster weights and it is implicitly related to the network topology. In fact, the matrix X_k is the solution of a Lyapunov's equation containing J_k, whose spectrum coincides with the stable eigenvalues of the negative Laplacian matrix of the k-th cluster. We refer the interested reader to the proof of Lemma 24. Finally, we remark that a result akin to Theorem 25 has been derived in [196], although for interconnected systems whose coupling functions are required to satisfy certain assumptions that fail to hold in the Kuramoto model.
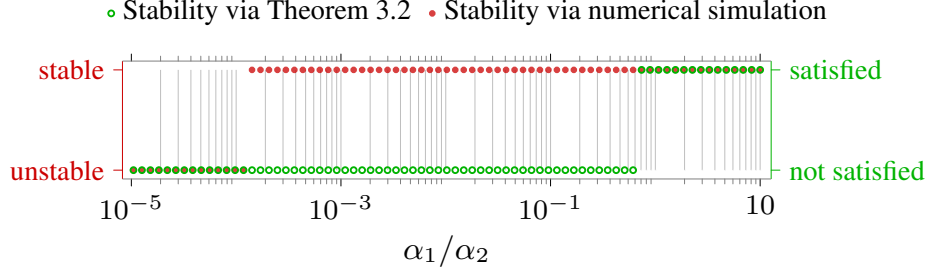
Figure 6.7: This Figure shows that the condition in Theorem 25 leads to conservative stability bounds. For the network in Example 36, we let $\beta = 0.1$ and plot, as a function of the ratio α_1/α_2, the stable configurations predicted by Theorem 25 (green) and those found numerically. For each value of α_1/α_2, we assess numerical stability by making use of the Floquet stability theory [197, Chapter 5] and by resorting to statement (i) in Lemma 29. This is possible because the partition in Example 36 has only two clusters.

Example 27 *(Tradeoff between intra- and inter-cluster weights) Consider the network in Fig. 6.12(a) with partition $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2\}$, where $\mathcal{P}_1 = \{1, 2, 3\}$ and $\mathcal{P}_2 = \{4, 5, 6\}$, natural frequencies $\omega_1 = 1$ and $\omega_2 = 6$ for the oscillators in $\mathcal{P}_1$ and $\mathcal{P}_2$, and adjacency matrix as in Fig. 6.6(b). The parameters $\alpha_1, \alpha_2 \in \mathbb{R}_{>0}$ and $\beta \in \mathbb{R}_{>0}$ denote the strength of the intra- and inter-cluster coupling, respectively. Let $\alpha_1 = \alpha_2$, and construct the matrix S as in Theorem 25:*

$$S = \begin{bmatrix} \lambda_{max}^{-1}(X_1) - \gamma_{11} & -\gamma_{12} \\ -\gamma_{12} & \lambda_{max}^{-1}(X_2) - \gamma_{22} \end{bmatrix},$$

where $X_k \succ 0$ is such that $J_k^\mathsf{T} X_k + X_k J_k = -I$, $\lambda_{max}^{-1}(X_1) = \lambda_{max}^{-1}(X_2) = 2\,\alpha_1$ and, from (6.12), $\gamma_{ij} = 4\,\beta$ for all i, j. By inspecting all leading principal minors, S is an M-matrix if $\alpha_1/\beta > 4$, and the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$ is exponentially stable (see Fig. 6.6(c)). We remark that, when $\alpha_1 \neq \alpha_2$, the synchronization manifold $\mathcal{S}_\mathcal{P}$ can become unstable, as we verify numerically in Fig. 6.6(d). □

The stability condition in Theorem 25 depends only on the network weights, and typically leads to conservative bounds (see also Fig. 6.7). To derive refined stability conditions, we next characterize how the natural frequencies of the oscillators affect stability of the cluster synchronization manifold.

112

6.3.2 Asymptotic stability of $\mathcal{S}_\mathcal{P}$ when the oscillators' natural frequencies in disjoint clusters are sufficiently different

Natural frequencies play a fundamental role for full and cluster synchronization of Kuramoto oscillators. However, while heterogeneity of the natural frequencies typically impedes full synchronization [28], we will show that cluster synchronization is in fact facilitated when the oscillators in different clusters have sufficiently different natural frequencies. We start with an asymptotic result that is valid for arbitrary networks and partitions, and then improve our results for the case of partitions containing only two clusters.

Theorem 28 *(Stability of $\mathcal{S}_\mathcal{P}$ for large natural frequency differences) Let $\mathcal{S}_\mathcal{P}$ be the cluster synchronization manifold associated with a partition $\mathcal{P} = \{\mathcal{P}_1, \dots, \mathcal{P}_m\}$ of the network $\mathcal{G}$ of Kuramoto oscillators. Let $\omega_i \in \mathbb{R}_{>0}$ be the natural frequency of the oscillators in the cluster $\mathcal{P}_i$, with $i \in \{1, \dots, m\}$. In the limit $|\omega_i - \omega_j| \to \infty$, for all $i, j \in \{1, \dots, m\}$, $i \neq j$, the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$ is locally exponentially stable.*

Theorem 28 shows that heterogeneity of the natural frequencies of the oscillators in different clusters facilitates cluster synchronization, independently of the network weights. We remark that a similar behavior was also identified in [198], albeit with a different method and definition of synchronization.

We next improve upon Theorem 28 by analyzing the case where the natural frequencies are finite and the partition $\mathcal{P}$ contains only two clusters. To this aim, let $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2\}$ and assume, without loss of generality, that $\omega_2 \geq \omega_1$, where ω_i is the natural frequency of the oscillators in $\mathcal{P}_i$. Define

$$\bar{\omega} = \omega_2 - \omega_1, \text{ and } \bar{a} = \sum_{k \in \mathcal{P}_2} a_{ik} + \sum_{k \in \mathcal{P}_1} a_{jk},$$

for any $i \in \mathcal{P}_1$ and $j \in \mathcal{P}_2$. The next result characterizes the inter-cluster phase difference when the network evolves on the cluster synchronization manifold.

Lemma 29 *(Nominal inter-cluster difference) Let $\mathcal{S}_{\mathcal{P}}$ be the cluster synchronization manifold associated with a partition $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2\}$ of the network $\mathcal{G}$ of Kuramoto oscillators. Let $\theta(0) \in \mathcal{S}_{\mathcal{P}}$ (equivalently, $x_{\mathrm{intra}}(0) = 0$). Then, if $x_{\mathrm{intra}} = 0$ at all times and $\bar{\omega} > \bar{a}$,*

$$x_{inter}(t) = \begin{cases} h(t), & \text{if } t \neq t_0 + kT, \ k \in \mathbb{Z}, \\ \pi, & \text{if } t = t_0 + kT, \ k \in \mathbb{Z}. \end{cases}$$

$$\triangleq x_{\mathrm{nom}}(t), \tag{6.15}$$

where

$$h(t) = 2 \tan^{-1}\left(\frac{\bar{a} + \sqrt{\bar{\omega}^2 - \bar{a}^2} \, \tan\left(\frac{\sqrt{\bar{\omega}^2 - \bar{a}^2}}{2}(t + \tau)\right)}{\bar{\omega}}\right),$$

$t_0 = -\tau + \pi/\sqrt{\bar{\omega}^2 - \bar{a}^2}$, $T = 2\pi/\sqrt{\bar{\omega}^2 - \bar{a}^2}$, *and* $\tau \in \mathbb{R}$ *is a constant that depends only on $\theta(0)$.*

Moreover,

  (i) x_{nom} is T-periodic with zero time average, and

  (ii) the following inequality holds:

$$\left| \int_0^t \cos(x_{nom}(\tau)) \, \mathrm{d}\tau \right| \leq \frac{1}{\bar{a}} \log\left(\frac{\bar{\omega} + \bar{a}}{\bar{\omega} - \bar{a}}\right). \tag{6.16}$$

Remark 30 *(Constant versus time-varying inter-cluster difference) The values of $\bar{\omega}$ and $\bar{a}$ determine the behavior of the inter-cluster phase difference. In particular, if $\bar{\omega} < \bar{a}$, then the inter-cluster difference evolves as in (6.15).[8] If $\bar{\omega} = \bar{a}$, (6.4) reduces to $\dot{x}_{inter} = \bar{a} - \bar{a} \sin(x_{inter})$, which can be integrated:*

$$\bar{a}t = \int_{x_{inter}(0)}^{x_{inter}(t)} (1 - \sin(s))^{-1} \mathrm{d}s$$

$$\bar{a}t = \frac{2 \sin(x_{inter}(t)/2)}{\cos(x_{inter}(t)/2) - \sin(x_{inter}(t)/2)} + \tau. \tag{6.17}$$

By substitution, it can be verified that

$$x_{inter}(t) = 2 \cos^{-1}\left(\frac{\bar{a}t - \tau + 2}{\sqrt{2(\bar{a}t - \tau + 1)^2 + 2}}\right),$$

---

[8]In fact, $\sqrt{\bar{\omega}^2 - \bar{a}^2}$ becomes a complex number and, by recalling that $\tan(\mathsf{i}\alpha) = \mathsf{i}\tanh(\alpha)$, where $\alpha \in \mathbb{R}$, in (6.15) we have $x_{\mathrm{inter}}(t) = 2\tan^{-1}((\bar{a} - \sqrt{\bar{a}^2 - \bar{\omega}^2} \, \tanh(\sqrt{\bar{a}^2 - \bar{\omega}^2}(t + \tau)/2))/\bar{\omega})$.
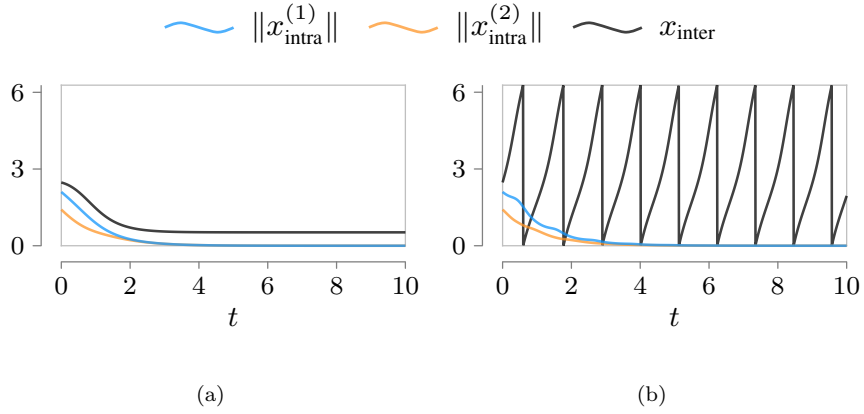
Figure 6.8: For the network in Example 36 with $\alpha_1 = \alpha_2 = \beta = 1$, $\bar{a} = 2$ and $\bar{\omega} = 1$, Fig. 6.8(a) shows that the clusters are synchronized (as $\|x_{\text{intra}}^{(1)}\|$ and $\|x_{\text{intra}}^{(2)}\|$ converge to zero), yet all oscillators remain phase locked (x_{inter} converges to a constant). Instead, Fig. 6.8(b) shows that the inter-cluster difference follows a limit cycle when $\alpha_1 = \alpha_2 = \beta = 1$, $\bar{a} = 2$ and $\bar{\omega} = 6$.

satisfies equation (6.17). In both cases ($\bar{\omega} \leq \bar{a}$), x_{inter} converges to the constant value $2\tan^{-1}((\bar{a} -$

$\sqrt{\bar{a}^2 - \bar{\omega}^2})/\bar{\omega})$ as t increases to infinity. In other words, if $\bar{\omega} \leq \bar{a}$, then the phases of the oscillators in

the two clusters evolve with the same frequency, and the oscillators are phase locked (see Fig. 6.8(a)

and [28, Remark 1]). Instead, if $\bar{\omega} > \bar{a}$, the clusters evolve with different frequencies, and the

inter-cluster phase difference follows a limit cycle (see Fig. 6.8(b) and [194, Chapter 2]). □

In the remainder of this section we assume that $\bar{\omega} > \bar{a}$, so that the clusters evolve with

different frequencies (see Remark 30). Leveraging Lemma 29, we next present a refined condition

for the stability of the cluster synchronization manifold.

Theorem 31 *(Sufficient condition on network weights and natural frequencies for the*

stability of $\mathcal{S}_{\mathcal{P}}$) Let $\mathcal{S}_{\mathcal{P}}$ be the cluster synchronization manifold associated with a partition $\mathcal{P} =$

$\{\mathcal{P}_1, \mathcal{P}_2\}$ of the network $\mathcal{G}$ of Kuramoto oscillators. Let $\omega_i \in \mathbb{R}_{>0}$ be the natural frequency of

the oscillators in the cluster $\mathcal{P}_i$, with $i \in \{1, 2\}$. Let J_{intra} be as in Lemma 24, and $J_{inter} =$

$\partial G(x_{intra}, x_{inter})/\partial x_{intra}$ along the trajectory $x_{intra} = 0$ and $x_{inter} = x_{nom}$. The cluster synchroniza-

tion manifold $\mathcal{S}_{\mathcal{P}}$ is locally exponentially stable if the following inequality holds:

$$\left(\frac{\bar{\omega} + \bar{a}}{\bar{\omega} - \bar{a}}\right)^{\frac{2}{\bar{a}}\|J_{inter}\|} < 1 + \frac{1}{2\lambda_{max}(X)\|J_{intra}\|}, \tag{6.18}$$

where $X \succ 0$ is the solution of $J_{intra}^{\mathsf{T}}X + XJ_{intra} = -I$.

Theorem 31 provides a quantitative condition on the network weights and the natural frequencies of the oscillators to ensure stability of the cluster synchronization manifold. It can be shown that (i) when the inter-cluster weights decrease to zero ($\bar{a} \to 0$) and $\bar{\omega}$ remains bounded, then $\|J_{\text{inter}}\|/\bar{a}$ remains bounded, the left-hand side of (6.18) converges to 1, and the inequality is automatically satisfied, and (ii) when $\bar{\omega}$ grows ($\bar{\omega} \to \infty$) and the inter-cluster weights remain bounded, the left-hand side of (6.18) converges to 1 and the inequality is automatically satisfied. The role of the intra-cluster connections on the stability of $\mathcal{S}_{\mathcal{P}}$ cannot be evaluated directly from (6.18) because of the dependency of the right-hand side on $\lambda_{\text{max}}(X)$. The following result, however, suggests that the synchronization manifold may remain exponentially stable when the intra-cluster weights are homogeneous, independently of the inter-cluster weights and the natural frequencies.

Theorem 32 *(Stability of $\mathcal{S}_{\mathcal{P}}$ with homogeneous clusters) Let $\mathcal{S}_{\mathcal{P}}$ be the cluster synchronization manifold associated with a partition $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2\}$ of the network $\mathcal{G}$ of Kuramoto oscillators. Let $\omega_i \in \mathbb{R}_{>0}$ be the natural frequency of the oscillators in the cluster $\mathcal{P}_i$, with $i \in \{1, 2\}$. If $J_{\text{intra}} = \alpha I$, for some constant $\alpha \in \mathbb{R}_{<0}$, then the cluster synchronization manifold $\mathcal{S}_{\mathcal{P}}$ is locally exponentially stable.*

We provide an example that illustrates the stability conditions derived in Theorem 31.

Example 33 *(Heterogeneity of natural frequencies improves stability of the cluster synchronization manifold) Consider the network of Kuramoto oscillators in Example 36. Fig. 6.9(a) illustrates that the cluster synchronization manifold is asymptotically stable when the condition in Theorem 31 is satisfied. Fig. 6.9(b) illustrates the tradeoff in the latter stability condition between*
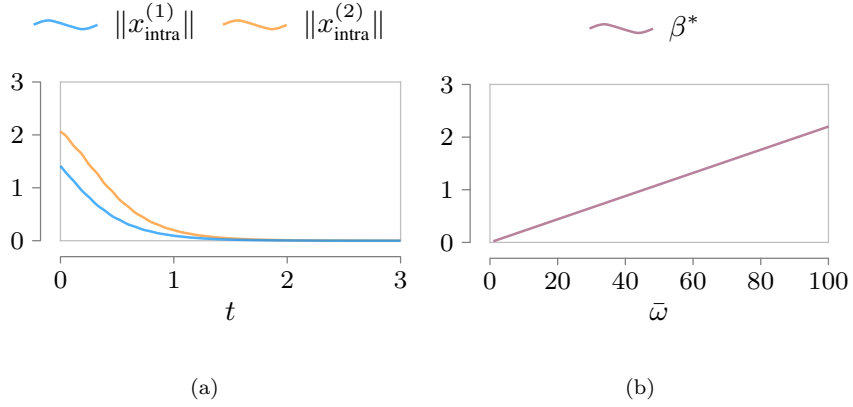
Figure 6.9: For the network in Example 36, Fig. 6.9(a) illustrates the stability of $\mathcal{S}_\mathcal{P}$ when $\alpha_1 = \alpha_2 = \beta = \omega_1 = 1$ and $\omega_2 = 47$, as predicted by the condition in Theorem 31. For the same network and weights, Fig. 6.9(b) shows the largest value of inter-cluster weights β^* that satisfies (6.18) with equality. As predicted by Theorem 28 and Theorem 31, stability of the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$ is preserved when $\bar{\omega}$ grows with the inter-cluster weights.

the natural frequency $\bar{\omega}$ and the inter-cluster strength measured by β^*, which denotes the largest inter-cluster weight β (see Example 36) such that (6.18) is still satisfied. Further, we show in Fig. 6.10 that, while being conservative, condition (6.18) captures the fact that stability of the cluster synchronization manifold can be recovered by increasing $\bar{\omega}$. Namely, choosing the same network weights that yield instability as in Fig. 6.6(d), we show that stability of the cluster synchronization manifold is recovered as the difference in natural frequencies grows. □

We conclude this section with a discussion of cluster synchronization in asymmetric networks and identical nodes.

Remark 34 *(Extension to networks with asymmetric weights) Symmetry of the network weights is typically exploited to provide conditions for the stability of the full synchronization manifold in networks of Kuramoto oscillators [28]. We rely on the symmetry assumption (A1) to derive statement (i) in Lemma 24, which supports our main theorems. However, these results remain valid for bidirected graphs,[9] provided that the Jacobian J_{intra} can be proven to be Hurwitz. In other words, Assumption (A1) is used to guarantee stability of the isolated clusters, and not of the cluster*

---

[9] A bidirected graph is a directed graph where $(i,j) \in \mathcal{E}$ implies $(j,i) \in \mathcal{E}$. The adjacency matrix of a bidirected graph needs not be symmetric.
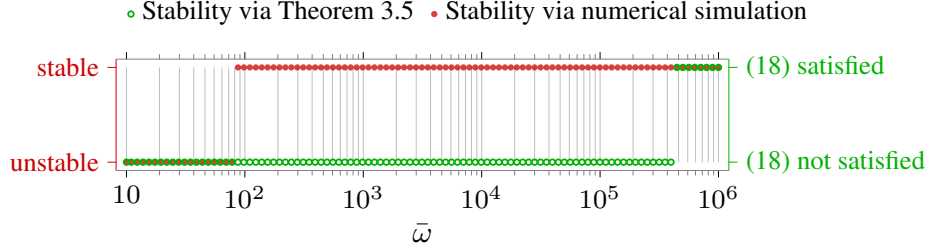
Figure 6.10: For the network in Example 36, we let $\alpha_1 = \beta = 1$ and $\alpha_2 = 10^{-4}$ and plot, as a function of $\bar{\omega}$, the stable configurations predicted by Theorem 31 (green) and those found numerically. For each value of $\bar{\omega}$, we assess numerical stability (in red) by making use of the Floquet stability theory [197, Chapter 5] and by resorting to statement (i) in Lemma 29. This is possible because the partition in Example 36 contains two clusters. Although condition (6.18) is conservative, it captures the effect of large $\bar{\omega}$ on the stability of $\mathcal{S}_{\mathcal{P}}$.
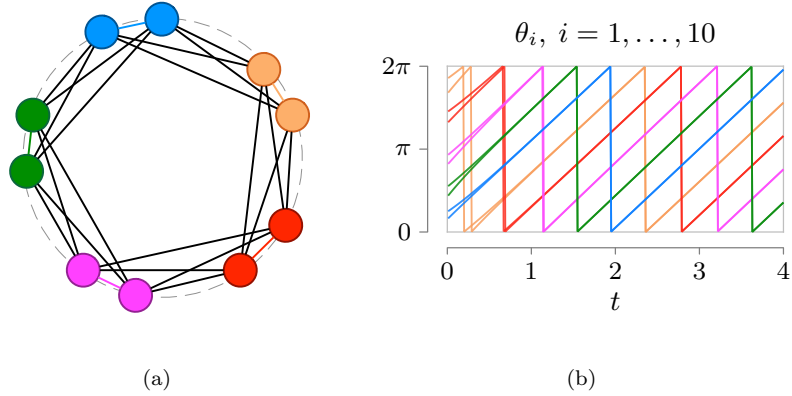
Figure 6.11: Fig. 6.11(a) illustrates a network of 10 Kuramoto oscillators with partition $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4, \mathcal{P}_5\}$, where each cluster is color-coded. All oscillators have identical natural frequency $\omega = 3$ and all edges have unit weight. As illustrated in Fig. 6.11(b), the cluster synchronization manifold associated to $\mathcal{P}$ is stable, showing that cluster synchronization is possible even in networks of identical Kuramoto oscillators with identical edge weights.

configuration. □

Remark 35 *(Cluster synchronization in networks of identical oscillators) This chapter*

focuses on heterogeneous oscillators and leverages mismatches in the natural frequencies and the

network weights to characterize the stability of the cluster synchronization manifold. Yet, cluster

synchronization can also arise in networks of homogeneous Kuramoto oscillators, where all units

have equal natural frequencies and all edges have equal weight (e.g., see Fig. 6.11). With the

exception of Theorem 28, which is also applicable in the case of identical edge weights, our stability

results cannot predict cluster synchronization in networks of identical oscillators, a question that we

leave as the subject of future investigation. □

6.4 Approximate conditions for the stability of cluster synchronization

In this section, we introduce approximate conditions for the stability of the cluster synchronization manifold in networks of heterogeneous Kuramoto oscillators. Before proceeding, we introduce a toy example to compare the exact stability conditions derived in the previous sections.

Example 36 *(**Comparison between stability conditions**) Consider the network in Fig. 6.12(a) with partition $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2\}$, where $\mathcal{P}_1 = \{1, 2\}$ and $\mathcal{P}_2 = \{3, 4\}$, and adjacency matrix as in Fig. 6.12(b). The parameters $\alpha_1, \alpha_2 \in \mathbb{R}_{>0}$ and $\beta \in \mathbb{R}_{>0}$ denote the intra- and inter-cluster couplings, respectively. The matrix S in Theorem 25 becomes $S = \begin{bmatrix} 4\alpha_1 - 2\beta & -2\beta \\ -2\beta & 4\alpha_2 - 2\beta \end{bmatrix}$. In Fig. 6.12(c) we compare stability conditions based on the matrix S in (7.3) and numerical conditions given by Floquet stability theory.[10] Notice that, for certain parameters, the synchronization manifold is unstable. Finally, Fig. 6.12(d) shows that, as the inter-cluster coupling β grows, the stability of $\mathcal{S}_{\mathcal{P}}$ is achieved by increasing the difference of the natural frequencies, as predicted by Theorem 28.* □

Stability of cluster synchronization is guaranteed by the stability of the time-varying *interconnected* system (6.33), where the subsystems are identified by the isolated clusters. For interconnected systems with time-varying dynamics

$$\dot{y}_i = A_i(t)y_i + \sum_{j \neq i} B_{ij}(t)y_j, \quad i = 1, \ldots, m, \tag{6.19}$$

where $A_i(t)$ and $B_{ij}(t)$ are time-varying matrices of suitable dimensions, a simplified version of the small-gain theorem can readily be derived from [199] and reads as follows.

---

[10]This comparison is possible for the case of two clusters, as we are able to explicitly derive the trajectory of the periodic inter-cluster difference.
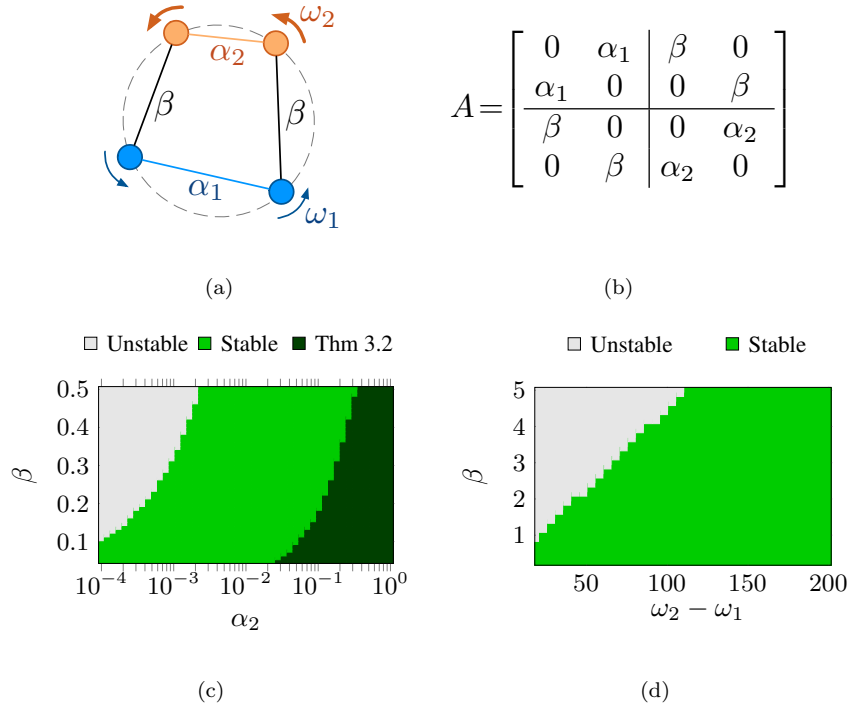
Figure 6.12: Fig. 6.12(a) illustrates the network of 4 Kuramoto oscillators in Example 36. We identify the clusters $\mathcal{P}_1$ and $\mathcal{P}_2$ in blue and orange, respectively. Fig. 6.12(c) compares the stability conditions of Theorem 25 and numerical stability via Floquet decomposition [197]. We fix $\alpha_1 = \omega_1 = 1$, $\omega_2 = 8$, and let α_2 and β vary. The condition in Theorem 25 (dark green) identifies a subset of the stable configurations (light green). Fig. 6.12(d) illustrates that stability is guaranteed for sufficiently heterogeneous natural frequencies, as predicted by Theorem 28 ($\alpha_1 = \omega_1 = 1$, $\alpha_2 = 0.001$).

Theorem 37 *(Small-gain stability test [199]) The origin of the system in* (6.19) *is (globally) exponentially stable if:*

(i) *the origin of each isolated subsystem* $\dot{y}_i = A_i(t)y_i$ *is (globally) exponentially stable;*

(ii) *there exist* $\xi_{ij} \in \mathbb{R}_{\geq 0}$ *(gains) s.t.,* $\forall\, t \geq 0$, $\|y_{i,f}(t)\| \leq \sum_{j \neq i} \xi_{ij} \sup_{\tau \in [0,t]} \|y_j(\tau)\|$, $i = 1, \ldots, m$, *where* $y_{i,f}(t)$ *is the forced response of the i-th subsystem;*

(iii) *the matrix* $\Xi \triangleq [\xi_{ij}] \in \mathbb{R}^{m \times m}$ *(gain matrix) satisfies*

$$\lambda_{max}(\Xi) < 1. \tag{6.20}$$

As conditions (i)-(ii) in Theorem 37 are generally difficult to verify in systems with time-varying dynamics, in what follows we propose and validate an approximation to the dynamics (6.33),

which allows us to apply Theorem 37 and derive (approximate) stability conditions for the cluster synchronization manifold. As we will show through numerical examples, our approximate stability conditions are tight.

To simplify the derivation of our approximate stability conditions, we start with the case of clusters with two nodes, so that the dynamics of the intra-cluster phase difference is scalar. This procedure extends directly to the general case. To obtain our stability condition, we follow three main steps.

(Approximate input-output response of the system (6.33)) The scalar intra-cluster dynamics of the k-th cluster reads as

$$\dot{x}^{(k)}_{\text{intra}} = (J_k + J_{\text{inter},k})x^{(k)}_{\text{intra}} + \sum_{\ell \neq k} J_{\text{inter},k\ell} x^{(\ell)}_{\text{intra}}, \tag{6.21}$$

where J_k, $J_{\text{inter},k} = \sum_{\ell \neq k} \eta_{k\ell} \cos(x^{(k\ell)})$, $J_{\text{inter},\ell z} = \zeta_{\ell z} \cos(x^{(\ell z)})$, $\eta_{k\ell}$, and $\zeta_{\ell z}$ are scalar quantities derived from (6.33). As $\omega^{(\ell z)} \triangleq \omega_z - \omega_\ell$ grows, we have [11] $x^{(\ell z)}(t) \approx \omega^{(\ell z)}t$. Thus, $J_{\text{inter},k} \approx \sum_{\ell \neq k} \eta_{k\ell} \cos(\omega^{(k\ell)}t)$, $J_{\text{inter},\ell z} \approx \zeta_{\ell z} \cos(\omega^{(\ell z)}t)$, and the approximate forced response of (6.21) to the input $x^{(\ell)}_{\text{intra}}$ becomes

$$x^{(k)}_{\text{intra},f} \approx \sum_{\ell \neq k} \int_0^t e^{J_k(t-\tau)} \zeta_{k\ell} \cos(\omega^{(k\ell)}\tau) x^{(\ell)}_{\text{intra}}(\tau) \, \mathrm{d}\tau, \tag{6.22}$$

where we have approximated the transition matrix of the linear time-varying system (6.21) as (see also [197]) $e^{\int_0^t J_k + J_{\text{inter},k}\mathrm{d}\tau} \approx e^{\int_0^t J_k \mathrm{d}\tau} = e^{J_k t}$. The last approximation is motivated by the fact that $J_{\text{inter},k}$ is a high-frequency signal with zero mean, so that its integral over time becomes negligible compared to the integral of J_k. Our approximation is validated in Fig. 6.13 for the network in Fig. 6.12(a), where we see that the input-output responses of the original (6.21) and approximated (6.22) systems remain close to each other for different values of the natural frequency.

(Computation of approximate input-output gains) Equation (6.22) can be viewed as the forced response of a linear time-invariant system with matrix J_k to the modulated input $\zeta_{k\ell}\cos(\omega^{(k\ell)}\tau)x^{(\ell)}_{\text{intra}}$.

---

[11]This approximation is reasonable for heterogeneous natural frequencies. The same approximation has been used also in [181], although for a different analysis of cluster synchronization.
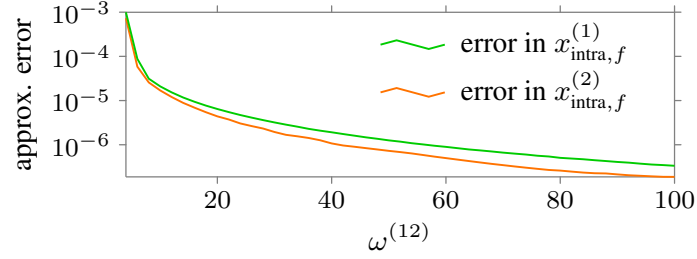
Figure 6.13: In this figure we plot the maximum error between the forced response of (6.21) and the proposed approximation in (6.22) as $\omega^{(12)}$ increases. For the simulation, $\alpha_1 = \beta_1 = \omega_1 = 1$, $\alpha_2 = 0.01$, and ω_2 varies as indicated. Initial conditions are chosen randomly in the interval $(0, 0.001]$.

Following [200], each term in the sum in (6.22) can be expressed as a Taylor series about the natural frequency of the modulating function. This yields

$$x^{(k)}_{\text{intra}, f}(t) = \sum_{\ell \neq k} \zeta_{k\ell} \underbrace{\sum_{r=0}^{\infty} c_r^{(k\ell)} \cos(\omega^{(k\ell)} t + \psi_r^{(k\ell)}) \frac{\mathrm{d}^r x^{(\ell)}_{\text{intra}}(t)}{\mathrm{d}t^r}}_{C_{k\ell}} \tag{6.23}$$

with $H_k(s) \triangleq \frac{1}{s - J_k}$, $c_r^{(k\ell)} = \left| \frac{\mathrm{i}^{-r}}{r!} \frac{\mathrm{d}^r H_k(\mathrm{i}\,\omega)}{\mathrm{d}\omega^r} \right|_{\omega = \omega^{(k\ell)}}$, and $\psi_r^{(k\ell)} = \text{angle}\left(\frac{\mathrm{i}^{-r}}{r!} \frac{\mathrm{d}^r H_k(\mathrm{i}\,\omega)}{\mathrm{d}\omega^r} \Big|_{\omega = \omega^{(k\ell)}} \right)$.

We propose the following first-order approximation of (6.23):

$$|C_{k\ell}| \approx \begin{cases} |H_k(\mathrm{i}\,\omega^{(k\ell)})| \; |x^{(\ell)}_{\text{intra}}(t)|, & \text{if } J_k \leq J_\ell, \\[2em] \frac{|H_k(0)|}{|H_\ell(0)|} |H_\ell(\mathrm{i}\,\omega^{(k\ell)})| \; |x^{(\ell)}_{\text{intra}}(t)|, & \text{if } J_\ell < J_k. \end{cases} \tag{6.24}$$

Loosely speaking, the former approximation is motivated by the fact that the modulated input coming from the ℓ-th system is "slower" than the k-th system. Instead, the second approximation is valid when the input from the ℓ-th system is "faster" than the k-th system, and it follows from [200]. Using (6.24), we define the approximate input-ouput gains $\xi_{k\ell}$, $k \neq \ell$:

$$\xi_{k\ell} = \begin{cases} |H_k(\mathrm{i}\,\omega^{(k\ell)})||\zeta_{k\ell}|, & \text{if } J_k \leq J_\ell, \\[2em] \frac{|H_k(0)|}{|H_\ell(0)|} |H_\ell(\mathrm{i}\,\omega^{(k\ell)})||\zeta_{k\ell}|, & \text{if } J_\ell < J_k. \end{cases} \tag{6.25}$$

Our approximation is validated in Fig. 6.14 for the network in Fig. 6.12(a), where it can be seen that, for the considered example, the proposed gains are accurate when $J_\ell > J_k$, and they constitute
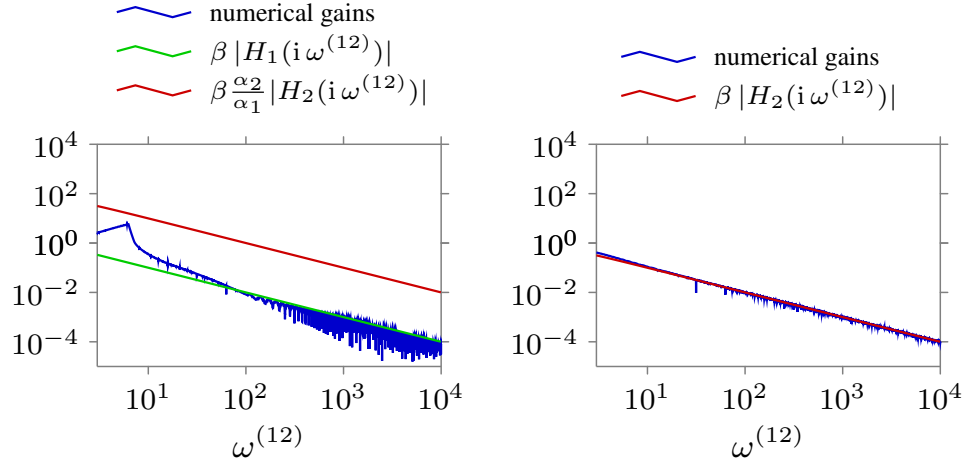
Figure 6.14: In this figure we plot the gains ξ_{12} (left panel) and ξ_{21} (right panel) in a logarithmic scale. For this comparison, we select $\beta = \alpha_2 = 1$ and $\alpha_1 = 0.01$; thus, $\zeta_{12} = \zeta_{21} = \beta$, $J_1 = -2\alpha_1$, and $J_2 = -2\alpha_2$. In the left panel, the additional green line represents the approximate gain obtained by truncating the series in (6.23) after the first term. Notice that, in such a case, the approximation is accurate only for large frequencies $\omega^{(12)}$. The red curve, instead, represent the proposed approximation detailed in (6.25).

a reasonable upper bound when $J_\ell < J_k$. Additional studies are required to further validate and support the proposed approximate gains (see also Example 38).

(Approximate stability test) We define the gain matrix $\Xi = [\xi_{k\ell}] \in \mathbb{R}^{m \times m}$, with $\xi_{k\ell} = 0$ if $k = \ell$, and $\xi_{k\ell}$ as in (6.25) if $k \neq \ell$. Finally, our approximate stability condition for the cluster synchronization manifold consists of applying Theorem 37 with the approximate gain matrix Ξ.

Example 38 *(Tightness of approximate stability condition) Consider the network in Example 36. In Fig. 6.15 we compare the proposed approximate stability condition with the numerical outcomes from Floquet stability theory. In Fig. 6.15(a) we use the following quantities: ω_{sg}^* is the smallest frequency difference $\omega^{(12)}$ such that condition (6.20) with gains as in (6.25) is satisfied, and ω_f^* is the smallest frequency difference $\omega^{(12)}$ such that the largest Floquet exponent of (6.33) is negative. Notice that the approximate condition in (6.20) closely reproduces the numerical instability-stability transition in a reasonable region of intra- and inter-cluster parameters; namely, for small values of β and $\alpha_2 - \alpha_1$. In Fig. 6.15(b) we show that our approximate stability condition*
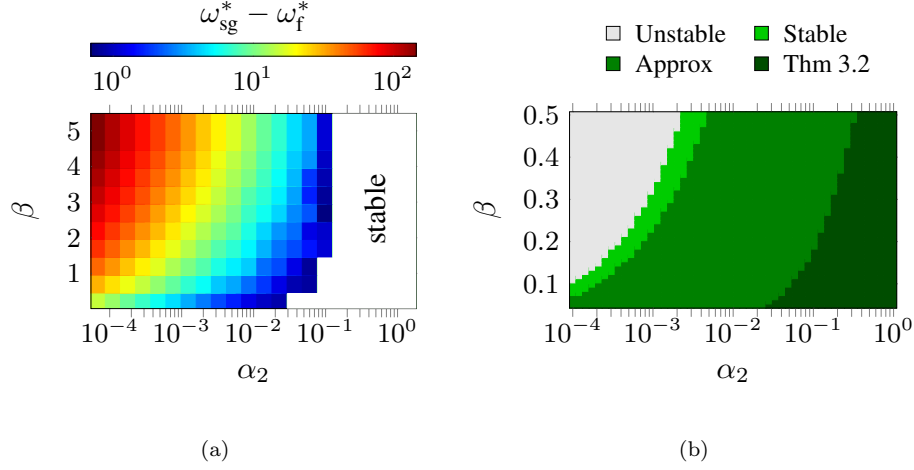
Figure 6.15: The heatmap in Fig. 6.15(a) represents the error between ω_{sg}^* and ω_{f}^* defined in Example 38. We let $\alpha_1 = \omega_1 = 1$, and β, α_2 vary as indicated. The white area contains stable network realizations predicted by both the Floquet exponents and the condition (6.20). Fig. 6.15(b) shows that, for the parameters in Example 36, the approximate stability condition based on Theorem 37 and (6.25) is much tighter than the condition in Theorem 25.

outperforms the analytical one derived in Theorem 25. □

Remark 39 *(Extension to clusters with multiple nodes) When the clusters contain more than*

two nodes, a gain matrix similar to (6.25) can be computed by using a suitable scalar approximation

of the transfer matrix $H_k(s) = (sI - J_k)^{-1}$. We propose the following stability test, which relies on

using the mean singular value of H_k to compute input-output gains:

$$
\xi_{k\ell} = \begin{cases} \nu_{k\ell}\overline{\sigma}(H_k(\mathrm{i}\,\omega^{(k\ell)})), & \text{if } \overline{\lambda}(J_k) \leq \overline{\lambda}(J_\ell), \\[2ex] \nu_{k\ell}\dfrac{\overline{\sigma}(H_k(0))}{\overline{\sigma}(H_\ell(0))}\overline{\sigma}(H_\ell(\mathrm{i}\,\omega^{(k\ell)})), & \text{if } \overline{\lambda}(J_\ell) < \overline{\lambda}(J_k), \end{cases}
\tag{6.26}
$$

where $\nu_{k\ell} = \|J_{inter,k\ell}\|$ at time $t = 0$. We validate the approximate stability condition (6.20) on

random networks with $2N$ nodes that are generated as follows. First, we choose two weighted undi-

rected and connected Erdös-Rényi graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ with cardinality N and edge probability $p = 0.5$.

To facilitate instability of $\mathcal{S}_{\mathcal{P}}$ for small natural frequency differences, we select small intra-cluster

weights in $\mathcal{G}_2$ (see Fig. 6.15(a)). Second, we connect $\mathcal{G}_1$ and $\mathcal{G}_2$ to satisfy Assumption (A2). Finally,

for each N, we compare the condition (6.20) with our approximate gains (9.2) and the condition

in Theorem 28 with the smallest value of the natural frequency ensuring stability, which is obtained
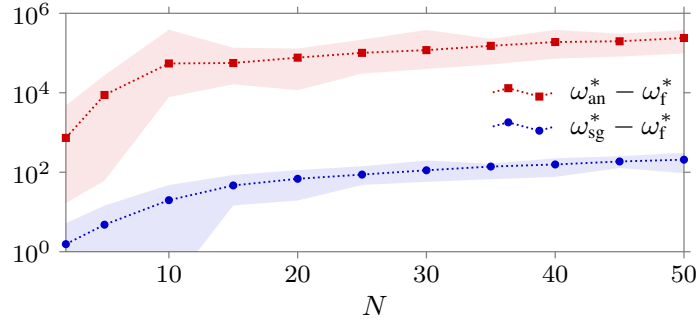
Figure 6.16: This figure shows the error between ω_{sg}^* and ω_{f}^* defined in Example 38 and using the approximate gain matrix (9.2), as the cardinality of the clusters increases. For each cardinality N, we generate 100 random realizations of the graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ described in Remark 39. The intra-cluster weights of $\mathcal{G}_1$ (resp. $\mathcal{G}_2$) are uniformly distributed in $[0, 1]$ (resp. $[0, 0.01]$), and the inter-cluster weights are uniformly distributed in $[0, 1]$. The dashed blue line represents the mean difference $\omega_{\text{sg}}^* - \omega_{\text{f}}^*$ for the approximate gains in (9.2), and the dashed red line represents the mean difference $\omega_{\text{an}}^* - \omega_{\text{f}}^*$ for the analytical test in Theorem 28. The shaded area contains the maximum and minimum values among all realizations.

using the Floquet exponents of (6.33). For different random realizations of $\mathcal{G}_1$ and $\mathcal{G}_2$, Fig. 6.16

shows that our heuristic test consistently performs better than its analytical counterpart. □

6.5 Conclusion

In this chapter we characterize conditions for the stability of cluster synchronization in networks of oscillators with Kuramoto dynamics, where multiple synchronized groups of oscillators coexist in a connected network. We derive conditions showing that the cluster synchronization manifold is locally exponentially stable when (i) the intra-cluster coupling is sufficiently stronger that the inter-cluster coupling, (ii) the differences of natural frequencies of the oscillators in disjoint clusters are sufficiently large, (iii) a combination of the two, or, (iv) in the case of two clusters, if the intra-cluster dynamics is homogeneous. To the best of our knowledge, our results are the first to characterize the stability of the cluster synchronization manifold in sparse and weighted networks of heterogeneous Kuramoto oscillators.

6.6 Appendix

In this section we provide the proofs of the results presented in Section 9.3, together with some instrumental lemmas.

6.6.1 Proofs of the results in Section 6.3.1

Proof of Lemma 24: *Proof of statement (i).* Notice that the block-diagonal form of the Jacobian matrix J_{intra} follows directly from the form of $F(x_{\text{intra}})$ in (6.8). Therefore, the stability of J_{intra} is equivalent to the stability of the diagonal blocks J_k. Let $\theta^{(k)}$ be the vector of θ_i, $i \in \mathcal{P}_k$ and, by Assumption (A2), let ω_k be the natural frequency of any oscillator in $\mathcal{P}_k$. From (9.1), we write the phase dynamics of the k-th cluster as (see [187])

$$\dot{\theta}^{(k)} = \omega_k \mathbb{1} - B_k \operatorname{diag}(\{a_{ij}\}_{(i,j)\in\mathcal{E}_k}) \sin(B_k^\mathsf{T} \theta^{(k)}).$$

Because the phase differences satisfy $x_{\text{intra}}^{(k)} = B_{\text{span},k}^\mathsf{T} \theta^{(k)}$ and $x^{(k)} = B_k^\mathsf{T} \theta^{(k)}$, we have

$$\dot{x}_{\text{intra}}^{(k)} = -B_{\text{span},k}^\mathsf{T} B_k \operatorname{diag}(\{a_{ij}\}_{(i,j)\in\mathcal{E}_k}) \sin(x^{(k)}), \qquad (6.27)$$

where we have used the property $B_{\text{span},k}^\mathsf{T} \mathbb{1} = 0$. Using (6.5), the Jacobian matrix of (6.27) computed at $x_{\text{intra}}^{(k)} = 0$ reads as

$$J_k = -B_{\text{span},k}^\mathsf{T} B_k \operatorname{diag}(\{a_{ij}\}_{(i,j)\in\mathcal{E}_k}) T_{\text{intra},k}. \qquad (6.28)$$

Recall that the Laplacian matrix of the graph $\mathcal{G}_k$ satisfies $\mathcal{L}_{\mathcal{G}_k} = B_k \operatorname{diag}(\{a_{ij}\}_{(i,j)\in\mathcal{E}_k}) B_k^\mathsf{T}$, and that, because $\mathcal{G}_k$ is connected, the eigenvalues of $-\mathcal{L}_{\mathcal{G}_k}$ have negative real part, except one single eigenvalue located at the origin with eigenvector $\mathbb{1}$. Define the matrix $W_k = [B_{\text{span},k} \ \mathbb{1}]^\mathsf{T}$ and notice that, because $B_{\text{span},k}^\mathsf{T} \mathbb{1} = 0$ and $B_{\text{span},k}$ being full column rank [42, Theorem 8.3.1], then W_k is invertible and $W_k^{-1} = [(B_{\text{span},k}^\mathsf{T})^\dagger \ (\mathbb{1}^\mathsf{T})^\dagger]$. Therefore we have

$$W_k(-\mathcal{L}_{\mathcal{G}_k})W_k^{-1} = \begin{bmatrix} J_k & 0 \\ 0 & 0 \end{bmatrix},$$

where we have used that $T_{\text{intra},k} = B_k^\mathsf{T}(B_{\text{span},k}^\mathsf{T})^\dagger$ in (6.28). This shows that J_k contains only the stable eigenvalues of $-\mathcal{L}_{\mathcal{G}_k}$.

Proof of statement (ii). Notice that, for any $(j,z) \in \mathcal{E}$ with $j \in \mathcal{P}_k$, $z \in \mathcal{P}_\ell$, and $k \neq \ell$, the difference $\text{diff}(\text{p}(j,z))$ in $G_{ij}^{(k)}(x_{\text{inter}}, x_{\text{inter}})$ in equation (6.6) can be rewritten as

$$\text{diff}(\text{p}(j,z)) = \text{diff}(\text{p}(j,k^*)) + \text{diff}(\text{p}(k^*,\ell^*)) + \text{diff}(\text{p}(\ell^*,z)),$$

where k^* and ℓ^* are such that $\text{p}(k^*,\ell^*)$ is the shortest path on $\mathcal{T}$ connecting the clusters $\mathcal{P}_k$ and $\mathcal{P}_\ell$. Then,

$$G_{ij}^{(k)}(x_{\text{inter}}, x_{\text{inter}}) = \sum_{\substack{\ell=1 \\ \ell \neq k}}^{m} \sum_{z \in \mathcal{P}_\ell} [a_{jz}\sin(\text{diff}(\text{p}(j,k^*)) + \text{diff}(\text{p}(k^*,\ell^*)) + \text{diff}(\text{p}(\ell^*,z)))$$

$$-a_{iz}\sin(\text{diff}(\text{p}(i,k^*)) + \text{diff}(\text{p}(k^*,\ell^*)) + \text{diff}(\text{p}(\ell^*,z)))].$$

Notice that $\text{diff}(\text{p}(i,k^*))$ and $\text{diff}(\text{p}(j,k^*))$ contain only differences in $x_{\text{intra}}^{(k)}$, and $\text{diff}(\text{p}(\ell^*,z))$ only differences in $x_{\text{intra}}^{(\ell)}$.

Notice that $\sin(a+b) = \sin(a) + \delta$, with $|\delta| \leq |b|$.[12] Then,

$$G_{ij}^{(k)}(x_{\text{intra}}, x_{\text{inter}}) = \sum_{\substack{\ell=1 \\ \ell \neq k}}^{m} \sum_{z \in \mathcal{P}_\ell} [a_{jz}(\sin(\text{diff}(\text{p}(k^*,\ell^*))) + \delta_{jz}) - a_{iz}(\sin(\text{diff}(\text{p}(k^*,\ell^*))) + \delta_{iz})]$$

$$= \sum_{\substack{\ell=1 \\ \ell \neq k}}^{m} \left(\sum_{z \in \mathcal{P}_\ell} [(a_{jz} - a_{iz})\sin(\text{diff}(\text{p}(k^*,\ell^*)))] + \sum_{z \in \mathcal{P}_\ell} [a_{jz}\delta_{jz} - a_{iz}\delta_{iz}] \right) \overset{\text{(A3)}}{=} \sum_{\substack{\ell=1 \\ \ell \neq k}}^{m} \sum_{z \in \mathcal{P}_\ell} [a_{jz}\delta_{jz} - a_{iz}\delta_{iz}],$$

where δ_{jz} and δ_{iz} are upper bounded by $\sqrt{n_{\text{intra},k}}\|x_{\text{intra}}^{(k)}\| + \sqrt{n_{\text{intra},\ell}}\|x_{\text{intra}}^{(\ell)}\|$. Therefore, we have the following bound:

$$|G_{ij}^{(k)}| \leq \sum_{\substack{\ell=1 \\ \ell \neq k}}^{m} \left(\sum_{z \in \mathcal{P}_\ell} a_{jz}|\delta_{jz}| + \sum_{z \in \mathcal{P}_\ell} a_{iz}|\delta_{iz}| \right)$$

$$\overset{\text{(A3)}}{\leq} 2 \sum_{\substack{\ell=1 \\ \ell \neq k}}^{m} \sum_{z \in \mathcal{P}_\ell} a_{jz} \left(\sqrt{n_{\text{intra},k}}\|x_{\text{intra}}^{(k)}\| + \sqrt{n_{\text{intra},\ell}}\|x_{\text{intra}}^{(\ell)}\| \right) = 2 \sum_{\ell=1}^{m} \sqrt{n_{\text{intra},\ell}} \, \tilde{\gamma}_{ij}^{(k\ell)}\|x_{\text{intra}}^{(\ell)}\|,$$

---

[12]Letting $\delta = \sin(a+b) - \sin(a)$, we have $|\delta| = |2\sin(\frac{b}{2})\cos(a+\frac{b}{2})| \leq |2\sin(\frac{b}{2})|$, from which the inequality $|\delta| \leq |b|$ follows.

where

$$\tilde{\gamma}_{ij}^{(k\ell)} = \begin{cases} \displaystyle\sum_{\substack{\ell=1 \\ \ell\neq k}}^{m} \sum_{z\in\mathcal{P}_\ell} a_{jz}, & \text{if } \ell = k, \\[2em] \displaystyle\sum_{z\in\mathcal{P}_\ell} a_{jz}, & \text{otherwise.} \end{cases}$$

To conclude, $\|G^{(k)}\| \leq \sqrt{n_{\text{intra},k}} \max_{(i,j)\in\mathcal{E}_{\text{span},k}} |G_{ij}^{(k)}|$, and, due to (A3), $\tilde{\gamma}_{ij}^{(k\ell)} = \tilde{\gamma}^{(k\ell)}$ is independent

of i and j. Thus, $\|G^{(k)}\| \leq \sum_{\ell=1}^{m} 2\max_r n_{\text{intra},r}\, \tilde{\gamma}^{(k\ell)} \|x_{\text{intra}}^{(\ell)}\|$, and this concludes the proof. $\blacksquare$

Proof of Theorem 25: The system (6.8) can be viewed as the perturbation via $G(x_{\text{intra}}, x_{\text{inter}})$ of

$\dot{x}_{\text{intra}} = F(x_{\text{intra}})$, which describes the dynamics of m disjoint networks of oscillators:

$$\dot{x}_{\text{intra}}^{(k)} = F^{(k)}(x_{\text{intra}}^{(k)}). \tag{6.29}$$

The origin of each system (6.29) is an exponentially stable equilibrium, which can be shown with

the Lyapunov candidate

$$V_k(x_{\text{intra}}) = x_{\text{intra}}^{(k)\mathsf{T}} P_k x_{\text{intra}}^{(k)},$$

where $P_k \succ 0$ is such that $J_k^\mathsf{T} P_k + P_k J_k = -Q_k$ for $Q_k \succ 0$. In fact, the derivative of V along the

trajectories (6.29) is

$$\dot{V}_k(x_{\text{intra}}^{(k)}) = F^{(k)\mathsf{T}}(x_{\text{intra}}^{(k)}) P_k x_{\text{intra}}^{(k)} + x_{\text{intra}}^{(k)\mathsf{T}} P_k F^{(k)}(x_{\text{intra}}^{(k)})$$

$$= x_{\text{intra}}^{(k)\mathsf{T}} (J_k^\mathsf{T} P_k + P_k J_k) x_{\text{intra}}^{(k)} + O(\|x_{\text{intra}}^{(k)}\|^3), \tag{6.30}$$

and the latter is strictly negative when $\|x_{\text{intra}}^{(k)}\| \leq r$ and $r \in \mathbb{R}_{>0}$ is sufficiently small. Further, it

holds that: (i) $\|\partial V_k/\partial x_{\text{intra}}^{(k)}\| \leq 2\lambda_{\max}(P_k)\|x_{\text{intra}}^{(k)}\|$, (ii) $\dot{V}_k(x_{\text{intra}}^{(k)}) \leq -\lambda_{\min}(Q_k)\|x_{\text{intra}}^{(k)}\|^2$, and (iii)

the perturbation terms $G^{(k)}(x_{\text{intra}}, x_{\text{inter}})$ are linearly bounded in $\|x_{\text{intra}}^{(k)}\|$ following statement (ii)

in Lemma 24.

Consider now the following Lyapunov candidate for (6.8):

$$V(x_{\text{intra}}) = \sum_{k=1}^{m} d_k V_k(x_{\text{intra}}^{(k)}), \quad d_k > 0.$$

From [194, Chapter 9.5] we have:

$$\dot{V}(x_{\mathrm{intra}}) \leq -\frac{1}{2}(DS + S^{\mathsf{T}}D)\|x_{\mathrm{intra}}\|^2, \tag{6.31}$$

where $D = \mathrm{diag}(d_1, \ldots, d_m)$, and S satisfies

$$S = [s_{k\ell}] = \begin{cases} \frac{\lambda_{\min}(Q_k)}{\lambda_{\max}(P_k)} - \gamma^{(kk)} & \text{if } k = \ell, \\[2mm] -\gamma^{(k\ell)} & \text{if } k \neq \ell. \end{cases} \tag{6.32}$$

The origin of (6.8) is locally exponentially stable if S is an M-matrix [194, Lemma 9.7 and Theorem 9.2]. Finally, choosing $Q_k = I$ in (6.32) yields condition (7.3) in Theorem 25. ∎

6.6.2 Proofs of the results in Section 6.3.2

Let $\mathcal{C}$ be the set of connected clusters pairs, that is,

$$\mathcal{C} = \{(\ell, z) \ : \ \exists \, (i, j) \in \mathcal{E} \text{ with } i \in \mathcal{P}_\ell, j \in \mathcal{P}_z, \text{ and } \ell < z\}.$$

With a slight abuse of notation, for any $(\ell, z) \in \mathcal{C}$, we define $x^{(\ell z)} = x_{ij}$, for any node $i \in \mathcal{P}_\ell$ and $j \in \mathcal{P}_z$.

Lemma 40 *(Linearized intra-cluster dynamics) The linearization of the intra-cluster dynamics (6.8) around the trajectory $x_{intra} = 0$ and $x_{inter} = x_{nom}$ reads as follows:*

$$\dot{x}_{intra} = (J_{intra} + J_{\mathrm{inter}}) \, x_{intra}, \tag{6.33}$$

where J_{intra} is defined in Lemma 24, and

$$J_{\mathrm{inter}} = \frac{\partial G}{\partial x_{intra}}\bigg|_{\substack{x_{intra}=0 \\ x_{inter}=x_{nom}}} \triangleq \sum_{(\ell,z)\in\mathcal{C}} \cos(x^{(\ell z)}) \, J_{inter}^{(\ell z)}.$$

Proof. Linearization of (6.8) around the trajectory $(x_{\mathrm{intra}}, x_{\mathrm{inter}}) = (0, x_{\mathrm{nom}})$ yields $\partial F/\partial x_{\mathrm{intra}} = J_{\mathrm{intra}}$ and $\partial G/\partial x_{\mathrm{intra}} = J_{\mathrm{inter}}$. The remaining derivatives vanish. That is, $\partial F/\partial x_{\mathrm{inter}} = 0$ because F does not depend on x_{inter}, and $\partial G/\partial x_{\mathrm{inter}} = 0$ because of Assumption (A3). In fact, for any

129

intra-cluster difference x_{ij} with $i, j \in \mathcal{P}_\ell$, $\ell \in \{1, \ldots, m\}$,

$$\frac{\partial G_{ij}}{\partial x_{\text{intra}}}\bigg|_{\substack{x_{\text{intra}}=0 \\ x_{\text{inter}}=x_{\text{nom}}}} = \sum_{(\ell,z)\in\mathcal{C}} \cos(x^{(\ell z)}) \underbrace{\sum_{k\in\mathcal{P}_z} [a_{jk} - a_{ik}]}_{=0} = 0.$$

This concludes the proof. ∎

We next characterize an asymptotic property of the inter-cluster differences through the following instrumental result.

Lemma 41 (Asymptotic behavior of the inter-cluster dynamics for large frequency differences) *Let* $i \in \mathcal{P}_\ell$, $j \in \mathcal{P}_z$, *and* $\ell \neq z$. *Then, the inter-cluster difference* x_{ij} *satisfies*

$$\lim_{|\omega_j - \omega_i| \to \infty} \frac{x_{ij}(t)}{\omega_j - \omega_i} = t. \tag{6.34}$$

Proof. Let $\bar{\omega}_{ij} = \omega_j - \omega_i$. We rewrite (6.4) as

$$\dot{x}_{ij} = \bar{\omega}_{ij} - (a_{ij} + a_{ji}) \sin(x_{ij}) + \sum_{k \neq i,j} [a_{jk} \sin(x_{jk}) - a_{ik} \sin(x_{ik})]. \tag{6.35}$$

From (6.35), let $\beta = \sum_{k \neq i,j}[a_{jk} + a_{ik}]$, and

$$\dot{\underline{x}}_{ij} = \bar{\omega}_{ij} - (a_{ij} + a_{ji}) \sin(\underline{x}_{ij}) - \beta, \tag{6.36}$$

$$\dot{\bar{x}}_{ij} = \bar{\omega}_{ij} - (a_{ij} + a_{ji}) \sin(\bar{x}_{ij}) + \beta, \tag{6.37}$$

with $\underline{x}_{ij}(0) = \bar{x}_{ij}(0) = x_{ij}(0)$. Integrating (6.36) yields

$$\int_{x_{ij}(0)}^{\underline{x}_{ij}(t)} \frac{dy}{\bar{\omega}_{ij} - (a_{ij} + a_{ji}) \sin(y) - \beta} = \int_0^t d\tau. \tag{6.38}$$

As $|\bar{\omega}_{ij}|$ grows, it holds that $|(a_{ij} + a_{ji}) + \beta| < |\bar{\omega}_{ij}|$. Therefore,

$$\frac{1}{\bar{\omega}_{ij} - (a_{ij} + a_{ji}) \sin(y) - \beta} = \frac{1}{\bar{\omega}_{ij}} \left[\frac{1}{1 - \frac{(a_{ij}+a_{ji})\sin(y)+\beta}{\bar{\omega}_{ij}}} \right] = \frac{1}{\bar{\omega}_{ij}} \sum_{k=0}^{\infty} \left[\frac{(a_{ij} + a_{ji}) \sin(y) + \beta}{\bar{\omega}_{ij}} \right]^k.$$

In view of the latter equality, (6.38) becomes

$$t = \frac{\underline{x}_{ij}(t) - x_{ij}(0)}{\bar{\omega}_{ij}} + \frac{1}{\bar{\omega}_{ij}} \underbrace{\int_{x_{ij}(0)}^{\underline{x}_{ij}(t)} \sum_{k=1}^{\infty} \left[\frac{(a_{ij} + a_{ji}) \sin(y) + \beta}{\bar{\omega}_{ij}} \right]^k dy}_{O(\bar{\omega}_{ij}^{-1})},$$

130

or, equivalently,

$$\underline{x}_{ij}(t) = \bar{\omega}_{ij}\, t + x_{ij}(0) + O\left(\bar{\omega}_{ij}^{-1}\right). \tag{6.39}$$

Similarly, the solution of (6.37) has the form in (6.39). Finally, using the Comparison Principle [194, Lemma 3.4], it holds that $\underline{x}_{ij}(t) \le x_{ij}(t) \le \overline{x}_{ij}(t)$ for all $t \ge 0$. Hence, $\frac{x_{ij}(t)}{\bar{\omega}_{ij}} \to t$ as $|\bar{\omega}_{ij}| \to \infty$ and this concludes the proof. ∎

We are now ready to prove Theorem 28.

Proof of Theorem 28: Consider the Lyapunov candidate $V(x_{\mathrm{intra}}, t) = x_{\mathrm{intra}}^{\mathsf{T}} \Gamma(t) x_{\mathrm{intra}}$, and notice that, using (6.33),

$$\begin{aligned}
\dot{V}(x_{\mathrm{intra}}, t) =\ & \dot{x}_{\mathrm{intra}}^{\mathsf{T}} \Gamma x_{\mathrm{intra}} + x_{\mathrm{intra}}^{\mathsf{T}} \Gamma \dot{x}_{\mathrm{intra}} + x_{\mathrm{intra}}^{\mathsf{T}} \dot{\Gamma} x_{\mathrm{intra}} \\
=\ & x_{\mathrm{intra}}^{\mathsf{T}} \left[J_{\mathrm{intra}}^{\mathsf{T}} \Gamma + \Gamma J_{\mathrm{intra}} + \dot{\Gamma} + \sum_{(\ell,z)\in\mathcal{C}} \cos(x^{(\ell z)}) \left(J_{\mathrm{inter}}^{(\ell z)\mathsf{T}} \Gamma + \Gamma J_{\mathrm{inter}}^{(\ell z)} \right) \right] x_{\mathrm{intra}} + O(\|x_{\mathrm{intra}}\|^3).
\end{aligned} \tag{6.40}$$

Let

$$\dot{\Gamma} = -\sum_{(\ell,z)\in\mathcal{C}} \cos(x^{(\ell z)}) \left(J_{\mathrm{inter}}^{(\ell z)\mathsf{T}} \Gamma + \Gamma J_{\mathrm{inter}}^{(\ell z)} \right). \tag{6.41}$$

When the inter-cluster natural frequencies satisfy $|\omega_i - \omega_j| \to \infty$ for all i, j, then $\Gamma(t) \to \Gamma(0)$ for all times t. In fact, integrating both sides of (6.41) and substituting $\Gamma(t) = \Gamma(0)$ yields

$$\begin{aligned}
\int_0^t \dot{\Gamma}\, d\tau =\ & \Gamma(t) - \Gamma(0) = \Gamma(0) - \Gamma(0) = 0 \\
=\ & -\sum_{(\ell,z)\in\mathcal{C}} \int_0^t \cos(x^{(\ell z)}) \left(J_{\mathrm{inter}}^{(\ell z)\mathsf{T}} \Gamma + \Gamma J_{\mathrm{inter}}^{(\ell z)} \right) d\tau \\
=\ & -\sum_{(\ell,z)\in\mathcal{C}} \left(J_{\mathrm{inter}}^{(\ell z)\mathsf{T}} \Gamma(0) + \Gamma(0) J_{\mathrm{inter}}^{(\ell z)} \right) \int_0^t \cos(x^{(\ell z)})\, d\tau,
\end{aligned}$$

which holds true because $\int \cos(x^{(\ell z)})\, d\tau = 0$ due to Lemma 41. Because J_{intra} is stable, we conclude that, when the inter-cluster natural frequencies satisfy $|\omega_i - \omega_j| \to \infty$ for all i, j, $\dot{\Gamma} = 0$, and there exists $\Gamma(0)$ such that (6.40) is strictly negative. This concludes the proof. ∎

Proof of Lemma 29: When $x_{\text{intra}} = 0$, the differential equation (6.35) reduces to $\dot{x}_{\text{inter}} = \bar{\omega} - \bar{a}\sin(x_{\text{inter}})$, which is a separable differential equation with solution as in (6.15). To show that the period of (6.15) is equal to $T = 2\pi/\sqrt{\bar{\omega}^2 - \bar{a}^2}$, we assume, without loss of generality, that $\tau = 0$. It is easy to see that, because $\tan(t)$ is π-periodic, $x_{\text{nom}}(t) = x_{\text{nom}}(t + 2\pi/\sqrt{\bar{\omega}^2 - \bar{a}^2})$. Further, notice that the variable substitution $z = x_{\text{nom}}$ in $\int_0^t \cos(x_{\text{nom}})\,d\tau$ yields

$$\int_0^t \cos(x_{\text{nom}}(\tau))\,d\tau = \int_{x_{\text{nom}}(0)}^{x_{\text{nom}}(t)} \frac{\cos(z)}{\bar{\omega} - \bar{a}\sin(z)}\,dz = \frac{1}{\bar{a}}\log\left(\frac{\bar{\omega} - \bar{a}\sin(x(0))}{\bar{\omega} - \bar{a}\sin(x_{\text{nom}}(t))}\right), \qquad (6.42)$$

which implies the bound (6.16). To prove that $\cos(x_{\text{nom}})$ has zero time average, it suffices to substitute $t = T$ in (6.42). ∎

Proof of Theorem 31: Consider the Lyapunov candidate $V(x_{\text{intra}}, t) = x_{\text{intra}}^{\mathsf{T}}\Gamma(t)x_{\text{intra}}$, and notice that, using (6.33),

$$\dot{V}(x_{\text{intra}}, t) = x_{\text{intra}}^{\mathsf{T}}[\,J_{\text{intra}}^{\mathsf{T}}\Gamma + \Gamma J_{\text{intra}} + \dot{\Gamma} + \cos(x_{\text{nom}})(J_{\text{inter}}^{\mathsf{T}}\Gamma + \Gamma J_{\text{inter}})\,]x_{\text{intra}} + O(\|x_{\text{intra}}\|^3).$$

$$(6.43)$$

Let $\dot{\Gamma} = -\cos(x_{\text{nom}})(J_{\text{inter}}^{\mathsf{T}}\Gamma + \Gamma J_{\text{inter}})$ and notice that, following [197, Exercise 3.9 and Property 4.2], its solution satisfies

$$\Gamma(t) = \exp\left[-\int_0^t \cos(x_{\text{nom}}(\tau))\,J_{\text{inter}}^{\mathsf{T}}\,d\tau\right]\Gamma(0)\cdot\exp\left[-\int_0^t \cos(x_{\text{nom}}(\tau))J_{\text{inter}}\,d\tau\right].$$

This implies that $V(x_{\text{intra}}, t)$ is a Lyapunov function for (6.33) because, by Lemma 29, $\int_0^t \cos(x_{\text{nom}}(\tau))\,d\tau$ is bounded. Furthermore, notice that

$$\exp\left[-\int_0^t \cos(x_{\text{nom}}(\tau))\,J_{\text{inter}}^{\mathsf{T}}\,d\tau\right] = I + \underbrace{\sum_{k=1}^{\infty} \frac{(J_{\text{inter}}^{\mathsf{T}})^k}{k!}\left(-\int_0^t \cos(x_{\text{nom}}(\tau))\,d\tau\right)^k}_{\Delta}.$$

Thus, (6.43) can equivalently be written as $\dot{V} = x_{\text{intra}}^{\mathsf{T}}[J_{\text{intra}}^{\mathsf{T}}\Gamma(0) + \Gamma(0)J_{\text{intra}} + M\,]x_{\text{intra}} + O(\|x_{\text{intra}}\|^3)$, where $M = J_{\text{intra}}^{\mathsf{T}}\Delta\Gamma(0)\Delta^{\mathsf{T}} + \Delta\Gamma(0)\Delta^{\mathsf{T}}J_{\text{intra}} + J_{\text{intra}}^{\mathsf{T}}(\Delta\Gamma(0) + \Gamma(0)\Delta) + (\Delta\Gamma(0) + \Gamma(0)\Delta)J_{\text{intra}}$. using

132

the triangle inequality and Lemma 29, we obtain

$$
\begin{aligned}
\|\Delta\| &= \left\| \sum_{k=1}^{\infty} \frac{(J_{\text{inter}}^{\mathsf{T}})^k}{k!} \left(-\int_0^t \cos(x_{\text{nom}}(\tau)) \, \mathrm{d}\tau \right)^k \right\| \\
&\leq \sum_{k=1}^{\infty} \frac{\|J_{\text{inter}}\|^k}{k!} \left| \int_0^t \cos(x_{\text{nom}}(\tau)) \, \mathrm{d}\tau \right|^k \\
&= e^{\left| \int_0^t \cos(x_{\text{nom}}(\tau)) \, \mathrm{d}\tau \right| \|J_{\text{inter}}\|} - 1 \leq e^{\frac{1}{\bar{a}} \log\left(\frac{\bar{\omega}+\bar{a}}{\bar{\omega}-\bar{a}} \right) \|J_{\text{inter}}\|} - 1.
\end{aligned}
$$

Because J_{intra} is stable, there always exists $\Gamma(0) \succ 0$ such that $J_{\text{intra}}^{\mathsf{T}}\Gamma(0) + \Gamma(0)J_{\text{intra}} = -Q$ for any $Q \succ 0$. Thus,

$$
\dot{V} \leq (-\lambda_{\min}(Q) + \|M\|)\|x_{\text{intra}}\|^2 + O(\|x_{\text{intra}}\|^3). \tag{6.44}
$$

By a simple Lyapunov argument, the cluster synchronization manifold $\mathcal{S}_{\mathcal{P}}$ is locally exponentially stable if $\|M\| < \lambda_{\min}(Q)$. In addition, $\|M\|$ can be upper bounded as $\|M\| \leq 2\|J_{\text{intra}}\|\|\Gamma(0)\|\|\Delta\|(\|\Delta\| + 2) \leq 2\lambda_{\max}(\Gamma(0))\|J_{\text{intra}}\| \left(e^{\frac{2}{\bar{a}} \log\left(\frac{\bar{\omega}+\bar{a}}{\bar{\omega}-\bar{a}} \right) \|J_{\text{inter}}\|} - 1 \right)$. Thus, a sufficient condition for local exponential stability is $2\lambda_{\max}(\Gamma(0))\|J_{\text{intra}}\| \left(e^{\frac{2}{\bar{a}} \log\left(\frac{\bar{\omega}+\bar{a}}{\bar{\omega}-\bar{a}} \right) \|J_{\text{inter}}\|} - 1 \right) < \lambda_{\min}(Q)$, and because the ratio $\lambda_{\min}(Q)/\lambda_{\max}(\Gamma(0))$ is maximized for $Q = I$ [194, Exercise 9.1], we have

$$
2\lambda_{\max}(\Gamma(0))\|J_{\text{intra}}\| \left(e^{\frac{2}{\bar{a}} \log\left(\frac{\bar{\omega}+\bar{a}}{\bar{\omega}-\bar{a}} \right) \|J_{\text{inter}}\|} - 1 \right) < 1,
$$

from which condition (6.18) follows. ∎

Proof of Theorem 32: From (6.43) and for $\beta \in \mathbb{R}_{>0}$ we have $\dot{V}(x_{\text{intra}}, t) = x_{\text{intra}}^{\mathsf{T}}[J_{\text{intra}}^{\mathsf{T}}\Gamma + \Gamma J_{\text{intra}}]x_{\text{intra}} + O(\|x_{\text{intra}}\|^3) = -\beta x_{\text{intra}}^{\mathsf{T}}\Gamma x_{\text{intra}} + O(\|x_{\text{intra}}\|^3)$, which is negative in a small neighborhood of the origin. ∎

Chapter 7

Relay Interactions Enable Remote Synchronization of Phase Oscillators

In this chapter, we study remote synchronization, which describes a fascinating phenomenon where oscillators that are not directly connected via physical links evolve synchronously. This phenomenon is thought to be critical for distributed information processing in the mammalian brain, where long-range synchronization is empirically observed between neural populations belonging to spatially distant brain regions. Inspired by the growing belief that remote synchronization may be prompted by intermediate mediating brain regions, such as the *thalamus*, in this chapter we derive a novel mechanism to achieve remote synchronization. This mechanism prescribes remotely synchronized oscillators to be stably connected to a cohesive relay in the network – a group of tightly connected oscillators mediating the distant ones. Remote synchronization unfolds whenever the stability of the subnetwork formed by relays and remotely synchronized oscillators is not affected by the rest of the oscillators. In accordance with our results, we find that remotely-synchronized cortico-

thalamo-cortical circuits in the brain posses strong interconnection profiles. Finally, we demonstrate that the absence of cohesive relays prevents stable remote synchronization in a large class of cases, further validating our results. We refer the reader to the published work [201] for a comprehensive discussion.

7.1 Introduction

Synchronization is a universal phenomenon intimately related to the functioning of many natural and engineered systems [145]. In the brain, synchronization phenomena are thought to constitute the neural basis of cognition, memory, and large-scale information processing [202–204]. Empirical evidence demonstrates that brain regions that are not physically interconnected are capable of synchronizing, giving rise to what is known as *remote* synchronization [205–207]. In this context, the thalamus is believed to be an enabler of remote synchronization by functioning as a central hub that relays information to distant cortical regions [205]. However, the mechanism underlying this compelling phenomenon has not been fully characterized yet, and studies on remote synchronization remain few and sparse.

To investigate this phenomenon, we model neural activity as the network-wide product of interacting oscillators, where each oscillator represents a brain region [57]. Preliminary work has probed remote synchronization in phase-amplitude oscillators, producing seminal results [208–210]. Yet, there is compelling evidence that most of the information in brain-wide interactions can be explained by the phases of brain signals, not their amplitude [202, 211]. Hence, phase oscillators lend themselves as an ideal candidate for the modeling and analysis of remotely synchronizing brain regions.

In this chapter, we utilize heterogeneous Kuramoto oscillators to investigate the role that network topology and parameters play in the emergence of remote synchronization. Specifically, we derive conditions to ensure stable remote synchronization that prescribe the existence of a strongly

connected set of oscillators acting as intermediate relay between remotely-synchronized nodes. Moreover, by analyzing human brain structural and functional data, we find evidence that the brain may enact a similar mechanism.

Related work. Kuramoto-like models, known for their rich dynamics and fascinating behaviors [28], have been widely used for the study of neural synchronization phenomena [212–215]. Besides emerging in brain recordings, remote synchronization finds applications in climate research [216] and in secure communication technologies [217]. In the latter, concurrent remote synchronization of distant network nodes and asynchronous behavior of the intermediate ones allows for the secure distribution of critical information. Despite its importance, the characterization of remote synchronization in phase oscillators has remained elusive.

Early attempts at the characterization of remote synchronization in phase oscillators have been made by employing phase shifts [218] and network symmetries (mathematically described by graph automorphisms) [219,220]. Here, we present a different mechanism for remote synchronization, and show that network symmetries, although beneficial, are not necessary for the emergence of remote synchronization.

Chapter contribution. The main contribution of this chapter is the derivation of a mechanism that guarantees the emergence of stable remote synchronization in networks of heterogeneous Kuramoto oscillators. We demonstrate that stability of remote synchronization is guaranteed whenever there exists a network partition where oscillators within the same group evolve cohesively, and at least one group consists of remotely-synchronized oscillators with strongly connected neighbors – a *cohesive relay*. We confirm the existence of a cohesive relay that enables remote synchronization in human brain data using a publicly available dataset. The mechanism proposed in this chapter builds upon and extends the results from the previous chapter (see also [142]), and complements previous work on relay synchronization [221].

Furthermore, we reveal that the absence of cohesive relays hinders the stability of the remote synchronization manifold in the class of networks comprising two groups of synchronized

136

oscillators. This important result suggests that our condition may be (almost) necessary. Finally, since Kuramoto oscillators are hard to analyze with the Master Stability Function formalism, our findings shed light to the challenging analysis of exotic synchronization phenomena in this class of oscillators.

Mathematical notation. $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, and $\mathbb{S}$ denote the real numbers, the nonnegative real numbers, the positive real numbers, and the unit circle, respectively. The set $\mathbb{T}^n = \mathbb{S} \times \cdots \times \mathbb{S}$ is the n-dimensional torus. We use $\mathbb{1}$ and e_i to represent the vector of all ones and the i-th canonical vector, respectively. The operation $A^\dagger$ denotes the Moore-Penrose pseudoinverse of the matrix A. An M-matrix is a real nonsingular matrix $A = [a_{ij}]$ such that $a_{ij} \leq 0$ for all $i \neq j$ and all leading principal minors are positive. Finally, $A \succ 0$ indicates that A is positive definite, and $D = \text{diag}(c_1, \ldots, c_n)$ represents a diagonal matrix with (i, i)-th entry c_i, $i = 1, \ldots, n$.

7.2 Problem setup and preliminary notions

In this chapter, we characterize the stability properties of remotely-synchronized trajectories arising in networks of oscillators with Kuramoto dynamics. To this aim, let $\mathcal{G} = \{\mathcal{O}, \mathcal{E}\}$ be the connected and weighted graph representing the network of oscillators, with $\mathcal{O} = \{1, \ldots, n\}$ being the oscillator set, and $\mathcal{E}$ being the edge set. Let $A = [a_{ij}]$ be the sparse adjacency matrix of the network, where $a_{ij} > 0$ whenever $(i, j) \in \mathcal{E}$, and $a_{ij} = 0$ otherwise. The dynamics of the i-the oscillator reads

$$\dot{\theta}_i = \omega_i + \sum_{j \neq i} a_{ij} \sin(\theta_j - \theta_i), \tag{7.1}$$

where $\omega_i \in \mathbb{R}_{>0}$ and $\theta_i \in \mathbb{S}$ denote the natural frequency and the phase of the i-th oscillator, respectively, and a_{ij} represents the coupling strength of the undirected edge between oscillators i and j. We assume that there are no self-loops ($a_{ii} = 0$).

In what follows, we distinguish between phase synchronization and frequency synchronization of the oscillators.

Definition 42 ((Remote) phase synchronization) *We say that oscillators* $i, j \in \mathcal{O}$ *are* phase

synchronized *if $\theta_i(t) = \theta_j(t)$ for all $t \geq 0$. Additionally, two phase-synchronized oscillators $i, j \in \mathcal{O}$* are remotely-synchronized *if $a_{ij} = 0$.* □

Definition 43 (Frequency synchronization (manifold)) *We say that oscillators $i, j \in \mathcal{O}$ are* frequency-synchronized *if $\dot{\theta}_i(t) = \dot{\theta}_j(t)$ for all $t \geq 0$. Additionally, the frequency synchronization manifold for $\mathcal{G}$ is $\mathcal{M}_{\mathcal{G}} = \{\theta \in \mathbb{T}^n : \omega_i + \sum_k a_{ik}(\theta_k - \theta_i) = \omega_j + \sum_k a_{jk}(\theta_k - \theta_j), \forall i, j \in \mathcal{O}\}$.* □

Conditions for the oscillators in $\mathcal{G}$ to have an asymptotically stable frequency synchronization manifold $\mathcal{M}$ can be found in [28, 222], and demand that the coupling strengths dominate the heterogeneity of the natural frequencies.

Remote synchronization can be studied as a special case of *cluster synchronization*, where the oscillators can be partitioned into clusters (possibly, singletons) so that the oscillators in each cluster evolve identically. To formalize the treatment, consider a network partition $\mathcal{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_p\}$, with clusters satisfying $\mathcal{C}_k \cap \mathcal{C}_\ell = \emptyset$ for all $k, \ell \in \{1, \ldots, p\}$, $k \neq \ell$, and $\bigcup_{k=1}^p \mathcal{C}_k = \mathcal{O}$. The cluster synchronization manifold is

$$\mathcal{S}_{\mathcal{C}} = \{\theta \in \mathbb{T}^n : \theta_i = \theta_j \text{ for all } i, j \in \mathcal{C}_\ell, \ \ell = 1, \ldots, p\}.$$

To focus on remote synchronization, we assume that the first $m \geq 1$ clusters contain remotely-synchronized oscillators:

(A1) there exists $1 \leq m \leq p$ such that $a_{ij} = 0$ for all $i, j \in \mathcal{C}_\ell$ and $\ell \in \{1, \ldots, m\}$.

Fig. 7.1 illustrates a network partitioned into clusters.

To be stable, the cluster synchronization manifold $\mathcal{S}_{\mathcal{C}}$ must be invariant. Sufficient conditions on the network weights and oscillators' natural frequencies for the invariance of $\mathcal{S}_{\mathcal{C}}$ have been derived elsewhere, and read as follows [142]:

(C1) the natural frequencies satisfy $\omega_i = \omega_j$ for every $i, j \in \mathcal{C}_k$ and $k \in \{1, \ldots, m\}$;

(C2) The network weights satisfy $\sum_{k \in \mathcal{C}_\ell} a_{ik} - a_{jk} = 0$ for every $i, j \in \mathcal{C}_z$ and $z, \ell \in \{1, \ldots, m\}$, with $z \neq \ell$.
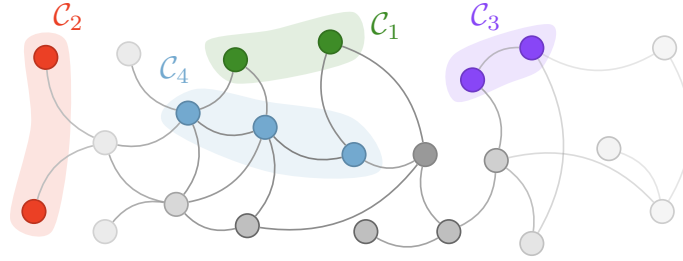
Figure 7.1: A network with (remotely-)synchronized clusters of oscillators. The clusters $\mathcal{C}_1$ and $\mathcal{C}_2$ comprise remotely-synchronized oscillators. Clusters $\mathcal{C}_3$ and $\mathcal{C}_4$ contain connected phase-synchronized oscillators. All the singletons $\mathcal{C}_5, \ldots, \mathcal{C}_p$ are depicted in shades of gray.

In the following, we assume that (C1) and (C2) are satisfied for the partition $\mathcal{C}$ being considered.

We conclude this section by stressing out that existing conditions for the stability of the cluster synchronization manifold $\mathcal{S}_{\mathcal{C}}$ require each cluster in $\mathcal{C}$ to be connected internally [142, 143], and do not cover the case of disconnected clusters. In the next section, we extend the results in [142] to account for the case of remote synchronization.

7.3 Stability of remote synchronization through relay interactions

In this section, we present a condition to ensure that remote synchronization emerges in a network of heterogeneous Kuramoto oscillators. We focus on local stability because $\mathcal{S}_{\mathcal{C}}$ is, in general, not globally asymptotically stable. As it is not clear where on the manifold the system stabilizes, we are interested in any trajectory that converges to $\mathcal{S}_{\mathcal{C}}$. Our condition reveals that the existence of a strongly connected relay promotes the stability of remotely-synchronized trajectories.

7.3.1 A sufficient condition for the stability of remote synchronization via perturbation theory and relays

In order to derive our stability condition, we introduce the incremental dynamics of (9.1). Let $x_{ij} = \theta_j - \theta_i$, so that $\dot{x}_{ij} = \dot{\theta}_j - \dot{\theta}_i$. Stacking all the differences x_{ij} with $i < j$ yields a vector of phase differences $x = B^\mathsf{T}\theta$, with B being the oriented incidence matrix of $\mathcal{G}$.[1]

It is worth noting that $n-1$ phase differences $x_{\min}$ encode all phase trajectories, and that there always exists a full column-rank submatrix $B_{\min} \in \mathbb{R}^{n \times (n-1)}$ of B such that

$$x = B^\mathsf{T}(B_{\min}^\mathsf{T})^\dagger x_{\min},$$

where $x_{\min}$ is a set of $n-1$ phase differences that can be used to quantify synchronization among all oscillators. To see this, let $x_{\min} = B_{\min}^\mathsf{T}\theta$, with $B_{\min}$ being any full column-rank submatrix of B (e.g., the incidence matrix of a spanning tree of $\mathcal{G}$ [42]). Because $\ker(B^\mathsf{T}) = \ker(B_{\min}^\mathsf{T}) = \mathbb{1}$ by definition, it holds that $x = B^\mathsf{T}\theta = B^\mathsf{T}(\theta + c\mathbb{1}) = B^\mathsf{T}(B_{\min}^\mathsf{T})^\dagger x_{\min}$.

Let $\dot{x}_{\min}^{(k)} = g_k(x_{\min}^{(k)})$ denote the incremental dynamics of (9.1) restricted to a subnetwork $\mathcal{G}_k = \{\mathcal{O}_k, \mathcal{E}_k\}$, with $\mathcal{O}_k \subseteq \mathcal{O}$ and $\mathcal{E}_k = \{(i,j) \in \mathcal{E} \ : \ i, j \in \mathcal{O}_k\} \subseteq \mathcal{E}$, and let

$$J_k(\bar{x}_{\min}^{(k)}) = \frac{\partial g_k}{\partial x_{\min}^{(k)}}(\bar{x}_{\min}^{(k)})$$

$$= -B_{\min,k}^\mathsf{T} B_k \, \mathrm{diag}(\{a_{ij}\cos(\bar{x}_{ij})\}_{(i,j)\in\mathcal{E}_k}) B_k^\mathsf{T}(B_{\min,k}^\mathsf{T})^\dagger \tag{7.2}$$

be the Jacobian matrix computed at $\bar{x}_{\min}^{(k)}$ (see Lemma 24), where each $\bar{x}_{ij}$ can be expressed as a function of $\bar{x}_{\min}^{(k)}$. The following theorem extends Theorem 25 to the case of disconnected clusters, thus providing a condition for the stability of the remote synchronization manifold.

Theorem 44 (Stable remote synchronization through relay interactions and weak outer couplings) *Let $\mathcal{S}_\mathcal{C}$ be the cluster synchronization manifold associated with a partition $\mathcal{C} = \{\mathcal{C}_1, \ldots \mathcal{C}_p\}$ of the network $\mathcal{G}$, with $\{\mathcal{C}_1, \ldots, \mathcal{C}_m\}$ comprising disconnected oscillators. The cluster synchronization*

---

[1]The $n \times |\mathcal{E}|$ oriented incidence matrix is defined entry-wise as $B_{k\ell} = -1$ if oscillator k is the source of the interconnection ℓ, $B_{k\ell} = 1$ if oscillator k is the sink of the interconnection ℓ, and $B_{k\ell} = 0$ otherwise.

manifold $\mathcal{S_C}$ is locally asymptotically stable if there exists a partition $\mathcal{F} = \{\mathcal{F}_1, \ldots, \mathcal{F}_r\}$ satisfying

the following conditions:

(i) $\mathcal{C}_\ell \subseteq \mathcal{F}_k$, for all $\ell \in \{1, \ldots, p\}$ and some $k \in \{1, \ldots, r\}$;

(ii) $\mathcal{G}_k = (\mathcal{F}_k, \mathcal{E}_k)$ is connected, for all $k \in \{1, \ldots, r\}$,

(iii) There exists a locally asymptotically stable $\mathcal{M}_{\mathcal{G}_k}$ (see Definition 43) for the oscillators in the

isolated subnetwork $\mathcal{G}_k = (\mathcal{F}_k, \mathcal{E}_k)$, for all $k \in \{1, \ldots, r\}$;

(iv) the matrix $S \in \mathbb{R}^{r \times r}$, defined as

$$S = [s_{k\ell}] = \begin{cases} \lambda_{max}^{-1}(P_k) - c^{(kk)} & \textit{if } k = \ell, \\ -c^{(k\ell)} & \textit{if } k \neq \ell, \end{cases} \tag{7.3}$$

is an M-matrix, where $P_k \succ 0$ is such that $J_k(0)P_k + P_k J_k(0)^\mathsf{T} = -I$, with J_k as in (7.2),

and, for any $i \in \mathcal{F}_k$,

$$c^{(k\ell)} = 2 \max_r |\mathcal{F}_r| \cdot \begin{cases} \displaystyle\sum_{j \in \mathcal{F}_\ell} a_{ij}, & \textit{if } \ell \neq k \\ \displaystyle\sum_{\ell \neq k} \sum_{j \in \mathcal{F}_\ell} a_{ij}, & \textit{otherwise.} \end{cases} \tag{7.4}$$

Proof. We will use perturbation theory of dynamical systems [194, ch. 9] to prove the sta-

bility of $\mathcal{S_C}$. Here, the nominal systems are the isolated sets $\mathcal{F}_1, \ldots, \mathcal{F}_r$, whose stability is perturbed

by the phase trajectories of the oscillators belonging to interconnected sets. We first show that,

because the oscillators in $\mathcal{G}_1, \ldots, \mathcal{G}_r$ have stable frequency synchronization manifolds, there exist r

quadratic Lyapunov functions for the linearized incremental dynamics of the isolated sets $\mathcal{F}_1, \ldots, \mathcal{F}_r$.

Notice that the frequency-synchronized trajectories of the oscillators in $\mathcal{F}_1, \ldots, \mathcal{F}_r$ uniquely identify

equilibria $\bar{x}_{\min}^{(k)}$, $k \in \{1, \ldots, r\}$, of their respective incremental dynamics on $\mathcal{G}_1, \ldots, \mathcal{G}_r$. By applying a

change of coordinates $y^{(k)} = x_{\min}^{(k)} - \bar{x}_{\min}^{(k)}$ such that the linearized incremental dynamics are centered

at the origin, we can define r Lyapunov functions that read as $V_k(y^{(k)}) = y^{(k)\,\mathsf{T}} P_k y^{(k)}$, with $P_k \succ 0$

such that $J_k(0)P_k + P_k J_k(0)^\mathsf{T} = -I$, and satisfy $V_k(0) = 0$, for all $k \in \{1, \ldots, r\}$. Next, we define

$\tilde{y}_{\min} = [y^{(1)}, \ldots, y^{(r)}]^{\mathsf{T}}$ as the minimum incremental variables for the entire partition $\mathcal{C}$ of $\mathcal{G}$, so that $\tilde{y}_{\min} = 0$ implies cluster synchronization. Let the Lyapunov candidate for the incremental dynamics $\dot{\tilde{y}}_{\min}$ be

$$V(\tilde{y}_{\min}) = \sum_{k=1}^{r} d_k V_k(\tilde{y}^{(k)}), \quad d_k > 0. \tag{7.5}$$

By the invariance condition (C2) and Lemma 24, we can apply perturbation theory of dynamical systems [194, ch. 9.5] to obtain that the derivative of (7.5) satisfies $\dot{V}(\tilde{y}_{\min}) \leq (DS + S^{\mathsf{T}}D)\|\tilde{y}_{\min}\|$, where $D = \mathrm{diag}(d_1, \ldots, d_r)$, and S is as in (7.3). Finally, [194, Lemma 9.7 and Theorem 9.2] define the constants $c^{(k\ell)}$ as in (7.4), the matrix S as in (7.3), and conclude on the origin of $\dot{\tilde{y}}_{\min}$ being locally stable if S is an M-matrix, thus proving the stability of $\mathcal{S}_{\mathcal{C}}$. ■

Theorem 44 introduces an additional partition of the oscillators besides $\mathcal{C}$, where the sets $\mathcal{F}_1, \ldots, \mathcal{F}_r$ contain one or more clusters each, and whose oscillators are required to be coupled strongly enough to achieve frequency synchronization when isolated from the other sets. That is, while the partition $\mathcal{C}$ encodes which oscillators are phase- and which are remotely-synchronized, $\mathcal{F}$ encodes the frequency synchronization behavior of the oscillators.

Remark 45 (Cohesive relays support stable remote synchronization) *Conditions (i) and (ii) in Theorem 44 require that any set $\mathcal{F}_k$ containing a cluster of remotely-synchronized oscillators $\mathcal{C}_\ell$, $\ell \in \{1, \ldots, m\}$ must also contain one or more connected clusters from $\{\mathcal{C}_{m+1}, \ldots, \mathcal{C}_p\}$, which in turn act as a relay for the remote oscillators in $\mathcal{F}_k$. Such relay is also* cohesive, *as all its oscillators behave cohesively due to its internal weights being large enough to guarantee frequency synchronization and the M-matrix condition. Taken together, these requirements reveal that stable remote synchronization is ensured by the existence of cohesive relays, which provide sufficient "inertia" to preserve remote synchronization despite the perturbations from other oscillator sets.* □

Fig. 7.2 illustrates remotely-synchronized oscillators supported by cohesive relays in a partition $\mathcal{F}$ for the network in Fig. 7.1. Finally, Fig. 7.3 presents an example for Theorem 44.
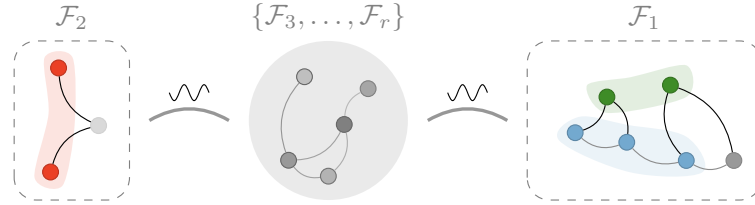
Figure 7.2: Remote synchronization through stable relay interconnections. This figure illustrates a partition $\mathcal{F} = \{\mathcal{F}_1, \ldots, \mathcal{F}_r\}$ of the network in Fig. 7.1 satisfying the conditions in Theorem 44. Requiring that the matrix in (7.3) is an M matrix guarantees that the stable frequency-synchronized trajectories in $\mathcal{F}_1$ and $\mathcal{F}_2$ are not perturbed by the rest of the oscillators.

(a)   (b)
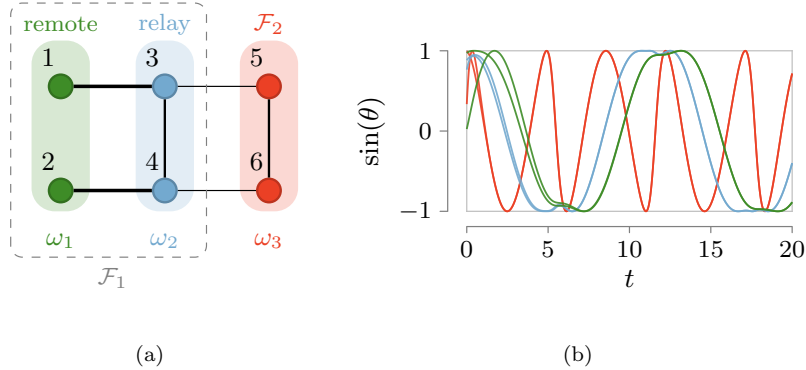
Figure 7.3: Fig. 7.3(a) illustrates a network partitioned into $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3\}$ that displays stable remote synchronization by satisfying Theorem 44 with $\mathcal{F} = \{\mathcal{F}_1, \mathcal{F}_2\}$, where $\mathcal{F}_1 = \{\mathcal{C}_1, \mathcal{C}_2\}$. The cluster $\mathcal{C}$ (in blue) being a cohesive relay for the remote cluster $\mathcal{C}_1$ (in green). In this example $\mathcal{F}_2 = \mathcal{C}_3$ (in red). We fix the network weights as $a_{13} = a_{24} = a_{34} = a_{56} = 10$ and $a_{35} = a_{46} = 1$, and the natural frequencies as $\omega_1 = 0$, $\omega_2 = 0.5$, $\omega_3 = 2.1$. It can be shown that the dynamics $\dot{x}_{\min}^{(1)} = [\dot{x}_{13}\ \dot{x}_{24}\ \dot{x}_{34}]^{\mathsf{T}}$ for $\mathcal{F}_1$ (in grey, dashed) has a stable equilibrium $\bar{x}_{\min}^{(1)} = [0\ 0.025\ 0]^{\mathsf{T}}$. The matrix $S = [[7.7139, -4]; [-4, 36]]$ in (7.3) is an M-matrix. Fig. 7.3(b) illustrates the phases evolution starting from random initial conditions close to $\mathcal{S}_{\mathcal{C}}$. The phases, which are color coded according to Fig. 7.3(a), converge to $\mathcal{S}_{\mathcal{C}}$, and the oscillators in $\mathcal{C}_1$ are remotely synchronized

.

7.3.2 Remotely-synchronized brain regions possess strong relay interconnections

While Theorem 44 does not provide a method to choose the partition $\mathcal{F}$ and, thus, which oscillators constitute relays, empirical studies have identified candidate relay regions in the brain to be the backbone of a set $\mathcal{F}_k \subset \mathcal{F}$. We interrogate the publicly available `NKI Rockland` dataset [223] to verify whether there exists a relay whose connections with brain regions that display remote synchronization are much stronger than the average connection strength across all brain regions. More in detail, we analyze cortico-thalamo-cortical circuits [205], where disconnected cortical regions synchronize via interconnections with the thalamus.

In the considered dataset, the anatomical organization of the brain is encoded by adjacency matrices $W_i \in \mathbb{R}_{\geq 0}^{188 \times 188}$. We analyze the first $N = 20$ subjects in the dataset, and find that each subject possesses pairs of disconnected cortical regions whose activity is highly synchronous (Pearson correlation coefficients ≥ 0.9). For each subject, we select the pair displaying the largest mean coupling strengths with the thalamus. We find that, across all 20 subjects, the mean cortico-thalamic interconnection of remotely synchronized regions is 4.079 ± 0.303 SEM (Standard Error of the Mean) times the mean network weight. Fig. 7.4 illustrates the distribution of cortico-thalamic interconnection weights divided by the average interconnection weight of each subject. Remarkably, even the smallest among these values is larger than one. Because a partition $\mathcal{F}$ cannot be uniquely identified from the available data, we cannot check whether Theorem 44 is satisfied. Nevertheless, the above finding is in line with the theoretical requirement of strong relay interconnections, and suggests that remote synchronization in the brain may be supported by strong interactions with a cohesive relay.

$$a_{\mathrm{CT}}/\mathrm{mean}(a_{ij})$$
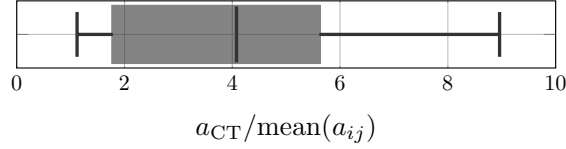
Figure 7.4: Distribution of cortico-thalamic interconnections of remotely synchronized cortical regions in anatomical brain networks of $N = 20$ subjects. For each subject, we plot the the cortico-thalamic interconnection weight of remotely-synchronized cortical regions a_{CT} divided by the subject-specific average interconnection weight $\mathrm{mean}(a_{ij})$.

7.4 Unstable remote synchronization in the absence of cohesive relays

In this section, we provide evidence that the absence of cohesive relays hinders the emergence of stable remote synchronization in a large class of networks. To do so, we first assess whether frequency-synchronized sets that include remotely-synchronized oscillators are necessary to enable stable remote synchronization of a single cluster of disconnected oscillators. We focus on the case of networks partitioned into two clusters ($\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2\}$), where $\dot{\theta}_i \neq \dot{\theta}_j$ for all $i \in \mathcal{C}_1$ and $j \in \mathcal{C}_2$. We set to zero all intra-cluster couplings in $\mathcal{C}_1$, so that $\mathcal{C}_2$ is a (non-cohesive) relay for the remote oscillators in $\mathcal{C}_1$. In this configuration, there does not exist a partition $\mathcal{F}$ satisfying Theorem 44.

The next theorem shows that this specific configuration yields at best *marginal stability* of $\mathcal{S}_{\mathcal{C}}$ – as phase trajectories that start in the vicinity of $\mathcal{S}_{\mathcal{C}}$ never converge to it as $t \to \infty$. To present our result, we define $\bar{\omega} = \omega_2 - \omega_1$, with ω_1, ω_2 being the natural frequencies of the oscillators in $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively. Finally, let $\bar{a} = \sum_{k \in \mathcal{C}_2} a_{ik} + \sum_{k \in \mathcal{C}_1} a_{jk}$, for any $i \in \mathcal{C}_1$ and $j \in \mathcal{C}_2$.

Theorem 46 (Instability of $\mathcal{S}_{\mathcal{C}}$ if common neighbors are not frequency-synchronized)
Consider the network $\mathcal{G}$ partitioned as $\mathcal{P} = \{\mathcal{C}_1, \mathcal{C}_2\}$, with $a_{ij} = 0$ for all $i, j \in \mathcal{C}_1$ and $a_{ij} = a_{ik} > 0$ for any $i \in \mathcal{C}_1$ and all $j, k \in \mathcal{C}_2$. If $\bar{\omega} > \bar{a}$, then the cluster synchronization manifold $\mathcal{S}_{\mathcal{C}}$ is not asymptotically stable. Furthermore, for any initial condition $\theta(0) = \tilde{\theta}(0) + \alpha e_i$, with $\tilde{\theta}(0) \in \mathcal{S}_{\mathcal{C}}$, $\alpha \in \mathbb{T}$, and $i \in \mathcal{C}_1$, the solution to (9.1) is periodic and does not belong to $\mathcal{S}_{\mathcal{C}}$.

145
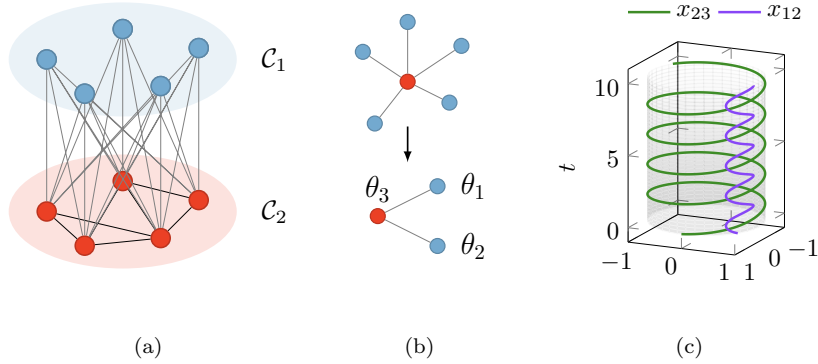
Figure 7.5: Fig. 7.5(a) illustrates a 2-cluster network where the remote oscillators in $\mathcal{C}_1$ are connected to all the oscillators in $\mathcal{C}_2$, thus satisfying the condition in Theorem 46. Fig. 7.5(b) illustrates that, due to the synchronization of all the oscillators in $\mathcal{C}_2$ (in red) whenever the oscillators in $\mathcal{C}_1$ share common neighbors, the network in Fig. 7.5(a) can be reduced to a star network in order to analyze its stability properties. When studying a specific perturbation of $\theta(0) \in \mathcal{S}_{\mathcal{C}}$ so that only the oscillators in $\mathcal{C}_1$ are outside of $\mathcal{S}_{\mathcal{C}}$ at $t = 0$, the star network in the top panel can be further reduced to a 3-node star. Fig. 7.5(c) depicts the periodic trajectories (in Cartesian coordinates) $x_{12} = \theta_2 - \theta_1$ and $x_{23} = \theta_3 - \theta_2$ of the 3-node star in Fig. 7.5(b) starting from $\theta_1(0) = -1$ and $\theta_2(0) = \theta_3(0) = 0$, which satisfy Theorem 46.

Proof. We prove that there exist an infinite number of time instants $t_1, t_2, \ldots$ such that $\theta(0) = \theta(t_1) = \theta(t_2) = \ldots$. Owing to Lemma 29, the condition $\bar{\omega} > \bar{a}$ implies that $\dot{\theta}_i \neq \dot{\theta}_j$ for all $i \in \mathcal{C}_1$ and $j \in \mathcal{C}_2$. Notice that any remote oscillator can equivalently be seen as a singleton cluster, and that the phases of such clusters cancel out in the dynamics of $\dot{x}_{ij}^{(2)}$ for all $i, j \in \mathcal{C}_2$. Hence, due to (C1) and the fact that $\mathcal{C}_2$ is internally connected, it holds that $|x_{ij}| \to 0$ for all $i, j \in \mathcal{C}_2$ [142]. In this case, we observe that any network akin to Fig. 7.5(a) can be conveniently analyzed as a star network, where $\mathcal{C}_2$ is considered as the oscillator at the center of the star because its oscillators synchronize (see Fig. 7.5(b)).

Let us reorder the oscillators so that the last one is the star center, and consider initial conditions $\theta(0) = c\mathbb{1} + [\alpha \; 0 \; \ldots \; 0]^\mathsf{T}$; that is, all oscillators start at the same value $c \in \mathbb{T}$ and $x_{1k} = \alpha$, for all $k = 2, 3, \ldots, |\mathcal{C}_1|$. This choice of $\theta(0)$ allows us to further reduce the star network to a 3-oscillator star, where synchronized oscillators $2, \ldots, |\mathcal{C}_1|$ become a single oscillator, as illustrated in Fig. 7.5(b). To see this, recall that the synchronized trajectories of $2, \ldots, |\mathcal{C}_1|$ are invariant due to

146

(C1), (C2). Finally, for simplicity, let us set $a_{ij} = 1$ for all a_{ij}.[2] We are left with studying a 3-oscillator star obeying equations $\dot{\theta}_1 = \sin(\theta_3 - \theta_1)$, $\dot{\theta}_2 = \sin(\theta_3 - \theta_2)$, $\dot{\theta}_3 = \omega_3 - \sin(\theta_3 - \theta_1) - \sin(\theta_3 - \theta_2)$.

We are now ready to demonstrate that there exist $t_1, t_2, \ldots$ such that $\theta(0) = \theta(t_1) = \theta(t_2) = \ldots$. We first define $\gamma = \frac{1}{2}(\theta_2 - \theta_1)$ and $\varphi = \omega_3 - \frac{\theta_1 + \theta_2}{2}$. On the unit circle, $\gamma(t)$ represents the evolution of half of the difference between the two outer oscillators, while $\varphi(t)$ represents the evolution of the difference between $\theta_3(t)$ and the center of the difference between the outer oscillators. Since $\theta_3 - \theta_1 = \gamma + \varphi$ and $\theta_3 - \theta_2 = \gamma - \varphi$, the time evolution of γ and φ can be written as $\dot{\gamma} = \frac{1}{2} [\sin(\varphi - \gamma) - \sin(\varphi + \gamma)]$, $\dot{\varphi} = \gamma_3 - \frac{3}{2} [\varphi - \gamma) + \sin(\varphi + \gamma)]$. Notice that $\dot{\gamma} = 0$ for $(\gamma, \varphi) = (\frac{\alpha}{2}, k\frac{\pi}{2})$ with k odd. Next, we get rid of the dependence of γ from time by considering $\frac{\dot{\gamma}}{\dot{\varphi}} = \frac{\partial \gamma}{\partial t} \frac{\partial t}{\partial \varphi} = \frac{\partial \gamma}{\partial \varphi}$. By recalling that φ is monotonically increasing (as $\bar{\omega} > \bar{a}$ implies that $\dot{\theta}_3 \neq \dot{\theta}_1$), we can study $\gamma(\varphi)$ and $\frac{\partial \gamma}{\partial \varphi}(\varphi)$ as a function of φ. In particular, showing that $\gamma(-\frac{\pi}{2}) = \gamma(\frac{3\pi}{2})$ is equivalent to show that x_{12} does not converge to zero. To do so, recall that $\gamma(\frac{3\pi}{2}) = \gamma(-\frac{\pi}{2}) + \int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{\partial \gamma}{\partial \varphi}(\tau) \mathrm{d}\tau$.

It holds that the integral in the above equation is zero, as it can be shown that

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\partial \gamma}{\partial \varphi}(\tau) \mathrm{d}\tau = -\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{\partial \gamma}{\partial \varphi}(\tau) \mathrm{d}\tau. \tag{7.6}$$

Specifically, we prove by mathematical induction that, for any $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\frac{\partial \gamma}{\partial \varphi}(\varphi) = \frac{\partial \gamma}{\partial \varphi}(\pi - \varphi)$. This can be done by letting $\varphi_1 = \frac{\pi}{2}$, and considering for the base step of the induction $\varphi_1^+ = \varphi + \delta$ and $\varphi_1^- = \varphi - \delta$. In the limit for $\delta \to 0$, it holds that $\frac{\partial \gamma}{\partial \varphi}(\varphi_1^+) = \frac{\partial \gamma}{\partial \varphi}(\varphi_1^-)$. The inductive step can be proven analogously, thus concluding on the symmetry of $\frac{\partial \gamma}{\partial \varphi}$ in the unit circle with respect to the axis $(-\frac{\pi}{2}, \frac{\pi}{2})$, which implies the equivalence in (7.6) since we have that

$$\int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{\partial \gamma}{\partial \varphi}(\tau) \mathrm{d}\tau = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\partial \gamma}{\partial \varphi}(\tau) \mathrm{d}\tau + \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{\partial \gamma}{\partial \varphi}(\tau) \mathrm{d}\tau$$
$$= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\partial \gamma}{\partial \varphi}(\tau) \mathrm{d}\tau + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\partial \gamma}{\partial \varphi}(\pi - \beta) \mathrm{d}\beta = 0.$$

This concludes the proof that x_{12} circles back to $x_{12}(0)$ an infinite amount of times, thus implying that the remote synchronization manifold is not asymptotically stable. ∎

---

[2]The reasoning below can be extended to any weights $a_{ij} \in \mathbb{R}_{\geq 0}$.
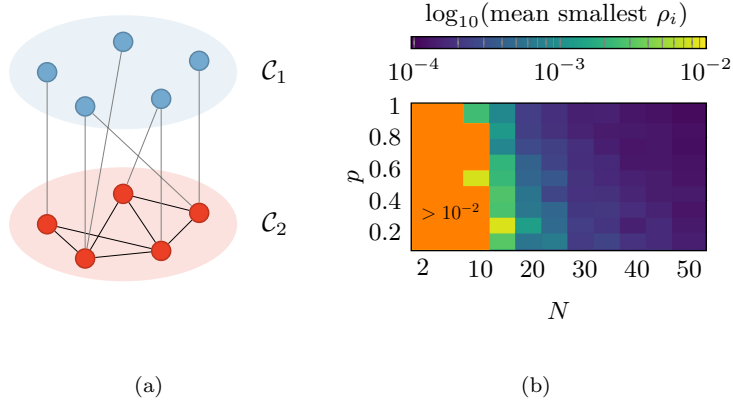
Figure 7.6: Fig. 7.6(a) illustrate a 2-cluster network with random inter-cluster coupling and no common neighbors. Fig. 7.6(b) depicts consistently unstable Floquet characteristic multipliers of a 2-cluster network for varying topologies and cluster sizes. We generated, for each cluster size $|\mathcal{C}_i| = N$ and probability p of interconnection between two nodes, 100 Erdös-Rényi graph topologies with weights drawn uniformly from $[0, 1]$. Each entry of the matrix plot represents the mean unstable characteristic multiplier ρ_i averaged over 100 different network instances satisfying: (i) clusters $\mathcal{C}_1$ and $\mathcal{C}_2$ consist of N oscillators each, (ii) the conditions for the invariance of $\mathcal{S}_{\mathrm{R}}$, (iii) $a_{ij} = 0$ for all $i, j \in \mathcal{C}_1$.

Theorem 46 shows that if every oscillator $i \in \mathcal{C}_1$ receives the same "input" $\sum_{j \in \mathcal{C}_2} a_{ij} \sin(\theta_j - \theta_i)$ from all the oscillators in $\mathcal{C}_2$, as illustrated in Fig. 7.5(a), the cluster synchronization manifold $\mathcal{S}_\mathcal{C}$ is at best marginally stable. Additionally, Theorem 46 demonstrates the existence of a family of initial conditions, which can be arbitrarily close to $\mathcal{S}_\mathcal{C}$, that yields periodic trajectories not belonging to $\mathcal{S}_\mathcal{C}$. Fig. 7.5(c) illustrates periodic trajectories as in Theorem 46.

We confirm by means of Floquet stability theory [224] that network topologies not as accurately crafted as the ones in Theorem 46 yield an unstable $\mathcal{S}_\mathcal{C}$. We remark that it is possible to employ Floquet stability theory because the phase differences between the oscillators in $\mathcal{C}_1$ and the ones in $\mathcal{C}_2$ are periodic whenever the two clusters are not frequency-synchronized (i.e., $\bar{\omega} > \bar{a}$) [142, Lemma 3.4].

We generated 10^4 2-cluster networks with varying size and connectivity profiles, where $a_{ij} = 0$ for all $i, j \in \mathcal{C}_1$, and $\mathcal{C}_2$ is a random connected topology (Erdös-Rényi graphs with edge probability $p \in [0.1, 1]$). Fig. 7.6 summarizes the stability analysis of $\mathcal{S}_\mathcal{C}$ in all these networks. While

not excluding that remote synchronization may arise in specific instances where multiple phase-unlocked clusters interact, our results suggest that the sufficient mechanism proposed in Section 7.3 may be (almost) necessary for the stability of $\mathcal{S_C}$.

7.5   Conclusion

We have studied remote synchronization in networks of heterogeneous Kuramoto oscillators. Motivated by the hypothesis that remote synchronization in the brain is promoted by intermediate brain regions that relay information between disconnected ones, we have proposed a novel mechanism to achieve stable remote synchronization. Our main result prescribes the existence of cohesive relays, which are required to be strongly coupled to remote oscillators and weakly perturbed by the other ones. We have also analyzed brain data and found that regions that remotely-synchronize are strongly connected to a common relay. Finally, we have demonstrated that the case of 2 clusters where no cohesive relay exists does not admit stable remote synchronization. The latter result suggests that our sufficient conditions may also be necessary. We leave the validation of this conjecture as a topic for future research.

Part III

Control of Synchronization in

Brain Networks

Chapter 8

Feedback Linearization of Nonlinear Network Systems

This chapter introduces the method of feedback linearization applied to network systems with the goal of achieving desired synchronization patterns. Feedback linearization allows for the local transformation of a nonlinear system to an equivalent linear one by means of a coordinate transformation and a feedback law. Feedback linearization of large-scale nonlinear network systems is typically difficult, as existing conditions become harder to check as the network size becomes larger. In this letter, we provide novel conditions to test whether a nonlinear network is feedback-linearizable. Specifically, given some dedicated control inputs injected to a set of network nodes, we derive an easy-to-check algebraic condition that can be tested on the Jacobian matrix of the network dynamics evaluated at some desired working point. Furthermore, our requirements are sufficient for (local) controllability, and thus provide a testable condition for controllability of large-scale nonlinear networks. Finally, we validate our findings by enforcing the formation of desired synchronization patterns in networks of coupled oscillators. We refer the interested reader to the publication [225] for a comprehensive discussion.

8.1 Introduction

The ability to effectively control complex nonlinear systems is still an outstanding engineering challenge. In fact, despite the ubiquitous presence of large-scale nonlinear network systems, both in the technological [151] and the natural [10] fields, a full characterization of controllability has remained elusive. This is due to the fact that, in general, checking the known sufficient conditions becomes harder (intractable, even) as the size of the system increases [226]. In this letter, we address this issue by providing algebraic conditions for nonlinear control problems that leverage the system's internal interconnection structure. Specifically, we resort to the theory of feedback linearization, which allows for the local transformation of a nonlinear system into an equivalent linear one by means of a coordinate transformation and a feedback loop [227]. This, in turn, enables the extensive array of control-theoretic tools for linear systems to be used for the control of nonlinear network systems.

While there exists a vast amount of literature on controllability of linear systems evolving on networks, the line of work studying the nonlinear counterpart is much narrower (e.g., [35, 228]). This letter complements the latter line of work and presents conditions to test whether a nonlinear network system is feedback-linearizable from a set of dedicated control inputs. Our conditions can be evaluated on the Jacobian matrix of the system computed at a desired working point, instead of complex differential geometric quantities as in classical tests, which also consist of more restrictive conditions. Finally, motivated by our interest in controlling the synchronization capabilities of neuronal networks in biological systems, we illustrate how our results can be used to achieve cluster synchronization in regular and multiplex networks of oscillatory neuronal ensembles [214, 229].

Related work. This work aims at narrowing the gap between controllability of large-scale network systems and the feedback linearization method. After the early theoretical developments, fewer works have surfaced on the topic of nonlinear controllability, [230, 231]. Recent papers address this problem for systems evolving on networks [31, 35, 228, 232–234]. For instance, [233, 234] study

accessibility of network systems. However, accessibility is a weaker notion than controllability, and thus may be of limited use in practice [235].

Feedback linearization is a classical topic in nonlinear control theory developed during the decades between the 1960 and the 1990 [236, 237]. Applications can be found in several engineering systems, including robotic mechanisms [238] and power networks [239]. Some recent work promotes the usage of data-driven methods to achieve feedback linearization whenever the model is not known exactly [240, 241]. Relevant studies on model-based feedback linearization include [242], where the authors give conditions based on differential geometry, and [243], where conditions are given for a single-input–single-output system with delays. The work that is most closely related to ours is [239], where the authors study feedback linearization of a chain network governed by a class of nonlinear dynamics.

Chapter contribution. The contribution of this chapter is two-fold. First, by exploiting a system's interconnected structure, we provide sufficient algebraic conditions to test for the feedback linearization of nonlinear systems controlled by linear vector fields. That is, given a set of dedicated control inputs, we derive conditions that can be tested on the Jacobian matrix of the nonlinear system. Our conditions can be used to test for (local) controllability of large-scale nonlinear network systems. Additionally, the use of feedback linearization enables the evaluation of the state space region in which the linearized system is defined; note that this is a computation not possible when using Jacobian linearization.

Second, we exploit our results to address the challenging problem of controlling oscillatory networks [214, 244, 245]. We show that a class of nonlinearly coupled oscillators is feedback-linearizable and, through feedback linearization, we control the formation of desired synchronization patterns.

153

8.2 Problem setup and preliminary notions

In this chapter, we study feedback linearization and controllability of nonlinear systems governed by the dynamics

$$\dot{x} = f(x) + Bu, \tag{8.1}$$

where $x = [x_1, \ldots, x_n]^\mathsf{T} \in \mathbb{R}^n$ is the the system state, f is a smooth vector field that describes the dynamics of the system, and $B \in \mathbb{R}^{n \times m}$ is the input matrix through which the control signals u are administered to the system. We let $\mathcal{K} = \{k_1, \ldots, k_m\} \subseteq \{1, \ldots, n\}$ be the control set, and let $B = \left[e_{k_1}, \ldots, e_{k_m} \right]$, with e_i denoting the i-th canonical vector. Without loss of generality, we choose $\mathcal{K} = \{1, \ldots, m\}$, thus $B = \left[e_1, \ldots, e_m \right]$. Finally, we assume that (8.1) has at least one equilibrium. That is, there exists $\bar{x}$ such that $f(\bar{x}) = 0$.

In this chapter, we make use of a graphical representation of the dynamic interdependence of the system's components. Namely, the *inference diagram* [246] consists of a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, with $\mathcal{V} = \{1, \ldots, n\}$ being the set of n nodes where each node corresponds to a state of the system, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ being the set of edges connecting the nodes as follows. For all $i, j \in \mathcal{V}$, there exists an unweighted directed edge from node j to node i if $i \neq j$ and x_j appears in x_i's differential equation. The adjacency matrix that describes the interconnection structure of the inference diagram is the sparse matrix $A = [a_{ij}]$, with $a_{ij} = 1$ if there is an edge between j and i, and $a_{ij} = 0$ otherwise. Fig. 8.1 illustrates the procedure to generate an inference diagram, which describes the underlying topology of the state interactions for the nonlinear system $\dot{x} = f(x)$ in (8.1), but can equivalently represent the interconnection structure of nonlinearly interacting agents in a network system.

Let A be the adjacency matrix of $\mathcal{G}$, let $\mathrm{path}(i, j)$ denote a path on $\mathcal{G}$ from node i to node j, and let $|\mathrm{path}(i, j)|$ be the number of edges of $\mathrm{path}(i, j)$. Notice that $\mathcal{K}$ represents the set of control nodes in $\mathcal{G}$. Define the distance between a subset of nodes $\mathcal{S} \subseteq \mathcal{V}$ and the control set $\mathcal{K}$ as $\mathrm{dist}(\mathcal{S}, \mathcal{K}) = \min\{|\mathrm{path}(i, j)| : i \in \mathcal{K}, j \in \mathcal{S}\}$. Without loss of generality, we order the nodes according to their distance from the set of control nodes. In particular, we define a positive integer

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} f_1(x_2, x_4) \\ f_2(x_1, x_2) \\ f_3(x_2, x_4) \\ f_4(x_1, x_3) \end{bmatrix} \xrightarrow{\frac{\partial \dot{x}_i}{\partial x_j} \neq 0}$$
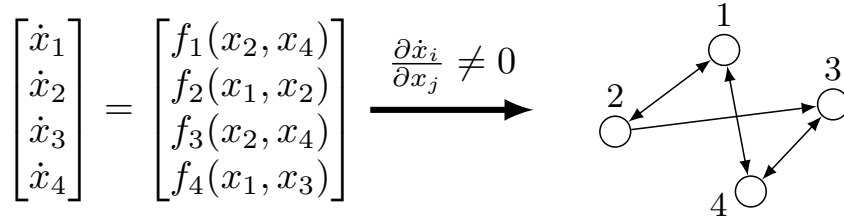
Figure 8.1: This figure illustrates the process of generating an inference diagram from a generic nonlinear system described by the vector field $f(x)$. The inference diagram depicted in this example is a graph in which there is an edge from node j to node i if $\dot{x}_i$ is a function of x_j and $i \neq j$. Notice that we do not allow self-loops in the inference diagram.

N so that $\mathcal{V} = \cup_{i=1}^{N} \mathcal{V}_i$, with $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset$ if $i \neq j$, and $\text{dist}(\mathcal{V}_i, \mathcal{K}) = i - 1$ for all $i \in \{1, \ldots, N\}$.

According to the partition $\{\mathcal{V}_1, \ldots, \mathcal{V}_N\}$, the adjacency matrix reads as

$$A = \begin{bmatrix} A_{11} & A_{12} & 0 & \cdots & & 0 \\ A_{21} & A_{22} & A_{23} & \cdots & & 0 \\ 0 & A_{32} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & A_{N-1,N-1} & A_{N-1,N} \\ 0 & \cdots & \cdots & A_{N,N-1} & A_{N,N} \end{bmatrix}, \tag{8.2}$$

where $A_{ii} \in \mathbb{R}^{|\mathcal{V}_i| \times |\mathcal{V}_i|}$, $A_{i-1,i} \in \mathbb{R}^{|\mathcal{V}_i| \times |\mathcal{V}_{i+1}|}$, and $A_{i,i-1} \in \mathbb{R}^{|\mathcal{V}_{i+1}| \times |\mathcal{V}_i|}$, with $|\mathcal{V}_i|$ denoting the cardinality of $\mathcal{V}_i$.

In this chapter, we address the following problem, whose solution is intimately tied to finding conditions for the local controllability of nonlinear systems.[1] Given a nonlinear system in the form (8.1), we investigate whether there exists a state feedback control law $u = \alpha(x) + \beta(x)v$ and a change of coordinates $z = \Phi(x)$ that transform the nonlinear system (8.1) into an equivalent controllable linear system of the form

$$\dot{z} = A_{\text{lin}}z + B_{\text{lin}}v,$$

where $A_{\text{lin}} \in \mathbb{R}^{n \times n}$ and $B_{\text{lin}} \in \mathbb{R}^{n \times m}$.

[1]A system in the form (8.1) is locally controllable at $\bar{x}$ if there exists a neighborhood $\mathcal{B}_{\bar{x}}$ of $\bar{x}$ such that for all $x_f \in \mathcal{B}_{\bar{x}}$, there exist $T > 0$ and a control input u that brings $x(0) = \bar{x}$ to $x(T) = x_f$.

To answer this question, we will make use of some notions from geometric control theory [226, 247]. Given two vector fields $f(x)$ and $g(x)$, both defined in an open subset of $\mathbb{R}^n$, we define the operation $[f, g]$ as the *Lie bracket* between $f(x)$ and $g(x)$, which yields the smooth vector field

$$[f, g](x) = \frac{\partial g(x)}{\partial x} f(x) - \frac{\partial f(x)}{\partial x} g(x).$$

To avoid confusing notation such as $[f, [f, \ldots, [f, g]]]$, we use the following recursive definition:

$$\mathrm{ad}_f^k g(x) = [f, \mathrm{ad}_f^{k-1} g](x),$$

where $\mathrm{ad}_f^0 g(x) = g(x)$. Note that the directions in which the state may be moved around an initial condition are those belonging to the set of all vector fields that can be obtained by iteratively computing the Lie brackets of the system's dynamics and the control vector fields. Further, given a real-valued function $\lambda(x)$ and a vector field $f(x)$, both defined in an open set of $\mathbb{R}^n$, we define the *derivative of λ along f*:

$$L_f \lambda(x) = \frac{\partial \lambda(x)}{\partial x} f(x).$$

If the function λ is differentiated k times along f, we write

$$L_f^k \lambda(x) = \frac{\partial \left(L_f^{k-1} \lambda(x) \right)}{\partial x} f(x).$$

For the sake of simplicity, we will omit the argument x when it is clear from the context. A (smooth) *distribution* is the assignment of the subspace spanned by the values at x of some smooth vector fields $f_1, \ldots, f_d$ that are defined in an open set $\mathcal{X} \subseteq \mathbb{R}^n$, and is denoted by $\mathscr{G} = \mathrm{span}\{f_1, \ldots, f_d\}$. In other words, a distribution assigns a vector space to each point x of the set $\mathcal{X}$. A distribution $\mathscr{G}$ is involutive if, whenever $f, g \in \mathscr{G}$, also $[f, g] \in \mathscr{G}$. Finally, the distribution $\mathscr{G}$ has constant dimension in a set $\mathcal{X} \subseteq \mathbb{R}^n$ whenever its dimension remains the same at all points in $\mathcal{X}$.

8.3 Structural conditions for feedback linearization

In this section, we show that if the inference diagram of a nonlinear system belongs to a well-defined class of networks, then there exists a change of coordinates such that the original controlled nonlinear system (8.1) can be transformed into a controllable linear system by means of a feedback law. Clearly, these conditions constitute a sufficient test for local controllability of nonlinear systems.

Before presenting our results, notice that the Jacobian $J(x) = \frac{\partial f(x)}{\partial x}$ of (8.1) reads as

$$
J(x) = \begin{bmatrix}
D_1(x) & U_1(x) & 0 & \cdots & & 0 \\
L_1(x) & D_2(x) & U_2(x) & \cdots & & 0 \\
0 & L_2(x) & \ddots & \ddots & & \vdots \\
\vdots & \ddots & \ddots & D_{N-1}(x) & U_{N-1}(x) \\
0 & \cdots & \cdots & L_{N-1}(x) & D_N(x)
\end{bmatrix},
\tag{8.3}
$$

where the blocks have the same size of the blocks in the matrix A in (8.2). We are now ready to present our main result.

Theorem 47 *(**Condition for feedback linearization**) Consider the dynamics (8.1). Let $\bar{x}$ be such that $f(\bar{x}) = 0$ and let the Jacobian $J(x) = \frac{\partial f(x)}{\partial x}$ read as in (8.3). The system is feedback-linearizable at $\bar{x}$ if $\mathrm{rank}(L_i(\bar{x})) = |\mathcal{V}_{i+1}|$ for all $i = \{1, \ldots, N-1\}$.*

Theorem 47 implicitly requires a certain network structure to hold true, as we elucidate in the following remark.

Remark 48 *(**Necessary network structure for the condition in Theorem 47**) Theorem 47 requires the subdiagonal blocks of $J(\bar{x})$ to be full row rank. Note that a necessary condition for this to hold true is that $|\mathcal{V}_{i+1}| \le |\mathcal{V}_i|$ for all $i \in \{1, \ldots, N-1\}$ (see Fig. 8.2 for an example). Furthermore, every node in partition $\mathcal{V}_{i+1}$ must have at least one incoming connection from nodes in $\mathcal{V}_i$. Clearly, this interconnection requirement prevents the network topology to produce zero rows in $L_i(\bar{x})$ for all $i \in \{1, \ldots, N-1\}$.* $\square$
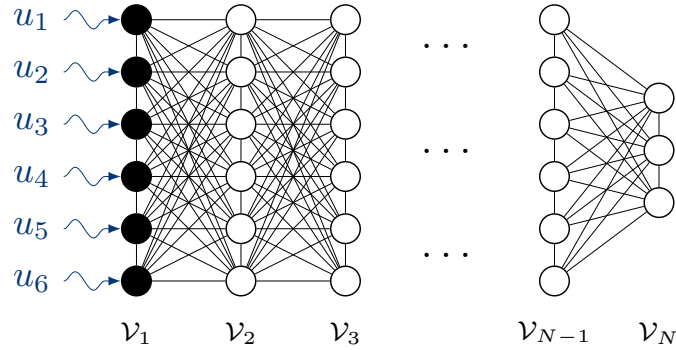
Figure 8.2: Network system with tridiagonal adjacency matrix as in (8.2) and node partition satisfying $|\mathcal{V}_1| \le |\mathcal{V}_2| \le \cdots \le |\mathcal{V}_N|$. Examples of networks with this topology are artificial neural networks, multiplex networks and, in general, all grid-like networks with layers of equal or decreasing cardinality.

Before proving Theorem 47 we introduce a lemma that relates the distribution of the Lie brackets of the dynamics (8.1) to the image of a block triangular matrix.

Lemma 49 (Block upper triangular distribution) *Consider the dynamics* (8.1). *Let* $\mathscr{G}_k = \text{span}\{\text{ad}_f^r e_\ell \ : \ 0 \le r \le k, 1 \le \ell \le m\}$, $k \in \{0, \ldots, N-1\}$. *The distribution* $\mathscr{G}_k(x)$ *can be written as*

$$\mathscr{G}_k(x) = \text{Im} \underbrace{\begin{bmatrix} E_{1,1}(x) & \otimes & \cdots & & \otimes \\ 0 & E_{2,2}(x) & \cdots & & \vdots \\ \vdots & & \ddots & \ddots & \otimes \\ 0 & \cdots & 0 & E_{k+1,k+1}(x) \\ 0 & \cdots & 0 & 0 \end{bmatrix}}_{G_k(x)},$$

where Im *denotes the image of a matrix,* $G_0(x) = B$, *so that* $E_{1,1}(x) = I$, $E_{i,i}(x) = L_{i-1}(x)E_{i-1,i-1}(x) \in \mathbb{R}^{|\mathcal{V}_i| \times |\mathcal{V}_i|}$, $\otimes$ *is any real matrix-valued function of x of suitable dimension, and* 0 *is a zero matrix of suitable dimension.*

Proof. Consider first $\mathscr{G}_0(x)$, and notice that, since $\text{ad}_f^0 e_\ell = e_\ell$, for $\ell = 1, \ldots, m$, $G_0(x) = [e_1, \ldots, e_m] = B$. Next, because of the definition of $\mathscr{G}_k(x)$, at each new step $k \in \{1, \ldots, N-1\}$, m

new vector fields of the form

$$\mathrm{ad}_f^k e_\ell = \frac{\partial \mathrm{ad}_f^{k-1} e_\ell}{\partial x} f - \frac{\partial f}{\partial x} \mathrm{ad}_f^{k-1} e_\ell, \tag{8.4}$$

with $\ell = 1, \ldots, m$, add to $\{\mathrm{ad}_f^r e_\ell \; : \; 0 \le r < k, 1 \le \ell \le m\}$ to generate the distribution $\mathscr{G}_k(x)$. The space spanned by these new vector fields corresponds to the space spanned by the last $|\mathcal{V}_{k+1}|$ columns of $G_k(x)$, which encompass the block $E_{k+1,k+1}(x)$. In what follows we show that, because of the (block) tridiagonal structure of A, only the second part of (8.4) contributes to the definition of these columns, and, therefore, of the block $E_{k+1,k+1}(x)$. To see this, consider the Lie bracket $\mathrm{ad}_f^k e_\ell(x)$ and notice that its first term reads as

$$\left[\frac{\partial \mathrm{ad}_f^{k-1} e_\ell}{\partial x} f \right](x) = \left[\begin{bmatrix} M_1 & 0 \\ 0 & 0 \end{bmatrix} f \right](x)$$

$$= [\underbrace{\otimes \cdots \otimes}_{\tilde{n}} \; 0 \; \cdots \; 0]^{\mathsf{T}},$$

where M_1 is $\tilde{n} \times \tilde{n}$, with $\tilde{n} = \sum_{i=1}^k |\mathcal{V}_i|$. Note that the definition of $\tilde{n}$ follows from the node labeling and the fact that, for all $j \in \{1, \ldots, N-1\}$, the nodes in $\mathcal{V}_j$ can only be connected with at most $|\mathcal{V}_{j+1}|$ nodes in $\mathcal{V}_{j+1}$. Thus, since the first term of each Lie bracket (8.4) does not affect the last $|\mathcal{V}_{j+1}|$ entries of the second term, the equation $E_{j,j}(x) = L_{j-1}(x) E_{j-1,j-1}(x)$ follows by direct computation of the second term in (8.4), and the claimed statement follows. ∎

In brief, the result in Lemma 49 states that each new $\mathscr{G}_k(x)$ "discovers" $|\mathcal{V}_{k+1}|$ new columns of $G_k(x)$ containing the entries of the block $E_{k+1,k+1}(x)$. We use this finding to prove our main result.

Proof of Theorem 47: For some nonnegative integer k, we let $\mathscr{G}_k = \mathrm{span}\{\mathrm{ad}_f^r e_\ell \; : \; 0 \le r \le k, 1 \le \ell \le m\}$. The system (8.1) is feedback-linearizable if and only if the following three conditions hold [226, Theorem 5.2.3]:

(i) $\mathscr{G}_{n-1}$ has dimension n at $\bar{x}$;

(ii) for each $0 \le k \le n-1$, $\mathscr{G}_k$ has constant dimension in a neighborhood of $\bar{x}$; and

(iii) for each $0 \leq k \leq n - 2$, $\mathscr{G}_k$ is involutive,

where $n = |\mathcal{V}|$.

Condition (i) follows directly from Lemma 49 and from the assumption $\mathrm{rank}(L_i(\bar{x})) = |\mathcal{V}_{i+1}|$ by noticing that each diagonal block $E_{i,i}(\bar{x})$ is full rank for all $i \in \{1, \ldots, k+1\}$. More in detail, let $P \in \mathbb{R}^{n_1 \times n_2}$ and $Q \in \mathbb{R}^{n_2 \times n_2}$; the fact $\mathrm{rank}(PQ) = \mathrm{rank}(P)$ if $\mathrm{rank}(Q) = n_2$ implies the full rank of $E_{i,i}(\bar{x}) = L_{i-1}(x)E_{i-1,i-1}(\bar{x})$.

In regards to condition (ii), from the assumption $\mathrm{rank}(L_i(\bar{x})) = |\mathcal{V}_{i+1}|$, the continuity of singular values, and the definition of $E_{i,i}$ in Lemma 49, it follows that the blocks $E_{i,i}(x)$ are full row rank in a neighborhood of $\bar{x}$. Thus, in such a neighborhood, the definition of the matrix G_k ensures that the dimension of $\mathscr{G}_k$, for each $0 \leq k \leq n - 1$, remains constant, and condition (ii) holds.

Finally, to show that condition (iii) holds, we observe that due to the structure of $\mathscr{G}_k(x)$ in Lemma 49, the Lie brackets between any two vector fields in $\mathscr{G}_k(x)$, $0 \leq k \leq N - 1$, cannot have nonzero rows greater than $\tilde{n} = \sum_{i=1}^{k+1} |\mathcal{V}_i|$. Thus, the Lie bracket $\mathrm{ad}_f^{j-1} e_\ell$ can only contain states up to $\tilde{n}$, and $\mathscr{G}_k(x)$ must be involutive for all $0 \leq k \leq N - 1$. Note that, for $k > N - 1$, $\mathscr{G}_k$ clearly remains involutive. This concludes the proof. ∎

A few comments are in order. First, Theorem 47 is constructive, can be used to check whether a network is feedback-linearizable, and also to design networks that satisfy such a property. Second, Theorem 47 implies that there exist a linearizing feedback of the form $u = \alpha(x) + \beta(x)v$ and a diffeomorphism $z = \Phi(x)$ that solve the state-space exact linearization problem [226, §5.2]. That is, it is possible to transform (8.1) into an equivalent linear system $\dot{z} = A_{\mathrm{lin}}z + B_{\mathrm{lin}}v$ with rank $\begin{bmatrix} B_{\mathrm{lin}} & A_{\mathrm{lin}}B_{\mathrm{lin}} & \ldots & A_{\mathrm{lin}}^{n-1}B_{\mathrm{lin}} \end{bmatrix} = n$. Thus, being the pair $(A_{\mathrm{lin}}, B_{\mathrm{lin}})$ controllable, Theorem 47 can also be used to assess local controllability of (8.1) at $\bar{x}$.

Corollary 50 (*Condition for nonlinear controllability*) *Consider the dynamics (8.1), and let* $\bar{x}$ *be such that* $f(\bar{x}) = 0$. *If the lower diagonal blocks* $L_i(\bar{x})$, $i = \{1, \ldots, N-1\}$, *of the Jacobian (8.3) have full row rank, then the system (8.1) is locally controllable at* $\bar{x}$.

$$\dot{x}_1 = x_3(1 + x_2) + u$$
$$\dot{x}_2 = x_1$$
$$\dot{x}_3 = x_2(1 + x_1)$$
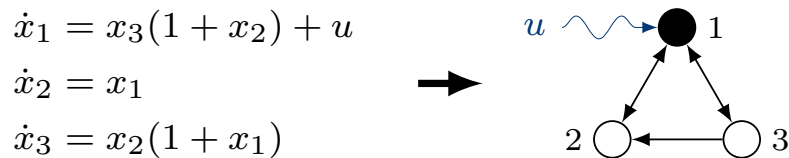
Figure 8.3: This figure illustrates the inference diagram obtained from the system in Example 52. The controlled node, which corresponds to x_1, is filled in black. Since $|\{2, 3\}| > |\{1\}|$, the condition in Theorem 47 cannot be satisfied. Yet, the system is feedback-linearizable, as we show in Example 52.

Notice that, because the linear system $\dot{z} = A_{\text{lin}}z + B_{\text{lin}}v$ is defined in an open set that depends on the nonlinear feedback and change of coordinates, we can evaluate the size of the region of the state space in which the linear transformation holds. In fact, if $\Phi(x)$, $\alpha(x)$, and $\beta(x)$ are defined in an open neighborhood $\mathcal{U}$ of $\bar{x}$, then the linear system is defined in the open set $\Phi(\mathcal{U})$. This fact enables the exact characterization of the operating regions for the feedback-linearized system. Conversely, Jacobian linearization is only exact at the equilibrium point at which the Jacobian matrix is computed. It is also worth noting that Jacobian linearization does not yield the same system as the feedback-linearized one.

Example 51 (*Difference in local controllability between Jacobian and feedback lineariza-tion*) *Consider the system (8.1), with $f(x) = \begin{bmatrix} 0 & x_1 & x_2 + x_1^2/2 \end{bmatrix}^{\mathsf{T}}$ and $b = e_1$. It is easy to see that the controllability matrix of the linearized system at the origin is full rank. Thus, the system is locally controllable around $\bar{x} = [0\ 0\ 0]^{\mathsf{T}}$. Yet, since the distribution $\mathcal{G}_1(x) = \text{span}\{e_1, [f, e_1]\}$ is not involutive, by [226, Theorem 5.2.3], the system is not feedback-linearizable at $\bar{x}$.* □

Is should be noted that the condition in Theorem 47 enables a general structural approach to assess feedback-linearizability of nonlinear systems with dynamics (8.1). Yet, it is only sufficient, as we show in the next example.

Example 52 (*Non-necessity of Theorem 47*) *Consider the system in Fig. 8.3 with control vector $b = e_1$. Notice that such a system does not satisfy Theorem 47 at $\bar{x} = 0$. Yet, it can*

be verified that the matrix[2] $G_2(\bar{x}) = \{b, [f, b], [f, [f, b]]\}(\bar{x}) = \mathrm{diag}(1, -1, 1)$, is full rank, and that x

$\mathscr{G}_1 = span\{b, [f, b]\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & x_2 \end{bmatrix}^\mathsf{T}$ is involutive. Thus, by [226, Theorem 5.2.3], this system is feedback

linearizable at the origin. □

8.4 Application to cluster synchronization of nonlinearly coupled oscillators

In this section, we apply the results developed in Section 8.3 to the important problem of controlling the emergence of clusters of neural units with synchronized activity. Specifically, we use feedback linearization to divide networks of nonlinear oscillators into distinct synchronized groups. We achieve this goal without resorting to the prescriptive conditions required by previous work [142, 143].

8.4.1 Network of Kuramoto dynamics

Patterns of correlated activity among neural units play a role in the correct execution of cognitive functions and in the abnormal dynamics of a class of neurological disorders, such as Parkinson's disease and epilepsy [161, 162]. A classical model used to represent the oscillatory behavior of brain activity is the Kuramoto model of diffusively-coupled oscillators [147, 214]. The dynamics of the i-th oscillator is

$$\dot{\theta}_i = \omega + \sum_{j \neq i} h_{ij} \sin(\theta_j - \theta_i), \tag{8.5}$$

where ω represents the natural frequency of the oscillators, θ_i is the phase of the i-th oscillator, and $h_{ij} \in \mathbb{R}_{>0}$ denotes the coupling strength between interconnected oscillators, with $\mathbb{R}_{>0}$ indicating the set of positive real numbers.

Interconnected systems of diffusively-coupled oscillators evolve on a network that is described by a weighted adjacency matrix $H = [h_{ij}]$, where $h_{ij} \in \mathbb{R}_{>0}$ if $(j, i) \in \mathcal{E}$, and $h_{ij} = 0$

---
[2]We denote with $\mathrm{diag}(\cdot)$ a diagonal matrix.
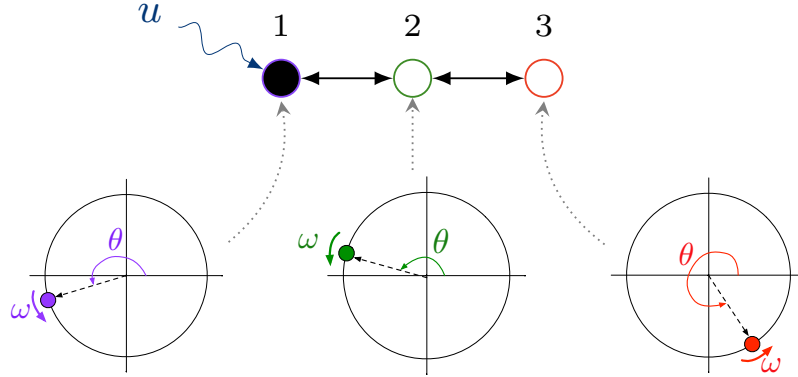
Figure 8.4: This figure represents the network of oscillators studied in Section 8.4.1. Each node represents an oscillator, whose dynamics are defined by the natural frequency ω and the interactions with neighboring oscillators. The first node receives the control input u, and is filled in black to represent $\mathcal{K} = \{1\}$. We choose a simple topology because it allows us to easily illustrate the results of Section 8.3, which can be extended to larger systems.

otherwise. Here, we choose the vector field

$$
f(\theta) = \begin{bmatrix} \omega + h_{12}\sin(\theta_2 - \theta_1) \\ \omega + h_{21}\sin(\theta_1 - \theta_2) + h_{23}\sin(\theta_3 - \theta_2) \\ \omega + h_{32}\sin(\theta_2 - \theta_3) \end{bmatrix},
$$

where $h_{12} = h_{21} = 1$, $h_{23} = h_{32} = 2$, and the control input is injected into the first oscillator through a control vector field $b = [1\ 0\ 0]^\mathsf{T}$. It is worth noting that for all diffusively-coupled oscillators, the inference diagram is also described by the adjacency matrix H after all weights are binarized.[3] The adjacency matrix for this illustrative example reads as $A = \begin{bmatrix} e_2 & e_1 + e_3 & e_2 \end{bmatrix}$. and the network is depicted in Fig. 8.4.

With the aid of a rotating reference frame with angular velocity ω, we can study the equilibria of (8.5) as fixed points. We choose to stabilize the unstable equilibrium $\bar{\theta} = [\pi\ \pi\ 0]^\mathsf{T}$, but we remark that this network satisfies Theorem 47 at all equilibria. We start the derivation of the linearizing feedback law from the Lie brackets $\mathrm{ad}_f b$ and $\mathrm{ad}_f^2 b$, which read as $\mathrm{ad}_f b =$

---

[3]That is, the weight associated with the edge (j, i) in the inference diagram adjacency matrix is set to 1 if $h_{ij} > 0$, and zero otherwise.

$$\begin{bmatrix} h_{12}\cos(\theta_2 - \theta_1) & -h_{21}\cos(\theta_1 - \theta_2) & 0 \end{bmatrix}^{\mathsf{T}}, \text{ and}$$

$$\mathrm{ad}_f^2 b = \begin{bmatrix} h_{12}(h_{12} + h_{21} - h_{23}\sin(\theta_2 - \theta_1)\sin(\theta_3 - \theta_2)) \\ \\ \star \\ \\ -h_{32}h_{21}\cos(\theta_1 - \theta_2)\cos(\theta_2 - \theta_3) \end{bmatrix},$$

with $\star = h_{21}(h_{12}\sin^2(\theta_2 - \theta_1) + h_{21} + 2h_{23}\sin(\theta_1 - \theta_2)\sin(\theta_3 - \theta_2) - h_{12}\cos^2(\theta_2 - \theta_1))$. To linearize

the Kuramoto network around $\bar{\theta}$, we must compute the linearizing feedback $u = \alpha(\theta) + \beta(\theta)v$,

where $\alpha(\theta)$ and $\beta(\theta)$ can be derived from [226, §4]: $\alpha(\theta) = -\frac{L_f^n \Phi_1}{L_b L_f^{n-1}\Phi_1}$, $\beta(\theta) = \frac{1}{L_b L_f^{n-1}\Phi_1}$, with

$\Phi_i(\theta) = L_f^{i-1}\Phi_1(\theta)$, and $\Phi_1(\theta)$ is such that

$$\frac{\partial \Phi_1}{\partial \theta}b = 0, \quad \frac{\partial \Phi_1}{\partial \theta}\mathrm{ad}_f b = 0, \quad \text{and} \quad \frac{\partial \Phi_1}{\partial \theta}\mathrm{ad}_f^2 b \neq 0. \tag{8.6}$$

The choice $\Phi_1(\theta) = z_1 = \theta_3$ satisfies (8.6), and yields $L_b L_f^2 \Phi_1 = h_{21}h_{32}\cos(\theta_1 - \theta_2)\cos(\theta_2 - \theta_3)$. From the latter, we derive $\alpha(\theta)$ and $\beta(\theta)$, which we omit here in the interest of space. As a

proof of concept, we can easily compute the state space region where the feedback-linearized system

is defined. Such a region corresponds the open set $\{\theta : |\theta_i - \bar{\theta}_i| < \frac{\pi}{2} \text{ for all } i\}$, the boundary points

of whose closure are the only coordinates for which $\alpha(\theta)$ and $\beta(\theta)$ are not defined. Finally, we can

assign the poles of $\dot{z} = A_{\mathrm{lin}}z + B_{\mathrm{lin}}v$ via classical static feedback $v = -Kz$ (see Fig. 8.5).

8.4.2 Multiplex network with multi-body interactions

Multiplex networks explicitly incorporate multiple channels of connectivity in a system. In

the following, we study a 2-layer multiplex network where the two layers contain neurons with three-

and two-body interaction, respectively [248, 249]. Importantly, three-body interactions are thought

to play a crucial role in heterosynaptic plasticity [250]. We represent neurons as oscillatory units

164
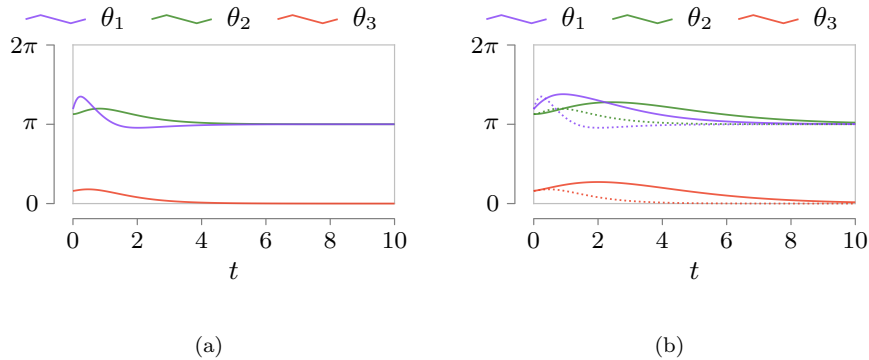
Figure 8.5: These figures show the phases evolution for the three oscillators in Fig. 8.4 in a reference framework that rotates with natural frequency ω. Fig.8.5(a) illustrates the convergence of the phases to the equilibrium $\bar{\theta} = [\pi\ \pi\ 0]^{\mathsf{T}}$ after a stabilizing feedback $v = -Kz$ is applied to the system. The eigenvalues for the matrix $A_{\mathrm{lin}} + bK$ are -1, -2, and -3, which are obtained with the feedback gains $K = [6\ 11\ 6]$. Fig.8.5(b) illustrates different convergence rates associated with different eigenvalues of $A_{\mathrm{lin}} + bK$. The dotted lines represent the phases of the three oscillators when the eigenvalues are the same as in Fig.8.5(a), whereas the solid lines are associated with eigenvalues -0.6, -0.7, and -0.8, and hence with slower dynamics.

that obey the dynamics

$$\dot{\theta}_i = \omega + \frac{\kappa_1}{m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} \sin(\theta_j + \theta_k - 2\theta_i),$$

$$\dot{\varphi}_i = \omega + \frac{\kappa_2}{m} \sum_{j=1}^{m} \sin(\varphi_j - \varphi_i) + d\sin(\theta_i - \varphi_i),$$

(8.7)

where ω is the natural frequency, and $\kappa_1, \kappa_2, d \in \mathbb{R}_{>0}$ are coupling strengths. We choose $m = 3$ and we apply a control input to the θ layer of the network (see Fig. 8.6(a)), so that $B = [e_1, e_2, e_3]$. Further, we fix the constants $\kappa_1 = \kappa_2 = d = 1$. By writing the state of the network as $x = [\theta_1\ \theta_2\ \theta_3\ \varphi_1\ \varphi_2\ \varphi_3]^{\mathsf{T}}$, it can be shown that $\bar{x} = [\pi\ \pi\ \pi\ 0\ 0\ 0]^{\mathsf{T}}$ is an unstable equilibrium point of (8.7). It holds that $J(\bar{x}) = \begin{bmatrix} \otimes & 0 \\ \mathrm{diag}(-d,-d,-d) & \otimes \end{bmatrix}$.

Hence, the condition in Theorem 47 is satisfied and the system is feedback-linearizable at $\bar{x}$. Along the lines of [226, Proof of Lemma 5.2.2], we compute the dummy output functions $\lambda_1(x) = \varphi_1$, $\lambda_2(x) = \varphi_2$, and $\lambda_3(x) = \varphi_3$, from which we obtain the local change of coordinates $\phi_k^i(x) = L_f^{k-1}\lambda_i(x)$, $1 \leq k \leq 2$, $1 \leq i \leq 3$, and finally the linearizing feedback law [226, §5]. The code for this simulation can be downloaded at [251]. In Fig. 8.6(b), the computed feedback partitions the network into two distinct clusters. This result extends existing work on the control of cluster

165
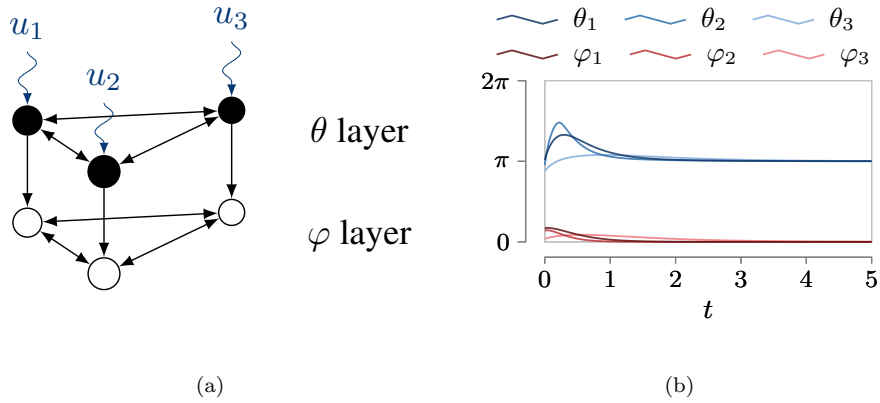
Figure 8.6: Fig. 8.6(a) illustrates the 2-layer multiplex network studied in Section 8.4.2. The θ layer consists of neurons connected with three-body interactions, whereas the φ layer consists of neurons connected with two-body interactions. We show in Fig. 8.6(b) the convergence of the system's state to the desired point after feedback linearization. The eigenvalues of $A_{\text{lin}} + B_{\text{lin}} K$ are assigned to the left half plane, thus making $\bar{x} = [\pi \ \pi \ \pi \ 0 \ 0 \ 0]^{\mathsf{T}}$ locally stable in the reference frame rotating with velocity ω.

synchronization in networks of nonlinearly-interacting oscillators (see, e.g., [214]) to more complex types of interactions.

8.5   Conclusion and future directions

We have derived sufficient structural conditions to test for feedback-linearizability of large-scale nonlinear network systems. These conditions also constitute a test for local controllability. The results contained in this chapter are particularly suited for large networks, such as multi-agent systems, neuronal networks, and artificial neural networks.

Because of striking topological similarities, future research could investigate the link between feedback-linearizable and strongly structurally controllable networks [252]. Finally, another interesting research direction is the characterization of the gap between locally controllable and feedback-linearizable systems [253].

Chapter 9

A Framework to Control

Functional Connectivity

In this chapter, we propose a framework to control brain-wide functional connectivity by selectively acting on the brain's structure and parameters. Functional connectivity, which measures the degree of correlation between neural activities in different brain regions, can be used to distinguish between healthy and certain diseased brain dynamics and, possibly, as a control parameter to restore healthy functions. We refer the reader to [214] for the complete work.

9.1   Introduction

The structural (i.e., matrix of anatomical connections between brain regions) and functional (i.e., matrix of correlation coefficients between the activity of brain regions) connectivity of the brain vary across healthy individuals and those affected by neurological or psychiatric disorders, and can be used as biomarkers to detect or predict pathological conditions. While structural connectivity changes rather slowly over time and can be measured accurately via diffusion imaging techniques [46], functional connectivity depends on the instantaneous neural activity and is affected, for instance,

by the tasks being performed and external stimuli [254]. Today, common measures of functional connectivity rely on resting-state functional magnetic resonance imaging (rs-fMRI) timeseries to quantify the level of correlated activity between brain regions. The relationships between structural and functional connectivity have recently received considerable attention [59,255], and the tantalizing idea of controlling functional states by leveraging or modifying brain structure has given birth to a new, thrilling, field of research [10, 12, 114].

In this paper, we leverage the connection between structural and functional connectivity, and propose a framework to control functional connectivity by selectively modifying structural connectivity and the regions' intrinsic frequencies (see Fig. 9.1). In particular, building on prior studies [147,165], we model the brain's neural activity as the phases of a collection of interconnected Kuramoto oscillators, and postulate that the level of functional connection between two regions is proportional to the level of synchronization between the phases of the oscillators associated with the two regions. Then, we derive conditions and methods to tune the oscillators' interconnection weights and natural frequencies so as to enforce arbitrary synchronization patterns and, consequently, brain-wide functional connectivity. We remark that the control mechanisms used in our framework are biologically plausible. For instance, changes in the spontaneous neural activity (i.e., oscillators' frequencies) are typical of the brain, involve natural modifications in regional metabolism of the neurons, and can alternatively be induced by a number of non-invasive stimulation techniques [256]. Changes to the structural interconnections (i.e., oscillators' interconnections), instead, can arise from different chemical or electrical mechanisms including, at the microscale, Hebbian plasticity [257] and short-term synaptic facilitation [258].

Related work. The discovery of oscillatory or rhythmic brain activity dates back almost a century. Yet, control-theoretic studies that exhaust the oscillatory nature of brain states have been sparse and of relatively recent date. Some authors focus on localized desynchronization of neural activity [173,259,260], which is desirable in individuals affected by epilepsy or Parkinson's disease, and others
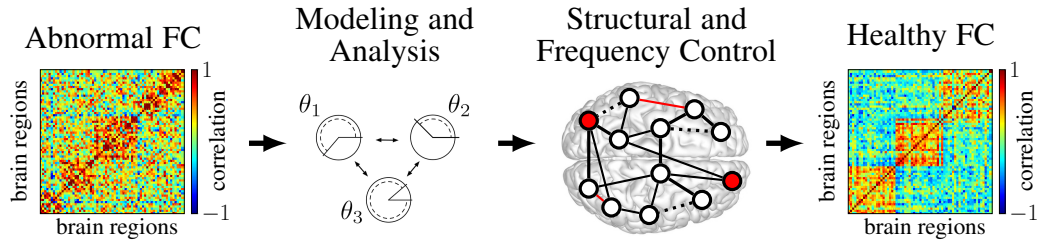
Figure 9.1: This chapter proposes a framework to restore healthy patterns of brain-wide functional connectivity by selectively acting on the brain's structure and parameters. Using a network of heterogeneous Kuramoto oscillators to model the brain's neural activity, we design and validate a minimally invasive method to correct the oscillators' interconnections and frequencies to obtain a desired and stable pattern of functional connectivity.

use synchronization phenomena to describe cognitive and functional brain states [182, 261, 262]. To the best of our knowledge, a framework to control the pattern of brain-wide functional connectivity is still missing, and is proposed for the first time in this chapter.

At the core of our framework to model and control functional connectivity is the concept of *cluster synchronization* in a network of oscillators, where groups of oscillators behave cohesively but independently from other clusters. For the case of oscillators with Kuramoto dynamics as used in this work, [181, 218] explore approximate notions of cluster synchronization in simplified configurations, while [184] provides exact invariance conditions for arbitrary cluster synchronization manifolds. Our recent work introduces rigorous [263] and approximate [143] stability conditions for cluster synchronization, which are also used here. Compared to the above references, this chapter focuses on the control of cluster synchronization, rather than on its enabling conditions.

Chapter contribution. The contributions of this chapter are twofold. On the technical side, we formulate and solve a network optimization problem to enforce stable cluster synchronization among interconnected Kuramoto oscillators (Section 9.3). We provide a two-step procedure to compute the smallest (as measured by the Frobenius norm) perturbation of the network weights and the oscillators' natural frequencies so as to achieve a desired and arbitrary synchronization pattern. Notably, the proposed algorithm allows for the modification of only a selected subset of the network parameters, as typically constrained in applications. We also prove that cluster synchronization

is robust to parameter mismatches and numerical inaccuracies, which complements the theoretical derivations in [143, 263], and strengthen the applicability of our control methods to work in practice.

On the application side, this work contains the first mathematically rigorous and neurologically plausible framework to control functional connectivity in the brain, and takes a significant step to fill the gap between empirical studies on oscillatory neural activity [147, 165, 264] and the recent technical body of work inspired by neural synchronization [143, 181, 184, 263]. In Section 9.5, we apply our control technique to an empirically-reconstructed structural brain network, and validate our results by computing the correlation of resting-state fMRI signals obtained through a realistic hemodynamic model. As a minor contribution, our work extends [147] by allowing heterogeneous Kuramoto dynamics.

Mathematical notation. The sets $\mathbb{R}_{>0}$, $\mathbb{S}^1$ and $\mathbb{T}^n$ denote the positive real numbers, the unit circle, and the n-dimensional torus, respectively. We represent the vector of all ones with $\mathbb{1}$. The Frobenius and ℓ_2 norms are denoted as $\|\cdot\|_{\mathrm{F}}$ and $\|\cdot\|$, respectively, and $A \circ B$ is the Hadamard product between matrices A and B. A (block-)diagonal matrix is denoted by (blk)diag$(\cdot)$. We let $\mathrm{i} = \sqrt{-1}$. Let $A \geq 0$ represent an element-wise inequality on the entries of A, A^+ the element-wise nonnegative part of A, and $A \succ 0$ a positive definite matrix A. We let $\lambda_i(A)$ and $\sigma_i(A)$ denote the i-th eigenvalue and the i-th singular value of $A \in \mathbb{R}^{n \times n}$, respectively, and $\lambda_{\max}(A) = \max_i |\lambda_i(A)|$ and $\lambda_{\min}(A) = \min_i |\lambda_i(A)|$. Finally, we let $\overline{\lambda}(A) = \frac{1}{n} \sum_i \lambda_i(A)$ and $\overline{\sigma}(A) = \frac{1}{n} \sum_i \sigma_i(A)$.

9.2 Problem setup and preliminary notions

The aim of this work is to control network parameters so that groups of brain regions exhibit a high degree of functional connectivity. In this context, functional interactions are defined as the pairwise correlation between hemodynamic signals recorded in two brain regions. One model used to simulate such hemodynamic signals is described by a set of nonlinear differential equations [190] that can be approximated in the frequency domain as a linear low-pass filter [147]. Because the

only input to such hemodynamic model is the oscillatory neural activity, the formation of strongly (functionally) connected brain regions can be promoted by controlling the synchronization level of their neural dynamics. We follow [147] to model such neural dynamics with a sparse network of heterogeneous Kuramoto oscillators that are connected to each other according to the anatomical architecture of the human brain, more specifically known as white matter tracts.[1] Ultimately, the problem of generating desired patterns of functional connectivity reduces to the one of controlling cluster synchronization in a network of heterogeneous Kuramoto oscillators.

To be precise, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a weighted digraph, where $\mathcal{V} = \{1, \ldots, n\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represent the oscillators, or nodes, and their interconnection edges, respectively. The i-th oscillator's dynamics reads as:

$$\dot{\theta}_i = \omega_i + \sum_{j \neq i} a_{ij} \sin(\theta_j - \theta_i), \tag{9.1}$$

where $\omega_i \in \mathbb{R}_{>0}$ denotes the natural frequency of the i-th oscillator, $\theta_i \in \mathbb{S}^1$ is its phase, $a_{ij} \in \mathbb{R}_{>0}$ is the weight of the edge $(j, i) \in \mathcal{E}$, with $a_{ij} = 0$ when $(j, i) \notin \mathcal{E}$, and $A = [a_{ij}]$ is the weighted adjacency matrix of $\mathcal{G}$.

To characterize synchronized trajectories among subsets of oscillators, let $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$ be a nontrivial partition of $\mathcal{V}$, where each cluster contains at least two oscillators and its graph is strongly connected.[2] We say that a network exhibits cluster synchronization when the oscillators can be partitioned so that the phases of the oscillators in each cluster evolve identically. Formally, we define the *cluster synchronization manifold* associated with the partition $\mathcal{P}$ as

$$\mathcal{S}_{\mathcal{P}} = \{\theta \in \mathbb{T}^n \ : \ \theta_i = \theta_j \text{ for all } i, j \in \mathcal{P}_k, k = 1, \ldots, m\}.$$

Then, the network is cluster-synchronized with partition $\mathcal{P}$ when the phases of the oscillators belong to $\mathcal{S}_{\mathcal{P}}$ at all times. Without loss of generality, the oscillators are labeled so that $\mathcal{P}_k = \{\sum_{\ell=1}^{k-1} |\mathcal{P}_\ell| + 1, \ldots, \sum_{\ell=1}^{k} |\mathcal{P}_\ell|\}$, where $|\mathcal{P}_\ell|$ denotes the cardinality of the set $\mathcal{P}_\ell$.

---

[1] We assume that at each node of a structural brain network there exists a community of excitatory and inhibitory neurons whose dynamical state is in a regime of self-sustained oscillation. In other words, the neurons' firing rates delineate a limit cycle, and their dynamics can be approximated by a single variable, which is the angle (or phase) on this cycle.

[2] As the brain is densely connected [265], this assumption is not restrictive.

Because our control framework leverages conditions for the invariance and stability of the cluster synchronization manifold to modify the network weights and oscillators' natural frequencies, we briefly recall useful preliminary results that have recently been established in [143, 184, 263]. Specifically, given a desired network partition $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$, invariance of $\mathcal{S}_\mathcal{P}$ is guaranteed by the following conditions:

(C1) The natural frequencies satisfy $\omega_i = \omega_j$ for every $i, j \in \mathcal{P}_k$ and $k \in \{1, \ldots, m\}$. Equivalently,

$$B_{\mathrm{span}}^\mathsf{T} \omega = 0,$$

where $B_{\mathrm{span}} \in \mathbb{R}^{|\mathcal{V}| \times |\bigcup_k \mathcal{E}_{\mathrm{span},k}|}$ is the incidence matrix of $\bigcup_{k=1}^m \mathcal{T}_k$, with $\mathcal{T}_k = (\mathcal{P}_k, \mathcal{E}_{\mathrm{span},k})$ being a spanning tree of the digraph $\mathcal{G}_k$ of the isolated cluster $\mathcal{P}_k$;

(C2) The network weights satisfy $\bar{V}_\mathcal{P}^\mathsf{T} \bar{A} V_\mathcal{P} = 0$,

where $V_\mathcal{P} \in \mathbb{R}^{n \times m}$ is the characteristic matrix of the network defined as $V_\mathcal{P} = [v_1/\|v_1\|, \ldots, v_m/\|v_m\|]$, with

$$v_i^\mathsf{T} = [\, \underbrace{0, \ldots, 0}_{\sum_{j=1}^{i-1} |\mathcal{P}_j|}, \underbrace{1, \ldots, 1}_{|\mathcal{P}_i|}, \underbrace{0, \ldots, 0}_{\sum_{j=i+1}^{n} |\mathcal{P}_j|} \,],$$

$\bar{V}_\mathcal{P} \in \mathbb{R}^{n \times (n-m)}$ is an orthonormal basis of the orthogonal subspace to the image of $V_\mathcal{P}$, and $\bar{A} = A - A \circ V_\mathcal{P} V_\mathcal{P}^\mathsf{T}$ is the matrix of inter-cluster connections only (see also [184]).

We assume that the *isolated* clusters are locally stable:

(A1) The dynamics (9.1), with $a_{ij} = 0$ when i, j belong to different clusters, converges exponentially fast to $\mathcal{S}_\mathcal{P}$.

Notice that Assumption (A1) is satisfied when $\mathcal{G}_k$ has symmetric weights and condition (C1) holds [263, Lemma 3.1] [28, Theorem 5.1]. In our case, (A1) is not restrictive because structural brain networks are typically symmetric [10].[3]

---

[3]In a general case, one can ensure that Assumption (A1) is satisfied simply by pairing the control mechanism developed in the next session with an independent one that makes intra-cluster connections symmetric.

Let $\omega^{(k\ell)}$ denote the natural frequency difference between any two nodes in disjoint clusters $\mathcal{P}_k$ and $\mathcal{P}_\ell$. If (C1) and (C2) hold, then a tight approximate condition for $\mathcal{S}_\mathcal{P}$ to be locally exponentially stable is [143]:

(C3) The natural frequencies and the network weights are such that $\lambda_{\max}(\Xi(A,\omega)) < 1$, with $\Xi = [\xi_{k\ell}]$ and

$$
\xi_{k\ell} = \begin{cases} \nu_{k\ell}\overline{\sigma}(G_k(\mathrm{i}\,\omega^{(k\ell)})), & \text{if } \overline{\lambda}(J_k) \leq \overline{\lambda}(J_\ell), \\[2ex] \nu_{k\ell}\dfrac{\overline{\sigma}(G_k(0))}{\overline{\sigma}(G_\ell(0))}\overline{\sigma}(G_\ell(\mathrm{i}\,\omega^{(k\ell)})), & \text{if } \overline{\lambda}(J_\ell) < \overline{\lambda}(J_k), \end{cases} \tag{9.2}
$$

where $G_k(s) = (sI - J_k)^{-1}$, J_k is the Hurwitz stable Jacobian matrix of the intra-cluster phase difference dynamics and $\nu_{k\ell}$ is a function of the inter-cluster weights. We refer the interested reader to the previous chapters (Section II of this thesis) for a detailed discussion on conditions (C1), (C2), and (C3).

9.3 Control of cluster synchronization

In this section, we propose a control mechanism to obtain a prescribed and robust configuration of synchronized oscillatory patterns. Towards this aim, we consider a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and an arbitrary partition $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$ of $\mathcal{V}$. The proposed control technique is minimally invasive in the sense that it looks for the smallest correction (in the Frobenius norm sense) of inter-cluster network weights and oscillators' natural frequencies that renders the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$ invariant and locally stable. In practice, a modification of the network parameters will require either the exploitation of neural plasticity or localized surgical intervention for the modification of the network weights and structure, and pharmacological or electromagnetic influence for the refinement of the brain regions' natural frequencies. In mathematical terms, the approach is

encoded into solving the following constrained minimization problem:

$$\min_{\Delta, \mu} \ \|[\Delta, \mu]\|_{\mathrm{F}}^2 \tag{9.3}$$

$$\text{s.t.} \ \ \bar{V}_{\mathcal{P}}^{\mathsf{T}}(\bar{A} + \Delta)V_{\mathcal{P}} = 0, \tag{9.3a}$$

$$B_{\mathrm{span}}^{\mathsf{T}}(\omega + \mu) = 0, \tag{9.3b}$$

$$H^{\mathrm{c}} \circ \Delta = 0 \tag{9.3c}$$

$$\bar{A} + \Delta \geq 0, \tag{9.3d}$$

$$\omega + \mu \geq 0, \tag{9.3e}$$

$$\lambda_{\mathrm{max}}(\Xi(A + \Delta, \omega + \mu)) < 1, \tag{9.3f}$$

where Δ is the correction of the network matrix, μ is the correction of the natural frequencies vector, and the (i, j)-th entry of Δ is zero if i, j belong to the same partition $\mathcal{P}_k$, $k \in \{1, \ldots, m\}$. Further, H is the 0-1 adjacency matrix of $\mathcal{H} = (\mathcal{V}, \mathcal{E}_{\mathcal{H}})$, which is the graph encoding the set of edges $\mathcal{E}_{\mathcal{H}} \subseteq \mathcal{E}$ that is allowed to be modified, and $H^{\mathrm{c}} = \mathbb{1}\mathbb{1}^{\mathsf{T}} - H$. That is, the (i, j)-th entry of a solution Δ^* to problem (9.3) is zero when the corresponding (i, j)-th entry of H is zero. The optimization problem (9.3) is illustrated in Fig. 9.2.

Constraints (9.3a) and (9.3b) are equivalent to conditions (C2) and (C1), respectively, for the invariance of $\mathcal{S}_{\mathcal{P}}$. Constraint (9.3c) restricts the corrective action to a subset of all the possible interconnections, in affinity with the practical limitations of localized interventions. Constraints (9.3d) and (9.3e) are due to biological compatibility and require the inter-cluster weights of the perturbed network and oscillators' natural frequencies to be nonnegative. Finally, Constraint (9.3f) corresponds to (C3) and guarantees the (local) stability of $\mathcal{S}_{\mathcal{P}}$. In particular, the latter constraint makes the above problem non-convex and, therefore, potentially intractable from a numerical viewpoint. To overcome this issue, we next propose a suboptimal, yet numerically more tractable, control strategy. Specifically, we decouple (9.3) into two simpler subproblems. The first one solves for the
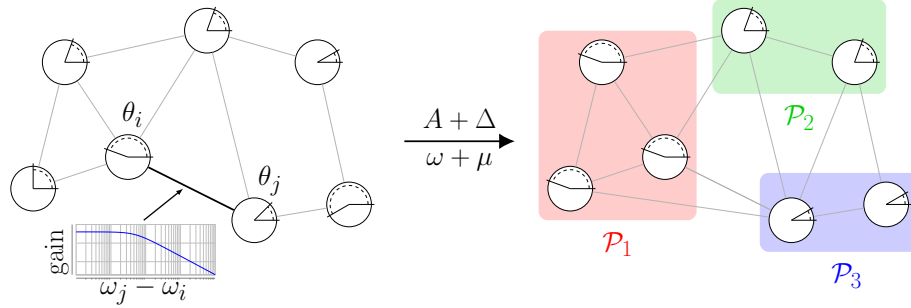
Figure 9.2: The left depicts a network of oscillators. The coupling strength between the oscillators depends on the network weights and the differences of their natural frequencies [143]. The optimization problem (9.3) seeks for the smallest modification of the network weights and the oscillators' natural frequencies to ensure a desired stable pattern of cluster synchronization (right panel). We remark that the techniques used in this chapter for cluster synchronization in frequency-weighted networks of Kuramoto oscillators, are applicable to a broad class of network optimization problems, e.g., [51].

smallest correction of inter-cluster weights satisfying (9.3a), (9.3c), and (9.3d), whereas the second one solves for the smallest correction of the oscillators' natural frequencies satisfying (9.3b), (9.3e) and (9.3f).

9.3.1 Inter-cluster structural control for invariance of $\mathcal{S}_{\mathcal{P}}$

We first address the problem of computing the smallest correction of inter-cluster weights such that constraints (9.3a), (9.3c), and (9.3d) are satisfied. Specifically, we focus on the following minimization problem:

$$\min_{\Delta} \ \|\Delta\|_{\mathrm{F}}^2 \tag{9.4}$$

$$\text{s.t.} \ \ \bar{V}_{\mathcal{P}}^{\mathsf{T}}(\bar{A} + \Delta)V_{\mathcal{P}} = 0, \tag{9.4a}$$

$$H^{\mathrm{c}} \circ \Delta = 0, \tag{9.4b}$$

$$\bar{A} + \Delta \geq 0. \tag{9.4c}$$

The optimization problem (9.4) is convex and, when feasible, it can be efficiently solved by means of standard optimization techniques. Feasibility of (9.4) depends on the constraint graph $\mathcal{H}$ (see Remark 55). In what follows, we present a simple and efficient projection-based algorithm to solve

this problem.

Theorem 53 (Smallest sparse inter-cluster correction) *Assume that the problem (9.4) is feasible, and consider the matrix sequence $\{Z_k\}_{k \geq 0}$ generated via the following iterative procedure:*

$$Y_k = H \circ (Z_k + T_k)^+ + H^c \circ \bar{A},$$

$$T_{k+1} = Z_k + T_k - Y_k,$$

$$Z_{k+1} = Y_k + Q_k - \bar{V}_{\mathcal{P}} \bar{V}_{\mathcal{P}}^{\mathsf{T}} (Y_k + Q_k) V_{\mathcal{P}} V_{\mathcal{P}}^{\mathsf{T}}, \qquad (9.5)$$

$$Q_{k+1} = Y_k + Q_k - Z_{k+1},$$

where $Z_0 = \bar{A}$, and $T_0 = Q_0 = 0$. Then, the sequence $\{Z_k\}_{k \geq 0}$ converges to a matrix Z^*, and a minimizer of (9.4) subject to (9.4a), (9.4b), and (9.4c), has the form $\Delta^* = Z^* - \bar{A}$.

Before proving Theorem 53, we introduce the following instrumental lemma.

Lemma 54 *Consider a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, and an arbitrary (nontrivial) partition $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$ of $\mathcal{V}$. Let $W \in \mathbb{R}^{n \times n}$ and consider the minimization problem*

$$\min_Z \|Z - W\|_F^2 \qquad (9.6)$$

$$s.t. \quad \bar{V}_{\mathcal{P}}^{\mathsf{T}} Z V_{\mathcal{P}} = 0, \qquad (9.6a)$$

The minimizer of the problem (9.6) subject to (9.6a) is

$$Z^* = W - \bar{V}_{\mathcal{P}} \bar{V}_{\mathcal{P}}^{\mathsf{T}} W V_{\mathcal{P}} V_{\mathcal{P}}^{\mathsf{T}}. \qquad (9.7)$$

Proof. We prove the result via the method of Lagrange multipliers. The Lagrangian of (9.6) subject to (9.6a) is $\mathcal{L}(Z, \Lambda) = \|Z - W\|_F^2 + \mathbb{1}^{\mathsf{T}} (\Lambda \circ \bar{V}_{\mathcal{P}}^{\mathsf{T}} Z V_{\mathcal{P}}) \mathbb{1} = \operatorname{tr}((Z - W)^{\mathsf{T}} (Z - W)) + \operatorname{tr}(\Lambda^{\mathsf{T}} \bar{V}_{\mathcal{P}}^{\mathsf{T}} Z V_{\mathcal{P}})$, where $\Lambda \in \mathbb{R}^{(n-m) \times m}$ is a matrix of Lagrange multipliers associated with Constraint (9.6a), and in the last equation we used that $\mathbb{1}^{\mathsf{T}} (A \circ B) \mathbb{1} = \operatorname{tr}(A^{\mathsf{T}} B)$. Equating the partial derivatives of $\mathcal{L}$ to zero yelds:

$$\frac{\partial \mathcal{L}}{\partial Z} = 2(Z - W) + \bar{V}_{\mathcal{P}} \Lambda V_{\mathcal{P}}^{\mathsf{T}} = 0, \qquad (9.8)$$

$$\frac{\partial \mathcal{L}}{\partial \Lambda} = \bar{V}_{\mathcal{P}}^{\mathsf{T}} Z V_{\mathcal{P}} = 0, \qquad (9.9)$$

We next pre- and post-multiply both sides of (9.8) by $\bar{V}_{\mathcal{P}}^{\mathsf{T}}$ and $V_{\mathcal{P}}$, respectively, and obtain

$$2\bar{V}_{\mathcal{P}}^{\mathsf{T}} Z V_{\mathcal{P}} = 2\bar{V}_{\mathcal{P}}^{\mathsf{T}} W V_{\mathcal{P}} - \bar{V}_{\mathcal{P}}^{\mathsf{T}} \bar{V}_{\mathcal{P}} \Lambda V_{\mathcal{P}}^{\mathsf{T}} V_{\mathcal{P}}$$

$$\Rightarrow \; 2\bar{V}_{\mathcal{P}}^{\mathsf{T}} Z V_{\mathcal{P}} = 2\bar{V}_{\mathcal{P}}^{\mathsf{T}} W V_{\mathcal{P}} - \Lambda \; \Rightarrow \; \Lambda = 2\bar{V}_{\mathcal{P}}^{\mathsf{T}} W V_{\mathcal{P}}, \tag{9.10}$$

where in the second implication we used $V_{\mathcal{P}}^{\mathsf{T}} V_{\mathcal{P}} = I_n$, $\bar{V}_{\mathcal{P}}^{\mathsf{T}} \bar{V}_{\mathcal{P}} = I_{n-m}$, and $\bar{V}_{\mathcal{P}}^{\mathsf{T}} V_{\mathcal{P}} = 0$, and in the last one we used (9.9). Finally, (9.7) follows by substituting (9.10) into (9.8). ∎

We are now ready to prove Theorem 53

Proof of Theorem 53: Let $\Pi_{\mathcal{Z}}(W) = \arg\min_{Z \in \mathcal{Z}} \|Z - W\|_{\mathrm{F}}^2$ denote the projection (in the Frobenius norm sense) of W onto a convex set $\mathcal{Z}$, and define the closed convex sets $\mathcal{Z}_1 = \{Z \in \mathbb{R}^{n \times n} \; : \; Z \geq 0 \text{ and } H^c \circ Z = \bar{A}\}$ and $\mathcal{Z}_2 = \{Z \in \mathbb{R}^{n \times n} \; : \; \bar{V}_{\mathcal{P}}^{\mathsf{T}} Z V_{\mathcal{P}} = 0\}$. Note that $\Pi_{\mathcal{Z}_1}(W) = H \circ W^+ + H^c \circ \bar{A}$ and, by Lemma 54 in the Appendix,

$$\Pi_{\mathcal{Z}_2}(W) = \arg\min_{Z \in \mathcal{Z}_2} \|Z - W\|_{\mathrm{F}}^2 = W - \bar{V}_{\mathcal{P}} \bar{V}_{\mathcal{P}}^{\mathsf{T}} W V_{\mathcal{P}} V_{\mathcal{P}}^{\mathsf{T}},$$

for any W. Hence, the sequence $\{Z_k\}_{k \geq 0}$ generated by (9.5) coincides with the sequence generated by Dykstra's projection algorithm [266] applied to the projections onto $\mathcal{Z}_1$ and $\mathcal{Z}_2$. Since the problem (9.4) is feasible, $\mathcal{Z}_1 \cap \mathcal{Z}_2 \neq \emptyset$, and the latter sequence converges to a matrix $Z^* = \Pi_{\mathcal{Z}_1 \cap \mathcal{Z}_2}(\bar{A}) = \arg\min_{Z \in \mathcal{Z}_1 \cap \mathcal{Z}_2} \|Z - \bar{A}\|_{\mathrm{F}}^2$ [266]. Finally,

$$Z^* = \arg\min_{Z \in \mathcal{Z}_1 \cap \mathcal{Z}_2} \|Z - \bar{A}\|_{\mathrm{F}}^2 = \bar{A} + \arg\min_{\Delta \text{ s.t. } (9.4a),\,(9.4b),\,(9.4c)} \|\Delta\|_{\mathrm{F}}^2,$$

and the statement follows. ∎

Remark 55 (Sufficient condition for the feasibility of (9.4)**)** *Recall from [184] that condition (C2) is equivalent to:*

$$\sum_{k \in \mathcal{P}_\ell} a_{ik} - a_{jk} = 0 \tag{9.11}$$

for every $i, j \in \mathcal{P}_z$ and for all $z, \ell \in \{1, \ldots, m\}$, with $z \neq \ell$. Notice that, if for every $i \in \mathcal{P}_z$ there exists at least one $(k, i) \in \mathcal{E}_{\mathcal{H}}$, $k \in \mathcal{P}_\ell$, a solution to (9.11) can always be found and problem (9.4) is feasible. □

9.3.2   Frequency tuning for local stability of $\mathcal{S}_\mathcal{P}$

We now turn to the problem of computing the smallest correction of natural frequencies such that constraints (9.3b), (9.3e), and (9.3f) are satisfied. That is,

$$\min_{\mu} \; \|\mu\|_\mathrm{F}^2 \tag{9.12}$$

$$\text{s.t.} \;\; B_{\mathrm{span}}^{\mathsf{T}}(\omega + \mu) = 0, \tag{9.12a}$$

$$\omega + \mu \geq 0, \tag{9.12b}$$

$$\lambda_{\max}(\Xi(A, \omega + \mu)) < 1. \tag{9.12c}$$

Theorem 56 (Feasibility of problem (9.12)) *There always exists a correction μ satisfying* (9.12a), (9.12b), *and* (9.12c).

Proof. Consider the vector $\mu = [\mu_1, \ldots, \mu_n]^{\mathsf{T}}$. Note that we can find some μ_i such that (i) $\omega_i + \mu_i = \omega_j + \mu_j > 0$ for all $i, j \in \mathcal{P}_k$, $k \in \{1, \ldots, m\}$, and (ii) $|\omega_i + \mu_i - (\omega_j + \mu_j)| > \eta$, for all $i \in \mathcal{P}_k$, $j \in \mathcal{P}_\ell$, $k, \ell \in \{1, \ldots, m\}$, $k \neq \ell$, and $\eta > 0$ arbitrarily large. From (i), μ satisfies (9.12a) and (9.12b). Further, since each nonzero entry of $\Xi(A, \omega + \mu)$ in (9.2) behaves as a low-pass filter, by fact (ii) $\lambda_{\max}(\Xi(A, \omega + \mu))$ can be made arbitrarily small. This implies that there always exists a vector μ satisfying (9.12c) and concludes the proof. ∎

An optimal solution to (9.12) is typically difficult to compute, because of the eigenvalue constraint (9.12c). However, several heuristics can be used to compute a suboptimal correction in (9.12). For instance, we next outline an effective procedure to find a suboptimal solution to (9.12). Let $\omega_{\mathrm{av}}^{(k)} = \frac{1}{|\mathcal{P}_k|} \sum_{i \in \mathcal{P}_k} \omega_i$ denote the average frequency within each cluster, and let

$$\omega_{\mathrm{av}} = [\omega_{\mathrm{av},1}, \ldots, \omega_{\mathrm{av},n}]^{\mathsf{T}} = [\underbrace{\omega_{\mathrm{av}}^{(1)}, \ldots, \omega_{\mathrm{av}}^{(1)}}_{|\mathcal{P}_1|}, \ldots, \underbrace{\omega_{\mathrm{av}}^{(m)}, \ldots, \omega_{\mathrm{av}}^{(m)}}_{|\mathcal{P}_m|}]^{\mathsf{T}}.$$

Further, define the quotient graph $\mathcal{Q} = (\mathcal{V}', \mathcal{E}')$ where each node in $\mathcal{V}'$ represents a cluster and each edge in $\mathcal{E}'$ an interconnection between two clusters. Our procedure leverages Theorem 56 and increases the frequency differences between pairs of connected clusters until constraint (9.12c) is satisfied. The procedure consists of four steps:

1. If $\lambda_{\max}(\Xi(A, \omega_{\mathrm{av}})) < 1$, then $\mu^* = [\mu_1^*, \ldots, \mu_n^*]$, with $\mu_i^* = \omega_{\mathrm{av},i} - \omega_i$ is an optimal correction to (9.12). Otherwise, proceed to the next step.

2. Construct a depth-first spanning tree $\mathcal{T}_{\mathcal{Q}}$ of $\mathcal{Q}$ rooted at $r = \arg\min_k \omega_{\mathrm{av}}^{(k)}$.[4]

3. Assign the frequency $\omega(k, \alpha) = \omega^{(r)} + k\alpha$, $\alpha > 0$, to each node of each cluster in $\mathcal{T}_{\mathcal{Q}}$ of depth k, $k = 1, 2, \ldots, k_{\max}$, where $k_{\max}$ denotes the height of $\mathcal{T}_{\mathcal{Q}}$.[5] Let $\omega(\alpha) = [\omega_1(\alpha), \ldots, \omega_n(\alpha)]^{\mathsf{T}}$ denote the resulting vector of modified frequencies.

4. Find the smallest α^* satisfying $\lambda_{\max}(A, \omega(\alpha^*)) < 1$. Then, $\mu^* = [\mu_1^*, \ldots, \mu_n^*]$, with $\mu_i^* = \omega_i(\alpha^*) - \omega_i$, is a (suboptimal) solution to (9.12).

9.4 Robustness of the control framework

In this section, we show that the control framework described in Section 9.3, and in fact the stability property of the cluster synchronization manifold $\mathcal{S}_{\mathcal{P}}$, is robust to perturbations of the network parameters. That is, small changes in the oscillators' natural frequencies and network weights yield a small deviation from cluster-synchronized trajectories. In light of this, the proposed control mechanism lends itself to practical applications, where the network parameters are not known exactly and the neural dynamics is subject to noise.

Consider the dynamics (9.1) with perturbed parameters:

$$\dot{\theta}_i = \tilde{\omega}_i + \sum_{j \neq i} \tilde{a}_{ij} \sin(\theta_j - \theta_i), \tag{9.13}$$

where $\tilde{\omega}_i = \omega_i + \delta\omega_i$ and $\tilde{a}_{ij} = a_{ij} + \delta a_{ij}$. Notice that, if $\delta\omega_i = 0$ and $\delta a_{ij} = 0$, the dynamics (9.13) is equivalent to (9.1). From (9.13), the perturbed intra-cluster difference dynamics of nodes $i, j \in \mathcal{P}_k$,

---

[4]Notice that such a spanning tree always exists, since $\mathcal{Q}$ is connected.
[5]Given a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and a spanning tree $\mathcal{T}$ of $\mathcal{G}$ rooted at $r \in \mathcal{V}$, the depth of a node $v \in \mathcal{V}$ is the length of the path in $\mathcal{T}$ from r to v, and the height of $\mathcal{T}$ is the maximum depth among the nodes in $\mathcal{V}$.

with $k \in \{1, \ldots, m\}$, reads as:

$$\dot{\theta}_j - \dot{\theta}_i = \omega_j + \delta\omega_i - \omega_i - \delta\omega_j$$

$$+ \sum_{z=1}^{n} [(a_{jz} + \delta a_{jz}) \sin(\theta_z - \theta_j) - (a_{iz} + \delta a_{iz}) \sin(\theta_z - \theta_i)]$$

$$= \omega_j - \omega_i + \sum_{z=1}^{n} [a_{jz} \sin(\theta_z - \theta_j) - a_{iz} \sin(\theta_z - \theta_i)] + \delta_{ij}, \tag{9.14}$$

where $\delta_{ij} = \delta\omega_j - \delta\omega_i + \sum_{z=1}^{n} [\delta a_{jz} \sin(\theta_z - \theta_j) - \delta a_{iz} \sin(\theta_z - \theta_i)]$. Finally, let δ be the vector of all δ_{ij} that affect the nominal intra-cluster dynamics as in (9.14).

We are now ready to present the main result of this section, which resorts to the prescriptive stability condition derived in [263].

Theorem 57 *(Robustness of cluster synchronization) Assume that the network weights satisfy Theorem 25, and consider any pair of nodes $i, j \in \mathcal{P}_k$, $k \in \{1, \ldots, m\}$. Then, for some finite $T > 0$ and for all initial conditions such that $|\theta_j(0) - \theta_i(0)| < \varepsilon$, with $\varepsilon > 0$ sufficiently small, the solution to the perturbed dynamics (9.13) satisfies*

$$|\theta_j(t) - \theta_i(t)| \le c\gamma \quad \forall t \ge T, \tag{9.15}$$

where $\gamma = \max_{\theta \in [0, 2\pi)} \|\delta\|$, and c is a constant that depends only on the network weights.

Proof. In the first part of the proof, we combine the Lyapunov functions for the isolated clusters $\mathcal{P}_k$, $k = 1, \ldots, m$, into a Lyapunov function for the intra-cluster differences dynamics of the whole network. In the second part of the proof, we show that such Lyapunov function satisfies certain bounds, so that the application of [194, Lemma 9.2] suffices to prove the claimed statement.

We let $x_{ij} = \theta_j - \theta_i$, and S, x_{intra} and J_k be as in the Appendix **??**. To combine the Lyapunov functions of the isolated clusters, we note that if S is an M-matrix, then, along the lines of [263, Proof of Theorem 3.2], the origin of the nominal intra-cluster dynamics of x_{intra} is locally exponentially stable with Lyapunov function

$$V(x_{\text{intra}}) = \sum_{k=1}^{m} d_k x_{\text{intra}}^{(k)\mathsf{T}} P_k x_{\text{intra}}^{(k)}, \tag{9.16}$$

where $P_k \succ 0$ satisfies $J_k^\mathsf{T} P_k + P_k J_k = -I$, and $d_k > 0$ are such that $DS + S^\mathsf{T} D \succ 0$, with $D = \text{diag}(d_1 \ldots, d_m)$ [194].

Consider now the Lyapunov function (9.16), and notice that: $c_1 \|x_{\text{intra}}\|^2 \leq \sum_{k=1}^m d_k x_{\text{intra}}^{(k)\mathsf{T}} P_k x_{\text{intra}}^{(k)} \leq c_2 \|x_{\text{intra}}\|^2$, with $c_1 = m d_{\min} \min_k \lambda_{\min}(P_k)$ and $c_2 = m d_{\max} \max_k \lambda_{\max}(P_k)$. Further, in the ball of radius r of the origin $\mathcal{B}_r = \{x_{\text{intra}} : \|x_{\text{intra}}\| < r, \dot{V}(x_{\text{intra}}) < 0\}$, it holds that $\dot{V}(x_{\text{intra}}) \leq -c_3 \|x_{\text{intra}}\|^2$, with $c_3 = \lambda_{\min}(DS + S^\mathsf{T} D)/2$. To see this, consider the derivative of the Lyapunov function $V(x_{\text{intra}})$ along the trajectories of the nominal system. Then, from [194, §9.5] and [263, proof of Theorem 3.2], the following inequality holds in $\mathcal{B}_r$:

$$\dot{V} \leq -\frac{1}{2} \begin{bmatrix} \|x_{\text{intra}}^{(1)}\| & \cdots & \|x_{\text{intra}}^{(m)}\| \end{bmatrix} (DS + S^\mathsf{T} D) \begin{bmatrix} \|x_{\text{intra}}^{(1)}\| \\ \vdots \\ \|x_{\text{intra}}^{(m)}\| \end{bmatrix} \leq -\frac{1}{2} \lambda_{\min}(DS + S^\mathsf{T} D) \|x_{\text{intra}}\|^2,$$

and c_3 follows. Further, since $\|\partial V / \partial x_{\text{intra}}\| = \|2 x_{\text{intra}}^\mathsf{T} P_k\| < 2 \lambda_{\max}(P_k) \|x_{\text{intra}}\|$, we have $\|\partial V / \partial x_{\text{intra}}\| \leq c_4 \|x_{\text{intra}}\|$, with $c_4 = 2 c_2$. Finally, once the constants c_1, c_2, c_3, and c_4 are computed, the definition of x_{intra} and [194, Lemma 9.2] conclude the proof. ∎

Importantly, Theorem 57 can be used to provide a quantitative bound on the asymptotic value of $|\theta_j - \theta_i|$. In fact, we can compute the constant c in (9.15) by exploiting [194, Lemma 9.2] and c_1, c_2, c_3, c_4 derived in the above proof.

9.5 Control of functional connectivity in an empirically reconstructed brain network

We conclude this chapter with the application of the control mechanism presented in Section 9.3 to the brain network estimated in [46], which is publicly available at: `http://umcd.humanconnectomeproject.org/umcd`. In these data, structural connectivity is proportional to large-scale connection pathways between cortical regions, and the gray matter is subdivided into $n = 66$ cortical regions (33 per hemisphere). To show the effectiveness of our proposed

method in enforcing desired functional connectivity by means of arbitrary synchronization patterns, we partition the structural brain network in 3 clusters, i.e. $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3\}$, each one comprising 22 regions that do not belong to any known functionally connected resting-state network. The three clusters are highlighted with different colors in Fig. 9.3(a) and Fig. 9.3(b). Furthermore, following our goal of providing a method to enhance the synchronization properties of a diseased or damaged brain, we simulate the effects of brain damage, e.g., a stroke, by damping the connectivity of one cluster [267]. That is, we weaken the intra-cluster connections of the first cluster by a scaling factor 10^{-2} to echo reduced structural connectivity, and we show that our technique can in fact recover stability of the cluster synchronization manifold associated with the desired network partition.[6]

Before presenting our results, we describe the methodology used to simulate human rs-fMRI functional connectivity.

9.5.1 Simulation of functional connectivity

The brain's neural activity is simulated through a network of coupled Kuramoto oscillators, where we randomly draw the natural frequencies of each oscillator from a uniform distribution in the range $[0, 60]$ [Hz] so as to include all meaningful neural frequency bands [190]. We set the initial phases in the interval $[0, 0.5]$ [rad]. The Kuramoto phases act as an input to the neurovascular coupling, which is modeled by the Balloon-Windkessel hemodynamic process [268], and whose output is the blood-oxygen-level dependent (BOLD) signal that is measured by rs-fMRI.

The neuronal activity z_i of the i-th brain region produces an increase in a vasodilatory signal s_i, which is subject to auto-regulatory feedback. The inflow f_i responds in proportion to this signal with concomitant changes in blood volume μ_i and deoxyhemoglobin content q_i. Mathematically, the

---

[6]Specifically, weakening the intra-cluster connections of one cluster is likely to make $\mathcal{S}_\mathcal{P}$ unstable [143].

(a)   (b)

(c)   (d)
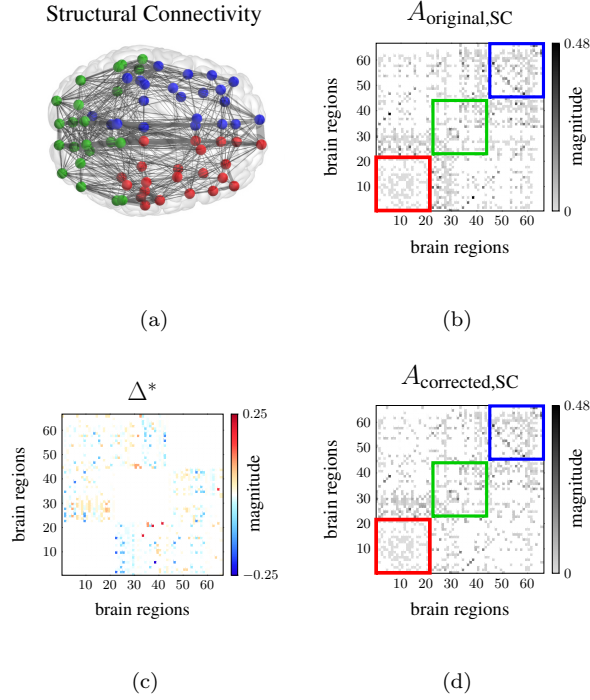
Figure 9.3: Fig. 9.3(a) depicts an axial view of the structural connectivity estimated in [46], and was obtained with *BrainNet Viewer* [52]. The edge thickness is proportional to the number of white matter streamlines connecting different regions. Fig. 9.3(b) represents the adjacency matrix of the structural brain network in Fig. 9.3(a), where the white entries correspond to zero, and the intra-cluster connections in the first cluster (red nodes in Fig. 9.3(a)) have been weakened to simulate the effect of brain damage. Fig. 9.3(c) represents the network matrix correction Δ^* solution to the iterative procedure (9.5) in Theorem 53. Finally, Fig. 9.3(d) represents the matrix $A_{\text{corrected,SC}} = A_{\text{original,SC}} + \Delta^*$, where the total change of the edge weights amounts to 17% (in the Frobenius norm) of the original ones.

dynamics of these quantities reads as:

$$\dot{s}_i = z_i - \kappa_i s_i - \gamma_i (f_i - 1), \quad \dot{f}_i = s_i,$$

$$\tau \dot{\mu}_i = f_i - \mu_i^{1/\alpha}, \quad \tau_i \dot{q}_i = f_i E(f_i, \rho_i)/\rho_i - \mu_i^{1/\alpha} q_i/\mu_i.$$

The oxygen extraction is a function of the flow $E(f, \rho) = 1 - (1 - \rho^f)$ where ρ denotes the resting oxygen extraction fraction. The biophysical parameters $\kappa, \gamma, \tau, \alpha,$ and ρ are exhaustively treated in [268]. Finally, the BOLD signal is described as a static nonlinear function:

$$y_i = V_0 (k_1(1 - q_i) + k_2(1 - q_i/v_i) + k_3(1 - \mu_i)),$$

where $V_0 = 0.02$ denotes the resting blood volume fraction, and $k_1 = 7\rho_i$, $k_2 = 2$, $k_3 = 2\rho_i - 0.2$.
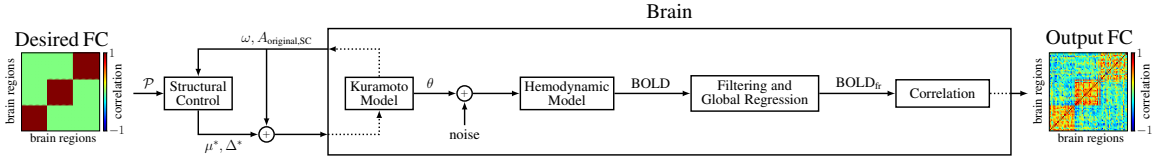
Figure 9.4: Schematic illustrating the pipeline to obtain desired functional connectivity (FC) from a structural connectivity matrix $A_{\text{original,SC}}$, the original natural frequencies ω, and a desired network partition $\mathcal{P}$. The addition of noise to the synchronized neural dynamics θ represents the presence of background noise. The output matrix depicts the brain regions' functional connectivity simulated by computing the correlations of filtered and regressed BOLD signals.

Following [147], we choose $z_i = \sin(\theta_i)$. Further, to account for the presence of background noise in the brain, we add white noise to the neural activity z_i with variance 10^{-2}. We simulate 2 minutes of BOLD signals and process the timeseries as explained below in order to compute functional connectivity estimates that closely resemble that of human rs-fMRI recordings.

To reduce the effect of spurious correlations from small and non-physiological high-frequency components, we filter the synthetic BOLD signals through a low-pass filter. Consequently, to improve the correspondence between resting-state correlations and anatomical connectivity, we process all of the simulated regional BOLD signals by a global signal regression [269] that averages the timeseries of all regions by removing spontaneous oscillations common to the whole brain. Next, we discard the first 40 seconds of all timeseries to eliminate the effect of initial transients. Finally, we compute the Pearson correlation of the filtered and regressed signals to obtain the synthetic functional connectivity. A pipeline describing the above process is illustrated in Fig. 9.4.

9.5.2 Application of the clustering control mechanism

In the remainder of this section, we apply the control method proposed in Section 9.3. We first solve the minimization problem (9.4) to find the optimal correction matrix Δ^* to be applied to $A_{\text{original,SC}}$ such that condition (C2) for the invariance of $\mathcal{S}_\mathcal{P}$ is satisfied. We choose to constrain the corrective action on a set of edges $\mathcal{E}_\mathcal{H} = \mathcal{E} \cup \tilde{\mathcal{E}}$ that includes the original set $\mathcal{E}$ and a minimal set $\tilde{\mathcal{E}}$ of randomly selected edges such that problem (9.4) is feasible (see Remark 55). Fig. 9.3(c) and
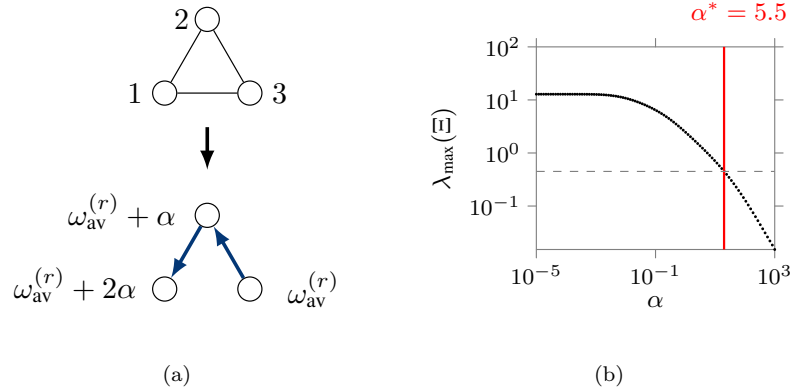
Figure 9.5: Fig. 9.5(a) depicts the quotient graph associated with the three clusters in partition $\mathcal{P}$ and the natural frequencies that follow from the procedure in Section 9.3.2. Fig. 9.5(b) shows the profile of $\lambda_{\max}(\Xi)$ as a function of the tuning parameter α on a logarithmic scale. The thick red line highlights the smallest value α^* for which the local stability of the cluster synchronization manifold $\mathcal{S}_\mathcal{P}$ is guaranteed according to condition (C3).

9.3(d) illustrate the corrective action Δ^* and the network matrix $A_{\text{corrected,SC}} = A_{\text{original,SC}} + \Delta^*$, respectively.

We proceed with the frequency tuning technique for invariance and stability of $\mathcal{S}_\mathcal{P}$ to $A_{\text{corrected,SC}}$ so that conditions (C1) and (C3) are satisfied. The first step involves computing the mean natural frequency $\omega_{\text{av}}^{(k)}$ among all oscillators belonging to the same cluster $\mathcal{P}_k$: $\omega_{\text{av}}^{(1)} = 199.2$, $\omega_{\text{av}}^{(2)} = 182.9$ and $\omega_{\text{av}}^{(3)} = 115.4$ [rad/s]. Next, we apply the procedure proposed in Section 9.3.2. We plot in Fig. 9.5(a) the spanning tree of the quotient graph $\mathcal{T}_\mathcal{Q}$, and in Fig. 9.5(b) the optimal α^* computed in step (iv) of our procedure. The final natural frequencies are $\omega^{(1)} = 131.8$, $\omega^{(2)} = 126.4$ and $\omega^{(3)} = 115.4$ [rad/s]. Notice that, although our frequency tuning procedure is sub-optimal, the outcome values remain well within the range of brain activity frequency bands and, based on numerical results, outperform the results of Matlab's `fmincon` function.

Finally, by following the pipeline described in the previous subsection, we compute the desired functional connectivity pattern, which we show in Fig. 9.6. Notably, the functional connectivity of the desired clusters is strong and the correlations between different clusters are negligible. Thus, the proposed method to control synchronization patterns of oscillatory neural activity lends
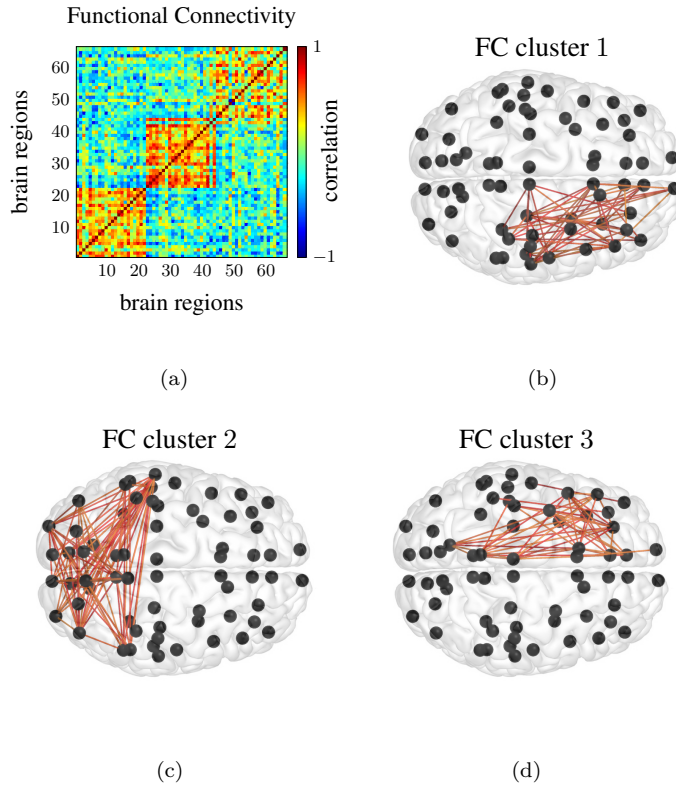
Figure 9.6: Fig. 9.6(a) represents the correlation matrix that encodes the output functional connectivity (FC) obtained with our control mechanism. Notably, the three clusters are mostly functionally disconnected. That is, there are very few functional connections between nodes belonging to different clusters. This implies that the outcome of our procedure is robust to noisy neural activity and faithfully reproduces synchronized BOLD signals. Fig. 9.6(b),9.6(c) and 9.6(d) illustrate the isolated functional connectivity of the desired clusters $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$, respectively, after the correlation matrix has been thresholded to 0.5 to show only the meaningful functional interactions. The functional edges are color-coded according to the colorbar of Fig. 9.6(a).

itself to a physiologically plausible framework and shows rather promising results.

9.6 Conclusion

In this work, we propose a minimally invasive technique to obtain robust synchronization patterns in sparse networks of heterogeneous Kuramoto oscillators. To the best of our knowledge, this is the first attempt at blending mathematically rigorous methods with physiological models of brain activity with the goal of steering whole-brain synchronization dynamics. Specifically, we cast a constrained optimization problem whose solution not only satisfies mathematical conditions for

invariance and stability of an arbitrary cluster synchronization manifold, but also meets biological constraints. We decompose the complete optimization problem into two simpler subproblems, and provide efficient methods to solve them. When applying our technique to correct the network parameters of empirically-reconstructed anatomical brain data, we find that our solution, although suboptimal, provides a result that is well within the range of physiologically plausible parameters. Additionally, we show that cluster synchronization is robust to small parameter mismatches and numerical inaccuracies. This result complements previous prescriptive studies on cluster synchronization and enables the use of our framework in practical situations.

Chapter 10

Conclusion

"There is nothing like looking,

if you want to find something.

You certainly usually find something,

if you look, but it is not always

quite the something you were after."

J. R. R. TOLKIEN [270, Ch. 4]

More often than not, the process of scientific discovery of a graduate student shares many similarities with the quest of the fictional characters in [270]. *Bilbo Baggins* has no idea that, by leaving the comfort of his home to help the band of *Thorin Oakenshield*, he will come into possession of immense riches and a magic ring that will set an extraordinary chain of events in motion. Akin to Bilbo's journey, this dissertation started with the rather "straightforward" goal of exploring the complex nature of the structure-function relationship in neural systems. Yet, working towards this goal, we have derived exciting theoretical results and discovered foundational working principles that find application not only in neural systems, but also in many other engineered and natural systems.

In the first part of this manuscript, we have explored the manifest controllability properties exhibited by state-of-the-art empirically reconstructed structural brain networks. By focusing on

the structural properties of brain networks, we have laid the ground to motivate further exploration of the structural backbone that supports brain functions, and to consider the brain as a dynamical system whose dynamics obey network-wide principles. Next, we have turned our attention to synchronization phenomena, which emerge ubiquitously in neural systems. Over short time frames, the brain is capable of exhibiting a rich repertoire of synchronization patterns while its anatomical interconnection scheme remains essentially unaltered. Such synchronization patterns not only underlie multiple cognitive processes, but can also be used as biomarkers for different psychiatric and neurological disorders. These observations motivated the final part of this dissertation, where we have designed control methods that allow us to optimally intervene on (neural) synchronization phenomena. We conclude this manuscript with some final remarks on each of these topics, together with a discussion of possible extensions.

10.1 Application of Control and Graph-Theoretic Methods To Structural Brain Networks

In Chapter 3, we have used a mathematically grounded approach to analyze whether structural brain networks are controllable from any single region. By resorting to the notion of structural controllability, we have demonstrated that networks with symmetric weights are structurally controllable from any region if and only if they are structurally controllable with asymmetric weights. We have applied our result to empirically reconstructed brain networks and found that they are all structurally controllable from any region. This result has important consequences for the study of cognitive control and exogenous interventions on brain states.

In Chapter 4, we have further investigated the controllability properties of structural brain networks. In this chapter, we have focused our attention on the controllability radius of networks, which measures how close the realization of a network is from uncontrollability. We have derived a method to compute the controllability radius for networks with symmetric weights and applied

our method to a set of empirically reconstructed brain networks. When compared with randomized networks, we found brain networks to have a consistently smaller controllability radius. Such a result may be compatible with substantial evidence supporting the hypothesis that the brain operates *at the edge of criticality* [271]. A possible direction for future research would explore the connection between criticality of the brain standard operating regime and its controllability properties.

10.2 Reverse Engineering Brain Network Dynamics Synchronization

We have considered two complementary approaches to investigate different synchrony configurations of brain activity: a bottom-up and a top-down explorations of the relationship between data, models, and paradigms of neural synchronization. In the bottom-up approach (presented in Chapter 5), we have inferred a data-driven dynamical model, used it to define brain dynamics fingerprints, and assessed the quality of multi-site repositories of brain activity recordings. Instead, in the top-down approach (presented in Chapter 6 and Chapter 7), we have made use of oscillator networks and mathematical abstractions to reverse engineer the synchronization of oscillatory brain rhythms. To do so, we have relied on previous extensive literature that promotes the use of Kuramoto phase oscillators to model brain activity [147, 166, 211, 213, 272]. The interaction between static large-scale structural architecture of the human brain and local oscillations of neural communities is a key factor in the functional connectivity patterns that are empirically observed in brain data. Building upon these assumptions and observations, we have explored the comprehensive framework of cluster synchronization and derived conditions that guarantee the emergence of such a phenomenon in networks of Kuramoto oscillators.

Possible extensions of our work may use higher-dimensional oscillators with a state that comprises phase and amplitude. For instance, we have studied the mechanisms underlying cross-frequency phase-amplitude coupling in Stuart-Laundau oscillators, which describes the phenomenon

where the power of a high-frequency oscillation evolves with the phase of a low-frequency one [273], and is relevant during memory and learning tasks. Further, the relationship between phase and amplitude is significantly altered between conscious and unconscious (e.g., anesthesia) states [274]. Finally, another possible extension would be the investigation of cluster synchronization in oscillator networks with asymmetric couplings.

10.3   Control of Brain Network Dynamics Synchronization

The final part of this dissertation relies on the insights from the previous chapters to derive control methods that allow us to prescribed desired functional relationships between neural units. Chapter 8 consists of a first attempt at dealing with the nonlinear dynamics of the Kuramoto equation. We have contributed to the challenging area of nonlinear systems control by deriving conditions for feedback linearization of network systems and applying our results to networks of oscillators in order to achieve desired steady state regimes.

In Chapter 9 we have used use a collection of interconnected Kuramoto oscillators to model oscillatory neural activity, and have developed an optimal control method to intervene on their functional connectivity, which is essentially regulated by the degree of synchronization between different clusters of oscillators. Our method computes the smallest adjustment of the network parameters that enforces stable cluster-synchronized trajectories. As a proof of concept, we have applied our procedure to an *in silico* model of whole-brain neural dynamics.

The above methods to control functional relations between components of a network have motivated our present efforts, which involve holistic control of synchronization phenomena in multiple domains [275]. Possible extensions in the area of synchronization control are multiple, and range from the expansion of the developed techniques to more realistic (although more complex) models of neuronal dynamics, to the engineering of exogenous signals to control state transitions.

Bibliography

[1] O. Sacks. The man who mistook his wife for a hat and other clinical tales, 1985.

[2] D.E. Bloom, E.T. Cafiero, E. Jané-Llopis, S. Abrahams-Gessel, L.R. Bloom, S. Fathima, A.B. Feigl, T. Gaziano, M. Mowafi, A. Pandya, K. Prettner, L. Rosenberg, B. Seligman, A.Z. Stein, and C. Weinstein. The global economic burden of noncommunicable diseases. *Geneva: World Economic Forum*, 2011.

[3] R. P. Feldman and J. T. Goodrich. The Edwin Smith surgical papyrus. *Child's Nervous System*, 15(6-7):281–284, 1999.

[4] T. Breitenfeld, M. J. Jurasic, and D. Breitenfeld. Hippocrates: the forefather of neurology. *Neurological Sciences*, 35(9):1349–1352, 2014.

[5] R. L. De Nó. Analysis of the activity of the chains of internuncial neurons. *Journal of Neurophysiology*, 1(3):207–244, 1938.

[6] T. G. Lewis. *Network science: Theory and applications*. John Wiley & Sons, 2011.

[7] T. Kailath. *Linear Systems*. Prentice-Hall, 1980.

[8] E Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.

[9] C. Tu, R. P. Rocha, M. Corbetta, S. Zampieri, M. Zorzi, and S. Suweis. Warnings and caveats in brain controllability. *NeuroImage*, 176:83–91, 2018.

[10] S. Gu, F. Pasqualetti, M. Cieslak, Q. K. Telesford, B. Y. Alfred, A. E. Kahn, J. D. Medaglia, J. M. Vettel, M. B. Miller, S. T. Grafton, and D. S. Bassett. Controllability of structural brain networks. *Nature Communications*, 6, 2015.

[11] T. Menara, S. Gu, D. S. Bassett, and F. Pasqualetti. On structural controllability of symmetric (brain) networks. *arXiv preprint arXiv:1706.05120*, 2017.

[12] T. Menara, D. S. Bassett, and F. Pasqualetti. Structural controllability of symmetric networks. *IEEE Transactions on Automatic Control*, 64(9):3740–3747, 2019.

[13] F. Pasqualetti, S. Gu, and D. S. Bassett. Re: Warnings and caveats in brain controllability. *NeuroImage*, 197:586–588, 2019.

[14] Y. Y. Liu, J. J. Slotine, and A. L. Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.

[15] G. Yan, J. Ren, Y.-C. Lai, C.-H. Lai, and B. Li. Controlling complex networks: How much energy is needed? *Physical Review Letters*, 108(21):218703, 2012.

[16] F.-J. Müller and A. Sachuppert. Few inputs can reprogram biological networks. *Nature*, 478(7369):E4–E4, 2011.

[17] F. Pasqualetti, S. Zampieri, and F. Bullo. Controllability metrics, limitations and algorithms for complex networks. *IEEE Transactions on Control of Network Systems*, 1(1):40–52, 2014.

[18] G. Notarstefano and G. Parlangeli. Controllability and observability of grid graphs via reduction and symmetries. *IEEE Transactions on Automatic Control*, 58(7):1719–1731, 2013.

[19] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt. Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM Journal on Control and Optimization*, 48(1):162–186, 2009.

[20] K. J. Reinschke. *Multivariable Control: A Graph-Theoretic Approach*. Springer, 1988.

[21] N. Monshizadeh, S. Zhang, and M. K. Camlibel. Zero forcing sets and controllability of dynamical systems defined on graphs. *IEEE Transactions on Automatic Control*, 59(9):2562–2567, 2014.

[22] S. D. Pequito, S. Kar, and A. P. Aguiar. A framework for structural input/output and control configuration selection in large-scale systems. *IEEE Transactions on Automatic Control*, 61(2):303–318, 2016.

[23] H. J. van Waarde, M. K. Camlibel, and H. L. Trentelman. A distance-based approach to strong target control of dynamical networks. *IEEE Transactions on Automatic Control*, 62(12):6266–6277, 2017.

[24] J. M. Dion, C. Commault, and J. van der Woude. Generic properties and control of linear structured systems: a survey. *Automatica*, 39(7):1125–1144, 2003.

[25] Z. Yuan, C. Zhao, Z. Di, W.-X. Wang, and Y.-C. Lai. Exact controllability of complex networks. *Nature communications*, 4, 2013.

[26] E. Scholtz. *Observer-based monitors and distributed wave controllers for electromechanical disturbances in power systems*. PhD thesis, Massachusetts Institute of Technology, 2004.

[27] F. Pasqualetti, F. Dörfler, and F. Bullo. Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems. *IEEE Control Systems Magazine*, 35(1):110–127, 2015.

[28] F. Dörfler and F. Bullo. Synchronization in complex networks of phase oscillators: A survey. *Automatica*, 50(6):1539–1564, 2014.

[29] F. Garin and L. Schenato. A survey on distributed estimation and control applications using linear consensus algorithms. In *Networked Control Systems*, LNCIS, pages 75–107. Springer, 2010.

[30] A. Chapman and M. Mesbahi. State controllability, output controllability and stabilizability of networks: A symmetry perspective. In *IEEE Conf. on Decision and Control*, pages 4776–4781, Osaka, Japan, 2015. IEEE.

[31] A. J. Whalen, S. N. Brennan, T. D. Sauer, and S. J. Schiff. Effects of symmetry on the structural controllability of neural networks: A perspective. In *American Control Conference*, pages 5785–5790, Boston, MA, USA, 2016.

[32] S. S. Mousavi, M. Haeri, and M. Mesbahi. On the structural and strong structural controllability of undirected networks. *IEEE Transactions on Automatic Control*, 63(7):2234–2241, 2017.

[33] F. Liu and A. S. Morse. Structural controllability of linear time-invariant systems. *arXiv preprint arXiv:1707.08243*, 2017.

[34] C. T. Lin. Structural controllability. *IEEE Transactions on Automatic Control*, 19(3):201–208, 1974.

[35] J. Sun and A. E. Motter. Controllability transition and nonlocality in network control. *Physical Review Letters*, 110(20):208701, 2013.

[36] W. M. Wonham. *Linear Multivariable Control: A Geometric Approach*. Springer, 3 edition, 1985.

[37] L. Markus and E. B. Lee. On the existence of optimal controls. *Journal of Basic Engineering*, 84(1):13–20, 1962.

[38] K. Tchoń. On generic properties of linear systems: An overview. *Kybernetika*, 19(6):467–474, 1983.

[39] E. Davison and S. Wang. Properties of linear time-invariant multivariable systems subject to arbitrary output and state feedback. *IEEE Transactions on Automatic Control*, 18(1):24–32, 1973.

[40] M. Egerstedt, S. Martini, M. Cao, K. Camlibel, and A. Bicchi. Interacting with networks: How does structure relate to controllability in single-leader, consensus networks? *IEEE Control Systems Magazine*, 32(4):66–73, 2012.

[41] F. Pasqualetti, S. Martini, and A. Bicchi. Steering a leader-follower team via linear consensus. In *Hybrid Systems: Computation and Control*, volume 4981, pages 642–645, Saint Louis, MO, April 2008.

[42] C. Godsil and G. F. Royle. *Algebraic Graph Theory*. Graduate Texts in Mathematics. Springer New York, 2001.

[43] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2001.

[44] H. Mayeda and T. Yamada. Strong structural controllability. *SIAM Journal on Control and Optimization*, 17(1):123–138, 1979.

[45] H. Mayeda. On structural controllability theorem. *IEEE Transactions on Automatic Control*, 26(3):795–798, 1981.

[46] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J Honey, V. J Wedeen, and O. Sporns. Mapping the structural core of human cerebral cortex. *PLoS Biology*, 6(7):e159, 2008.

[47] R. F. Galán. On how network architecture determines the dominant patterns of spontaneous neural activity. *PloS one*, 3(5):e2148, 2008.

[48] L. Wiles, S. Gu, F. Pasqualetti, B. Parvesse, D. Gabrieli, D. S. Bassett, and D. F. Meaney. Autaptic connections shift network excitability and bursting. *Scientific Reports*, 7, 2017.

[49] J. Goñi, M. P. Van Den Heuvel, A. Avena-Koenigsberger, N. V. de Mendizabal, R. F. Betzel, A. Griffa, P. Hagmann, B. Corominas-Murtra, J. Thiran, and O. Sporns. Resting-brain functional connectivity predicted by analytic measures of network communication. *Proceedings of the National Academy of Sciences*, 111(2):833–838, 2014.

[50] S. F. Muldoon, F. Pasqualetti, S. Gu, M. Cieslak, S. T. Grafton, J. M. Vettel, and D. S. Bassett. Stimulation-based control of dynamic brain networks. *PLoS Computational Biology*, 12(9):e1005076, 2016.

[51] T. Menara, V. Katewa, D. S. Bassett, and F. Pasqualetti. The structured controllability radius of symmetric (brain) networks. In *American Control Conference*, pages 2802–2807, Milwaukee, WI, USA, June 2018.

[52] M. Xia, J. Wang, and Y. He. Brainnet viewer: a network visualization tool for human brain connectomics. *PLoS ONE*, 8(7):e68910, 2013.

[53] K. Murota. *Systems analysis by graphs and matroids: structural solvability and controllability*, volume 3. Springer Science & Business Media, 2012.

[54] S. Gu, R. F. Betzel, M. G. Mattar, M. Cieslak, P. R. Delio, S. T. Grafton, F. Pasqualetti, and D. S. Bassett. Optimal trajectories of brain state transitions. *NeuroImage*, 148:305–317, 2017.

[55] D. Sritharan and S. V. Sarma. Fragility in dynamic networks: application to neural networks in the epileptic cortex. *Neural computation*, 26(10):2294–2327, 2014.

[56] S. Pequito, A. Ashourvan, D. S. Bassett, B. Litt, and G. J. Pappas. Spectral control of cortical activity. In *American Control Conference*, pages 2785–2791, Seattle, WA, USA, 2017. IEEE.

[57] D. S. Bassett and O. Sporns. Network neuroscience. *Nature Neuroscience*, 20(3):353–364, 2017.

[58] R. F. Betzel, S. Gu, J. D. Medaglia, F. Pasqualetti, and D. S. Bassett. Optimally controlling the human connectome: the role of network topology. *Scientific Reports*, 6:30770, 2016.

[59] C. O. Becker, S. Pequito, G. J. Pappas, M. B. Miller, S. T. Grafton, D. S. Bassett, and V. M. Preciado. Spectral mapping of brain functional connectivity from diffusion imaging. *Scientific Reports*, 8(1):1411, 2018.

[60] S. Martini, M. Egerstedt, and A. Bicchi. Controllability analysis of multi-agent systems using relaxed equitable partitions. *International Journal of Systems, Control and Communications*, 2(1-3):100–121, 2010.

[61] O. Sporns. Structure and function of complex brain networks. *Dialogues in Clinical Neuroscience*, 15(3):247, 2013.

[62] R. E. Kalman, Y. C. Ho, and S. K. Narendra. Controllability of linear dynamical systems. *Contributions to Differential Equations*, 1(2):189–213, 1963.

[63] F. L. Cortesi, T. H. Summers, and J. Lygeros. Submodularity of energy related controllability metrics. In *IEEE Conf. on Decision and Control*, pages 2883–2888, Los Angeles, CA, USA, December 2014.

[64] C. Paige. Properties of numerical algorithms related to computing controllability. *IEEE Transactions on Automatic Control*, 26(1):130–138, 1981.

[65] R. Eising. The distance between a system and the set of uncontrollable systems. *memo COSOR*, pages 82–19, 1982.

[66] M. Karow and D. Kressner. On the structured distance to uncontrollability. *Systems & Control Letters*, 58(2):128–132, 2009.

[67] G. Bianchin, P. Frasca, A. Gasparri, and F. Pasqualetti. The observability radius of networks. *IEEE Transactions on Automatic Control*, 62(6):3006–3013, 2017.

[68] D. Boley and W.-S. Lu. Measuring how far a controllable system is from an uncontrollable one. *IEEE Transactions on Automatic Control*, 31(3):249–251, 1986.

[69] D. Kressner and M. Voigt. Distance problems for linear dynamical systems. In *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*, pages 559–583. Springer, 2015.

[70] D. G. Luenberger. *Linear and Nonlinear Programming.* Addison-Wesley, 2 edition, 1984.

[71] J. L. Lagrange. *Analytical Mechanics.* Number 191 in Boston Studies in the Philosophy of Science. Kluwer Academic Publishers, 1997. Translation of the 1811 edition in French by A. Boissonnade and V. N. Vagliente.

[72] D. P. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, 1995.

[73] D. D. Olesky, M. Tsatsomeros, and P. Van Den Driessche. Qualitative controllability and uncontrollability by a single entry. *Linear algebra and its applications*, 187:183–194, 1993.

[74] J. D. Medaglia, S. Gu, F. Pasqualetti, R. L. Ashare, C. Lerman, J. Kable, and D. S. Bassett. Cognitive control in the controllable connectome. *Journal of Neuroscience*, 2017. Submitted.

[75] J. Kim, J. M. Soffer, A. E. Kahn, J. M. Vettel, F. Pasqualetti, and D. S. Bassett. Role of graph architecture in controlling dynamical networks with applications to neural systems. *Nature Physics*, 14:91–98, 2018.

[76] A. Fornito, A. Zalesky, and E. Bullmore. *Fundamentals of Brain Network Analysis.* Academic Press, 2016.

[77] T. Menara, G. Lisi, F. Pasqualetti, and A. Cortese. Brain network dynamics fingerprints are resilient to data heterogeneity. *Journal of Neural Engineering*, 18(2):026004, 2021.

[78] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman. A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(2):676–682, 2001.

[79] J. S. Damoiseaux, S. A. R. B. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith, and C. F. Beckmann. Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences*, 103(37):13848–13853, 2006.

[80] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun. Tracking Whole-Brain Connectivity Dynamics in the Resting State. *Cerebral Cortex*, 24(3):663–676, 11 2012.

[81] B. J. He. Scale-free properties of the functional magnetic resonance imaging signal during rest and task. *Journal of Neuroscience*, 31(39):13786–13795, 2011.

[82] A. Zalesky, A. Fornito, L. Cocchi, L. L. Gollo, and M. Breakspear. Time-resolved resting-state brain networks. *Proceedings of the National Academy of Sciences*, 111(28):10341–10346, 2014.

[83] S. M. Smith, C. F. Beckmann, J. Andersson, E. J. Auerbach, J. Bijsterbosch, G. Douaud, E. Duff, D. A. Feinberg, L. Griffanti, M. P. Harms, Mi. Kelly, T. Laumann, K. L. Miller, S. Moeller, S. Petersen, J. Power, G. Salimi-Khorshidi, A. Z. Snyder, A. T. Vu, M. W. Woolrich, J. Xu, E. Yacoub, K. Uğurbil, D. C. Van Essen, and M. F. Glasser. Resting-state fMRI in the Human Connectome Project. *NeuroImage*, 80:144 – 168, 2013. Mapping the Connectome.

[84] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub, and K. Ugurbil. The WU-Minn Human Connectome Project: An overview. *NeuroImage*, 80:62 – 79, 2013.

[85] C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, B. Liu, P. Matthews, G. Ong, J. Pell, A. Silman, A. Young, T. Sprosen, T. Peakman, and R. Collins. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Medicine*, 12(3):1–10, 03 2015.

[86] C. R. Jack Jr., M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward, A. M. Dale, J. P. Felmlee, J. L. Gunter, D. L. G. Hill, R. Killiany, N. Schuff, S. Fox-Bosetti, C. Lin, C. Studholme, C. S. DeCarli, Gunnar K., H. A. Ward, G. J. Metzger, K. T. Scott, R. Mallozzi, D. Blezek, J. Levy, J. P. Debbins, A. S. Fleisher, M. Albert, R. Green, G. Bartzokis, G. Glover, J. Mugler, and M. W. Weiner. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging*, 27(4):685–691, 2008.

[87] N. Yahata, J. Morimoto, R. Hashimoto, G. Lisi, K. Shibata, Y. Kawakubo, H. Kuwabara, M. Kuroda, T. Yamada, F. Megumi, H. Imamizu, J. Náñez Sr, H. Takahashi, Y. Okamoto, K. Kasai, N. Kato, Y. Sasaki, T. Watanabe, and M. Kawato. A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nature Communications*, 7(1):11254, 2016.

[88] A. T. Drysdale, L. Grosenick, J. Downar, K. Dunlop, F. Mansouri, Y. Meng, R. N. Fetcho, B. Zebley, D. J. Oathes, A. Etkin, A. F. Schatzberg, K. Sudheimer, J. Keller, H. S. Mayberg, F. M. Gunning, G. S. Alexopoulos, M. D. Fox, A. Pascual-Leone, H. U. Voss, B. J. Casey, M. J. Dubin, and C. Liston. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, 23(1):28–38, 2017.

[89] A. Abraham, M. P. Milham, A. Di Martino, R. C. Craddock, D. Samaras, B. Thirion, and G. Varoquaux. Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. *NeuroImage*, 147:736 – 745, 2017.

[90] G. Varoquaux and R. C. Craddock. Learning and comparing functional connectomes across subjects. *NeuroImage*, 80:405 – 415, 2013. Mapping the Connectome.

[91] R. A. Poldrack, K. Whitaker, and D. Kennedy. Introduction to the special issue on reproducibility in neuroimaging. *NeuroImage*, page 116357, 2019.

[92] S. Bari, E. Amico, N. Vike, T. M. Talavage, and J. Goñi. Uncovering multi-site identifiability based on resting-state functional connectomes. *NeuroImage*, 202:115967, 2019.

[93] A. Badhwar, Y. Collin-Verreault, P. Orban, S. Urchs, I. Chouinard, J. Vogel, O. Potvin, S. Duchesne, and P. Bellec. Multivariate consistency of resting-state fmri connectivity maps acquired on a single individual over 2.5 years, 13 sites and 3 vendors. *NeuroImage*, 205:116210, 2020.

[94] R. A. Feis, S. M. Smith, N. Filippini, G. Douaud, E. G. P. Dopper, V. Heise, A. J. Trachtenberg, J. C. van Swieten, M. A. van Buchem, S. A. R. B. Rombouts, and C. E. Mackay. ICA-based artifact removal diminishes scan site differences in multi-center resting-state fMRI. *Frontiers in Neuroscience*, 9:395, 2015.

[95] K. Jann, D. G. Gee, E. Kilroy, S. Schwab, R. X. Smith, T. D. Cannon, and D. J. J. Wang. Functional connectivity in BOLD and CBF data: Similarity and reliability of resting brain networks. *NeuroImage*, 106:111 – 122, 2015.

[96] S. Noble, D. Scheinost, E. S. Finn, X. Shen, X. Papademetris, S. C. McEwen, C. E. Bearden, J. Addington, B. Goodyear, K. S. Cadenhead, H. Mirzakhanian, B. A. Cornblatt, D. M. Olvet,

D. H. Mathalon, T. H. McGlashan, D. O. Perkins, A. Belger, L. J. Seidman, H. Thermenos, M. T. Tsuang, T. G. M. van Erp, E. F. Walker, S. Hamann, S. W. Woods, T. D. Cannon, and R. T. Constable. Multisite reliability of MR-based functional connectivity. *NeuroImage*, 146:959 – 970, 2017.

[97] C. Hawco, J. D. Viviano, S. Chavez, E. W. Dickie, N. Calarco, P. Kochunov, M. Argyelan, J. A. Turner, A. K. Malhotra, R. W. Buchanan, and A. N. Voineskos. A longitudinal human phantom reliability study of multi-center t1-weighted, dti, and resting state fmri data. *Psychiatry Research: Neuroimaging*, 282:134 – 142, 2018.

[98] J. P. Fortin, N. Cullen, Y. I. Sheline, W. D. Taylor, I. Aselcioglu, P. A. Cook, P. Adams, C. Cooper, M. Fava, P. J. McGrath, M. McInnis, M. L. Phillips, M. H. Trivedi, M. M. Weissman, and R. T. Shinohara. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, 167:104 – 120, 2018.

[99] A. Yamashita, N. Yahata, T. Itahashi, G. Lisi, T. Yamada, N. Ichikawa, M. Takamura, Y. Yoshihara, A. Kunimatsu, N. Okada, H. Yamagata, K. Matsuo, R. Hashimoto, G. Okada, Y. Sakai, J. Morimoto, J. Narumoto, Y. Shimada, K. Kasai, N. Kato, H. Takahashi, Y. Okamoto, S. C. Tanaka, M. Kawato, O. Yamashita, and H. Imamizu. Harmonization of resting-state functional MRI data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. *PLOS Biology*, 17(4):1–34, 04 2019.

[100] C. G. Yan, R. C. Craddock, X. N. Zuo, Y. F. Zang, and M. P. Milham. Standardizing the intrinsic brain: Towards robust measurement of inter-individual variation in 1000 functional connectomes. *NeuroImage*, 80:246 – 262, 2013. Mapping the Connectome.

[101] K. J. Friston. Transients, metastability, and neuronal dynamics. *NeuroImage*, 5(2):164 – 171, 1997.

[102] P. Faure and H. Korn. Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie*, 324(9):773 – 793, 2001.

[103] H. Markram. The Blue Brain Project. *Nature Reviews Neuroscience*, 7(2):153, 2006.

[104] K. Bansal, J. D. Medaglia, D. S. Bassett, J. M. Vettel, and S. F. Muldoon. Data-driven brain network models differentiate variability across language tasks. *PLOS Computational Biology*, 14(10):1–25, 10 2018.

[105] D. Vidaurre, L. T. Hunt, A. J. Quinn, B. A. E. 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.

[106] M. F. Singh, T. S. Braver, M. W. Cole, and S. Ching. Estimation and validation of individualized dynamic brain models with resting state fMRI. *NeuroImage*, page 117046, 2020.

[107] G. Varoquaux and B. Thirion. How machine learning is shaping cognitive neuroimaging. *GigaScience*, 3(1):28, 2014.

[108] 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.

[109] T. Kaufmann, D. van der Meer, N. T. Doan, E. Schwarz, M. J. Lund, I. Agartz, D. Alnæs, D. M. Barch, R. Baur-Streubel, A. Bertolino, F. Bettella, M. K. Beyer, E. Bøen, S. Borgwardt, C. L. Brandt, J. Buitelaar, E. G. Celius, S. Cervenka, A. Conzelmann, A. Córdova-Palomera,

and et al. Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature Neuroscience*, 22(10):1617–1623, 2019.

[110] 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.

[111] F Lotte, L Bougrain, A Cichocki, M Clerc, M Congedo, A Rakotomamonjy, and F Yger. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3):031005, apr 2018.

[112] G. Lisi, D. Rivela, A. Takai, and J. Morimoto. Markov switching model for quick detection of event related desynchronization in EEG. *Frontiers in Neuroscience*, 12:24, 2018.

[113] T. J. Mitchell, C. D. Hacker, J. D. Breshears, N. P. Szrama, M. Sharma, D. T. Bundy, M. Pahwa, M. Corbetta, A. Z. Snyder, J. S. Shimony, and E. C. Leuthardt. A Novel Data-Driven Approach to Preoperative Mapping of Functional Cortex Using Resting-State Functional Magnetic Resonance Imaging. *Neurosurgery*, 73(6):969–983, 09 2013.

[114] J. Stiso, A. N. Khambhati, T. Menara, A. E. Kahn, J. M. Stein, S. R. Das, R. Gorniak, J. Tracy, B. Litt, K. A. Davis, F. Pasqualetti, T. H. Lucas, and D. S. Bassett. White matter network architecture guides direct electrical stimulation through optimal state transitions. *Cell Reports*, 28(10):2554 – 2566.e7, 2019.

[115] M. Charquero-Ballester, B. Kleim, D. Vidaurre, C. Ruff, E. Stark, J. J. Tuulari, H. McManners, Y. Bar-Haim, L. Bouquillon, A. Moseley, S. C. R. Williams, M. Woolrich, M. L. Kringelbach, and A. Ehlers. Effective psychological treatment for ptsd changes the dynamics of specific large-scale brain networks. *bioRxiv*, 2020.

[116] C. Gratton, T. O. Laumann, A. N. Nielsen, D. J. Greene, E. M. Gordon, A. W. Gilmore, S. M. Nelson, R. S. Coalson, A. Z. Snyder, B. L. Schlaggar, N. U. F. Dosenbach, and S. E. Petersen. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron*, 98(2):439–452, 2018.

[117] S. M. Smith, D. Vidaurre, C. F. Beckmann, M. F. Glasser, M. Jenkinson, K. L. Miller, T. E. Nichols, E. C. Robinson, G. Salimi-Khorshidi, M. W. Woolrich, D. M. Barch, K. Uğurbil, and D. C. Van Essen. Functional connectomics from resting-state fmri. *Trends in Cognitive Sciences*, 17(12):666 – 682, 2013. Special Issue: The Connectome.

[118] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, and M. Jenkinson. The minimal preprocessing pipelines for the human connectome project. *NeuroImage*, 80:105–124, 2013.

[119] L. Griffanti, G. Salimi-Khorshidi, C. F. Beckmann, E. J. Auerbach, G. Douaud, C. E. Sexton, E. Zsoldos, K. P. Ebmeier, N. Filippini, C. E. Mackay, S. Moeller, J. Xu, E. Yacoub, G. Baselli, K. Ugurbil, K. L. Miller, and S. M. Smith. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*, 95:232 – 247, 2014.

[120] R. Ciric, D. H. Wolf, J. D. Power, D. R. Roalf, G. L. Baum, K. Ruparel, R. T. Shinohara, M. A. Elliott, S. B. Eickhoff, C. Davatzikos, R. C. Gur, R. E. Gur, D. S. Bassett, and T. D. Satterthwaite. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *NeuroImage*, 154:174 – 187, 2017.

[121] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3):2142 – 2154, 2012.

[122] C. G. Yan, X. D. Wang, X. N. Zuo, and Y. F. Zang. DPABI: Data processing & analysis for (resting-state) brain imaging. *Neuroinformatics*, 14(3):339–351, Jul 2016.

[123] 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.

[124] K. Friston, J. Mattout, N. Trujillo-Barreto, J. Ashburner, and W. Penny. Variational free energy and the Laplace approximation. *NeuroImage*, 34(1):220 – 234, 2007.

[125] I. Rezek and S. Roberts. *Ensemble Hidden Markov Models with Extended Observation Densities for Biosignal Analysis*, pages 419–450. Springer London, London, 2005.

[126] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct):2825–2830, 2011.

[127] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560, 1997.

[128] 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.

[129] D. Yekutieli and Y. Benjamini. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference*, 82(1):171 – 196, 1999.

[130] R. Halır and J. Flusser. Numerically stable direct least squares fitting of ellipses. In *Proc. 6th International Conference in Central Europe on Computer Graphics and Visualization. WSCG*, volume 98, pages 125–132. Citeseer, 1998.

[131] C. E. Curtis and M. D'Esposito. Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7(9):415 – 423, 2003.

[132] E. M. Gordon, T. O. Laumann, B. Adeyemo, and S. E. Petersen. Individual Variability of the System-Level Organization of the Human Brain. *Cerebral Cortex*, 27(1):386–399, 10 2015.

[133] T. Chiba, K. Ide, J. E. Taylor, S. Boku, H. Toda, T. Kanazawa, S. Kato, Y. Horiuchi, A. Hishimoto, T. Maruyama, T. Yamamoto, M. Shirakawa, I. Sora, M. Kawato, and A. Koizumi. A reciprocal inhibition model of alternations between under-/overemotional modulatory states in patients with PTSD. *Molecular Psychiatry*, pages 1–17, 2020.

[134] D. J. Lurie, D. Kessler, D. S. Bassett, R. F. Betzel, M. Breakspear, S. Keilholz, A. Kucyi, R. Liégeois, M. A. Lindquist, A. R. McIntosh, R. A. Poldrack, J. M. Shine, W. H. Thompson, N. Z. Bielczyk, L. Douw, D. Kraft, R. L. Miller, M. Muthuraman, L. Pasquini, A. Razi, D. Vidaurre, H. Xie, and V. D. Calhoun. Questions and controversies in the study of time-varying functional connectivity in resting fMRI. *Network Neuroscience*, 2019. In Press.

[135] 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.

[136] M. Bonomi, C. Camilloni, A. Cavalli, and M. Vendruscolo. Metainference: A Bayesian inference method for heterogeneous systems. *Science Advances*, 2(1), 2016.

[137] D. Bzdok and B. T. T. Yeo. Inference in the age of big data: Future perspectives on neuroscience. *NeuroImage*, 155:549 – 564, 2017.

[138] A. Cortese, K. Amano, A. Koizumi, M. Kawato, and H. Lau. Multivoxel neurofeedback selectively modulates confidence without changing perceptual performance. *Nature Communications*, 7(1):1–18, 2016.

[139] V. Taschereau-Dumouchel, A. Cortese, H. Lau, and M. Kawato. Conducting decoded neurofeedback studies. *Social Cognitive and Affective Neuroscience*, May 2020.

[140] V. Taschereau-Dumouchel, A. Cortese, T. Chiba, J. D. Knotts, M. Kawato, and H. Lau. Towards an unconscious neural reinforcement intervention for common fears. *Proceedings of the National Academy of Sciences*, 115(13):3470–3475, 2018.

[141] L.E. Stoeckel, K.A. Garrison, S.S. Ghosh, P. Wighton, C.A. Hanlon, J.M. Gilman, S. Greer, N.B. Turk-Browne, M.T. deBettencourt, D. Scheinost, C. Craddock, T. Thompson, V. Calderon, C.C. Bauer, M. George, H.C. Breiter, S. Whitfield-Gabrieli, J.D. Gabrieli, S.M. LaConte, L. Hirshberg, J.A. Brewer, M. Hampson, A. Van Der Kouwe, S. Mackey, and A.E. Evins. Optimizing real time fMRI neurofeedback for therapeutic discovery and development. *NeuroImage: Clinical*, 5:245 – 255, 2014.

[142] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. Stability conditions for cluster synchronization in networks of heterogeneous Kuramoto oscillators. *IEEE Transactions on Control of Network Systems*, 7(1):302 – 314, 2020.

[143] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. Exact and approximate stability conditions for cluster synchronization of Kuramoto oscillators. In *American Control Conference*, pages 205 – 210, Philadelphia, PA, USA, July 2019.

[144] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.

[145] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press, 2003.

[146] F. L. Lewis, H. Zhang, K. Hengster-Movric, and A. Das. Introduction to synchronization in nature and physics and cooperative control for multi-agent systems on graphs. In *Cooperative Control of Multi-Agent Systems*, pages 1–21. Springer, 2014.

[147] J. Cabral, E. Hugues, O. Sporns, and G. Deco. Role of local network oscillations in resting-state functional connectivity. *NeuroImage*, 57(1):130–139, 2011.

[148] A. Moiseff and J. Copeland. A new type of synchronized flashing in a north american firefly. *Journal of Insect Behavior*, 13(4):597–612, 2000.

[149] I. Giardina. Collective behavior in animal groups: theoretical models and empirical studies. *HFSP Journal*, 2(4):205–219, 2008.

[150] O. Simeone, U. Spagnolini, Y. Bar-Ness, and S. H. Strogatz. Distributed synchronization in wireless networks. *IEEE Signal Processing Magazine*, 25(5):81–97, 2008.

[151] F. Dörfler, M. Chertkov, and F. Bullo. Synchronization in complex oscillator networks and smart grids. *Proceedings of the National Academy of Sciences*, 110(6):2005–2010, 2013.

[152] F. Sorrentino and E. Ott. Network synchronization of groups. *Physical Review E*, 76(5):056114, 2007.

[153] 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.

[154] N. Chopra and M. W. Spong. On exponential synchronization of Kuramoto oscillators. *IEEE Transactions on Automatic Control*, 54(2):353–357, 2009.

[155] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Physics Reports*, 469(3):93–153, 2008.

[156] K. Kaneko. Relevance of dynamic clustering to biological networks. *Physica D: Nonlinear Phenomena*, 75(1-3):55–73, 1994.

[157] L. Stone, R. Olinky, B. Blasius, A. Huppert, and B. Cazelles. Complex synchronization phenomena in ecological systems. In *AIP Conference Proceedings*, volume 622, pages 476–488. AIP, 2002.

[158] I. Belykh and M. Hasler. Mesoscale and clusters of synchrony in networks of bursting neurons. *Chaos*, 21(1):016106, 2011.

[159] J. R. Terry, K. S. Thornburg Jr, D. J. DeShazer, G. D. VanWiggeren, S. Zhu, P. Ashwin, and R. Roy. Synchronization of chaos in an array of three lasers. *Physical Review E*, 59(4):4036, 1999.

[160] A. Schnitzler and J. Gross. Normal and pathological oscillatory communication in the brain. *Nature Reviews Neuroscience*, 6(4):285, 2005.

[161] K. Lehnertz, S. Bialonski, M.-T. Horstmann, D. Krug, A. Rothkegel, M. Staniek, and T. Wagner. Synchronization phenomena in human epileptic brain networks. *Journal of Neuroscience Methods*, 183(1):42–48, 2009.

[162] C. Hammond, H. Bergman, and P. Brown. Pathological synchronization in Parkinson's disease: networks, models and treatments. *Trends in Neurosciences*, 30(7):357–364, 2007.

[163] Y. Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, pages 420–422, Berlin, Heidelberg, 1975.

[164] A. Daffertshofer and B. van Wijk. On the influence of amplitude on the connectivity between phases. *Frontiers in Neuroinformatics*, 5:6, 2011.

[165] G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, and R. Kötter. Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences*, 106(25):10302–10307, 2009.

[166] F. Váša, M. Shanahan, P. J. Hellyer, G. Scott, J. Cabral, and R. Leech. Effects of lesions on synchrony and metastability in cortical networks. *Neuroimage*, 118:456–467, 2015.

[167] V. N. Belykh, I. Belykh, and M. Hasler. Hierarchy and stability of partially synchronous oscillations of diffusively coupled dynamical systems. *Physical Review E*, 62(5):6332, 2000.

[168] A. Y. Pogromsky, G. Santoboni, and H. Nijmeijer. Partial synchronization: from symmetry towards stability. *Physica D: Nonlinear Phenomena*, 172(1-4):65–87, 2002.

[169] I. Belykh, V. N. Belykh, K. Nevidin, and M. Hasler. Persistent clusters in lattices of coupled nonidentical chaotic systems. *Chaos*, 13(1):165–178, 2003.

[170] I. Stewart, M. Golubitsky, and M. Pivato. Symmetry groupoids and patterns of synchrony in coupled cell networks. *SIAM Journal on Applied Dynamical Systems*, 2(4):609–646, 2003.

[171] A. Y. Pogromsky. A partial synchronization theorem. *Chaos*, 18(3):037107, 2008.

[172] D. Fiore, G. Russo, and M. di Bernardo. Exploiting nodes symmetries to control synchronization and consensus patterns in multiagent systems. *Control Systems Letters*, 1(2):364–369, 2017.

[173] L. M. Pecora, F. Sorrentino, A. M. Hagerstrom, T. E. Murphy, and R. Roy. Cluster synchronization and isolated desynchronization in complex networks with symmetries. *Nature Communications*, 5(1):1–8, 2014.

[174] F. Sorrentino, L. M. Pecora, A. M. Hagerstrom, T. E. Murphy, and R. Roy. Complete characterization of the stability of cluster synchronization in complex dynamical networks. *Science Advances*, 2(4), 2016.

[175] L. M. Pecora and T. L. Carroll. Master stability functions for synchronized coupled systems. *Physical Review Letters*, 80(10):2109, 1998.

[176] G. Russo and J.-J E. Slotine. Symmetries, stability, and control in nonlinear systems and networks. *Physical Review E*, 84(4):041929, 2011.

[177] Q. C. Pham and J.-J. Slotine. Stable concurrent synchronization in dynamic system networks. *Neural Networks*, 20(1):62–77, 2007.

[178] Z. Aminzare, B. Dey, E. N. Davison, and N. E. Leonard. Cluster synchronization of diffusively-coupled nonlinear systems: A contraction based approach. *Journal of Nonlinear Science*, pages 1–23, 2018.

[179] W. Wu, W. Zhou, and T. Chen. Cluster synchronization of linearly coupled complex networks under pinning control. *IEEE Transactions on Circuits and Systems*, 56(4):829–839, 2009.

[180] W. Lu, B. Liu, and T. Chen. Cluster synchronization in networks of coupled nonidentical dynamical systems. *Chaos*, 20(1):013120, 2010.

[181] C. Favaretto, A. Cenedese, and F. Pasqualetti. Cluster synchronization in networks of Kuramoto oscillators. In *IFAC World Congress*, pages 2433–2438, Toulouse, France, July 2017.

[182] Y. Qin, Y. Kawano, and M. Cao. Partial phase cohesiveness in networks of communitinized Kuramoto oscillators. In *European Control Conference*, pages 2028–2033, Limassol, Cyprus, 2018.

[183] M. T. Schaub, N. O'Clery, Y. N. Billeh, J.-C. Delvenne, R. Lambiotte, and M. Barahona. Graph partitions and cluster synchronization in networks of oscillators. *Chaos*, 26(9):094821, 2016.

[184] L. Tiberi, C. Favaretto, M. Innocenti, D. S. Bassett, and F. Pasqualetti. Synchronization patterns in networks of Kuramoto oscillators: A geometric approach for analysis and control. In *IEEE Conf. on Decision and Control*, pages 481–486, Melbourne, Australia, December 2017.

[185] I. V. Belykh, B. N. Brister, and V. N. Belykh. Bistability of patterns of synchrony in Kuramoto oscillators with inertia. *Chaos*, 26(9):094822, 2016.

[186] Y. S. Cho, T. Nishikawa, and A. E. Motter. Stable chimeras and independently synchronizable clusters. *Physical Review Letters*, 119(8):084101, 2017.

[187] A. Jadbabaie, N. Motee, and M. Barahona. On the stability of the Kuramoto model of coupled nonlinear oscillators. In *American Control Conference*, pages 4296–4301, Boston, MA, USA, June 2004.

[188] F. Dörfler and F. Bullo. Exploring synchronization in complex oscillator networks. In *IEEE Conf. on Decision and Control*, pages 7157–7170, Maui, Hawaii, USA, 2012. IEEE.

[189] A. N. Michel, L. Hou, and D. Liu. *Stability of Dynamical Systems*. Springer, 2008.

[190] D. Mantini, M. G. Perrucci, C. Del Gratta, G. L. Romani, and M. Corbetta. Electrophysiological signatures of resting state networks in the human brain. *Proceedings of the National Academy of Sciences*, 104(32):13170–13175, 2007.

[191] Z. Ma, Z. Liu, and G. Zhang. A new method to realize cluster synchronization in connected chaotic networks. *Chaos*, 16(2):023103, 2006.

[192] V. N. Belykh, G. V. Osipov, V. S. Petrov, J. A. K. Suykens, and J. Vandewalle. Cluster synchronization in oscillatory networks. *Chaos*, 18(3), 2008.

[193] A. B. Siddique, L. Pecora, J. D. Hart, and F. Sorrentino. Symmetry-and input-cluster synchronization in networks. *Physical Review E*, 97(4):042217, 2018.

[194] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 3 edition, 2002.

[195] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.

[196] J. Qin, Q. Ma, H. Gao, Y. Shi, and Y. Kang. On group synchronization for interacting clusters of heterogeneous systems. *IEEE Transactions on Cybernetics*, 47(12):4122–4133, 2017.

[197] W. J. Rugh. *Linear System Theory*. Information and System Sciences Series. Prentice Hall, New Jersey, 1993.

[198] C. Favaretto, D. S. Bassett, A. Cenedese, and F. Pasqualetti. Bode meets Kuramoto: Synchronized clusters in oscillatory networks. In *American Control Conference*, pages 2378–5861, Seattle, WA, May 2017.

[199] S. Dashkovskiy, B. S. Rüffer, and F. R. Wirth. An ISS small gain theorem for general networks. *Mathematics of Control, Signals, and Systems*, 19(2):93–122, 2007.

[200] K. K. Stevens. On the response of linear systems to modulated sinusoidal inputs. *J. Sound Vib.*, 21(3):295–306, 1972.

[201] T. Menara, Y. Qin, D. S. Bassett, and F. Pasqualetti. Relay interactions enable remote synchronization in networks of phase oscillators. *IEEE Control Systems Letters*, 6:500–505, 2022.

[202] F. Varela, J. P. Lachaux, E. Rodriguez, and J. Martinerie. The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2(4):229–239, 2001.

[203] J. Fell and N. Axmacher. The role of phase synchronization in memory processes. *Nature Reviews Neuroscience*, 12(2):105–118, 2011.

[204] J. F. Hipp, D. J. Hawellek, M. Corbetta, M. Siegel, and A. K. Engel. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience*, 15(6):884–890, 2012.

[205] R. Vicente, L. L. Gollo, C. R. Mirasso, I. Fischer, and G. Pipa. Dynamical relaying can yield zero time lag neuronal synchrony despite long conduction delays. *Proceedings of the National Academy of Sciences*, 105(44):17157–17162, 2008.

[206] L. L. Gollo, C. R. Mirasso, and A. E. P. Villa. Dynamic control for synchronization of separated cortical areas through thalamic relay. *NeuroImage*, 52(3):947–955, 2010.

[207] G. Arnulfo, S. H. Wang, V. Myrov, B. Toselli, J. Hirvonen, M. M. Fato, L. Nobili, F. Cardinale, A. Rubino, A. Zhigalov, S. Palva, and J. M. Palva. Long-range phase synchronization of high-frequency oscillations in human cortex. *Nature Communications*, 11(1):5363, 2020.

[208] A. Bergner, M. Frasca, G. Sciuto, A. Buscarino, E. J. Ngamga, L. Fortuna, and J. Kurths. Remote synchronization in star networks. *Physical Review E*, 85(2):026208, 2012.

[209] L. V. Gambuzza, A. Cardillo, A. Fiasconaro, L. Fortuna, J. Gómez-Gardenes, and M. Frasca. Analysis of remote synchronization in complex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(4):043103, 2013.

[210] J. Emenheiser, A. Salova, J. Snyder, J. P. Crutchfield, and R. M. D'Souza. Network and phase symmetries reveal that amplitude dynamics stabilize decoupled oscillator clusters. *arXiv preprint arXiv:2010.09131*, 2020.

[211] A. Ponce-Alvarez, G. Deco, P. Hagmann, G. L. Romani, D. Mantini, and M. Corbetta. Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity. *PLoS Computational Biology*, 11(2):1–23, 02 2015.

[212] D. Cumin and C. P. Unsworth. Generalising the Kuramoto model for the study of neuronal synchronisation in the brain. *Physica D: Nonlinear Phenomena*, 226(2):181–196, 2007.

[213] A. Politi and M. Rosenblum. Equivalence of phase-oscillator and integrate-and-fire models. *Physical Review E*, 91(4):042916, 2015.

[214] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. A framework to control functional connectivity in the human brain. In *IEEE Conf. on Decision and Control*, pages 4697–4704, Nice, France, December 2019.

[215] Y. Qin, M. Cao, B. D. O. Anderson, D. S. Bassett, and F. Pasqualetti. Mediated remote synchronization: the number of mediators matters. *IEEE Control Systems Letters*, 5(3):767–772, 2020.

[216] D. Maraun and J. Kurths. Epochs of phase coherence between El Niño/Southern Oscillation and Indian monsoon. *Geophysical Research Letters*, 32(15):L15709, 2005.

[217] L. Zhang, A. E. Motter, and T. Nishikawa. Incoherence-mediated remote synchronization. *Physical Review Letters*, 118:174102, 2017.

[218] Y. Qin, Y. Kawano, and M. Cao. Stability of remote synchronization in star networks of Kuramoto oscillators. In *IEEE Conf. on Decision and Control*, pages 5209–5214, Miami Beach, FL, USA, Dec 2018.

[219] V. Nicosia, M. Valencia, M. Chavez, A. Díaz-Guilera, and V. Latora. Remote synchronization reveals network symmetries and functional modules. *Phys. Rev. Lett.*, 110:174102, Apr 2013.

[220] V. Vuksanović and P. Hövel. Functional connectivity of distant cortical regions: role of remote synchronization and symmetry in interactions. *NeuroImage*, 97:1–8, 2014.

[221] L. V. Gambuzza, M. Frasca, L. Fortuna, and S. Boccaletti. Inhomogeneity induces relay synchronization in complex networks. *Physical Review E*, 93:042203, Apr 2016.

[222] S. Jafarpour and F. Bullo. Synchronization of Kuramoto oscillators via cutset projections. *IEEE Transactions on Automatic Control*, 64(7):2830–2844, 2018.

[223] K. Nooner et al. The NKI-Rockland sample: A model for accelerating the pace of discovery science in psychiatry. *Frontiers in Neuroscience*, 6:152, 2012.

[224] W. J. Rugh. *Linear System Theory*. Information and System Sciences Series. Prentice Hall, New Jersey, 1993.

[225] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. Conditions for feedback linearization of network systems. *IEEE Control Systems Letters*, 4(3):578–583, 2020.

[226] A. Isidori. *Nonlinear Control Systems*. Communications and Control Engineering Series. Springer, 3 edition, 1995.

[227] A. J. Krener. On the equivalence of control systems and the linearization of nonlinear systems. *SIAM Journal on Control*, 11(4):670–676, 1973.

[228] C. O. Aguilar and B. Gharesifard. Necessary conditions for controllability of nonlinear networked control systems. In *American Control Conference*, pages 5379–5383, June 2014.

[229] R. F. Betzel and D. S. Bassett. Multi-scale brain networks. *NeuroImage*, 160:73 – 83, 2017.

[230] G. Conte, C. H. Moog, and A. M. Perdon. *Algebraic methods for nonlinear control systems*. Springer, 2007.

[231] I. Mir, H. Taha, S. A. Eisa, and A. Maqsood. A controllability perspective of dynamic soaring. *Nonlinear Dynamics*, 94(4):2347–2362, 2018.

[232] I. Klickstein, A. Shirin, and F. Sorrentino. Locally optimal control of complex networks. *Physical Review Letters*, 119:268301, Dec 2017.

[233] M. T. Angulo, A. Aparicio, and C. H. Moog. Structural accessibility and structural observability of nonlinear networked systems. *IEEE Transactions on Network Science and Engineering*, 2019. In Press.

[234] Y. Kawano and M. Cao. Structural accessibility and its applications to complex networks governed by nonlinear balance equations. *IEEE Transactions on Automatic Control*, 64(11):4607–4614, Nov 2019.

[235] E. D. Sontag. Controllability is harder to decide than accessibility. *SIAM Journal on Control and Optimization*, 26(5):1106–1118, 1988.

[236] R. W. Brockett. Feedback invariants for nonlinear systems. In *Int. Congress of Mathematicians*, Helsinki, Finland, 1978.

[237] B. Jakubczyk and W. Respondek. On linearization of control systems. *Bulletin de l' Académie Polonaise des Sciences*, 28(9-10):517–522, 1980.

[238] B. d'Andréa Novel, G. Campion, and G. Bastin. Control of nonholonomic wheeled mobile robots by state feedback linearization. *The International Journal of Robotics Research*, 14(6):543–559, 1995.

[239] E. Barany, S. Schaffer, K. Wedeward, and S. Ball. Nonlinear controllability of singularly perturbed models of power flow networks. In *IEEE Conf. on Decision and Control*, volume 5, pages 4826–4832, Atlantis, Paradise Island, Bahamas, Dec 2004.

[240] P. Tabuada, W. Ma, J. Grizzle, and A. D. Ames. Data-driven control for feedback linearizable single-input systems. In *IEEE Conf. on Decision and Control*, pages 6265–6270, Melbourne, Australia, December 2017.

[241] T. Westenbroek, D. Fridovich-Keil, E. Mazumdar, S. Arora, V. Prabhu, S. S. Sastry, and C. J. Tomlin. Feedback linearization for unknown systems via reinforcement learning. *arXiv preprint arXiv:1910.13272*, 2019.

[242] B. Charlet, J. Lévine, and R. Marino. Sufficient conditions for dynamic state feedback linearization. *SIAM Journal on Control and Optimization*, 29(1):38–57, 1991.

[243] L. A. Marquez-Martinez and C. H. Moog. Input-output feedback linearization of time-delay systems. *IEEE Transactions on Automatic Control*, 49(5):781–785, May 2004.

[244] X. Liu and T. Chen. Synchronization of nonlinear coupled networks via aperiodically intermittent pinning control. *IEEE Transactions on Neural Networks and Learning Systems*, 26(1):113–126, 2015.

[245] P. S. Skardal and A. Arenas. Control of coupled oscillator networks with application to microgrid technologies. *Science Advances*, 1(7):e1500339, 2015.

[246] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási. Observability of complex systems. *Proceedings of the National Academy of Sciences*, 110(7):2460–2465, 2013.

[247] H. Nijmeijer and A. Van der Schaft. *Nonlinear Dynamical Control Systems*, volume 175. Springer, 1990.

[248] A. E. Sizemore, C. Giusti, A. E. Kahn, J. M. Vettel, R. F. Betzel, and D. S. Bassett. Cliques and cavities in the human connectome. *Journal of Computational Neuroscience*, 44(1):115–145, Feb 2018.

[249] P. S. Skardal and A. Arenas. Abrupt desynchronization and extensive multistability in globally coupled oscillator simplexes. *Physical Review Letters*, 122(24):248301, 2019.

[250] J. J. O'Connor, M. J. Rowan, and R. Anwyl. Long-lasting enhancement of NMDA receptor-mediated synaptic transmission by metabotropic glutamate receptor activation. *Nature*, 367(6463):557, 1994.

[251] https://github.com/tommasomenara/feedbacklinearization.

[252] T. Menara, G. Bianchin, M. Innocenti, and F. Pasqualetti. On the number of strongly structurally controllable networks. In *American Control Conference*, pages 340–345, Seattle, WA, USA, 2017.

[253] W. Respondek and I. A. Tall. Nonlinearizable single-input control systems do not admit stationary symmetries. *Systems & Control Letters*, 46(1):1–16, 2002.

[254] A. Zalesky, A. Fornito, and E. Bullmore. On the use of correlation as a measure of network connectivity. *Neuroimage*, 60(4):2096–2106, 2012.

[255] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, 104(24):10240–10245, 2007.

[256] R. Polanía, M. A. Nitsche, and C. C. Ruff. Studying and modifying brain function with non-invasive brain stimulation. *Nature Neuroscience*, 21(2):174–187, 2018.

[257] A. Kirkwood and M. F. Bear. Hebbian synapses in visual cortex. *Journal of Neuroscience*, 14(3):1634–1645, 1994.

[258] A. M. Thomson. Facilitation, augmentation and potentiation at central synapses. *Trends in Neurosciences*, 23(7):305–312, 2000.

[259] P. A. Tass. A model of desynchronizing deep brain stimulation with a demand-controlled coordinated reset of neural subpopulations. *Biological Cybernetics*, 89(2):81–88, 2003.

[260] A. Franci, A. Chaillet, E. Panteley, and F. Lamnabhi-Lagarrigue. Desynchronization and inhibition of Kuramoto oscillators by scalar mean-field feedback. *Mathematics of Control, Signals, and Systems*, 24(1-2):169–217, 2012.

[261] M. Jafarian, X. Yi, M. Pirani, H. Sandberg, and K. H. Johansson. Synchronization of Kuramoto oscillators in a bidirectional frequency-dependent tree network. In *IEEE Conf. on Decision and Control*, pages 4505–4510, Miami Beach, FL, USA, Dec 2018.

[262] E. Nozari and J. Cortés. Oscillations and coupling in interconnections of two-dimensional brain networks. In *American Control Conference*, pages 193–198, Philadelphia, PA, USA, Jul 2019.

[263] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. Stability conditions for cluster synchronization in networks of heretogeneous Kuramoto oscillators. *IEEE Transactions on Control of Network Systems*, 2019. In press.

[264] C. D. Hacker, A. Z. Snyder, M. Pahwa, M. Corbetta, and E. C. Leuthardt. Frequency-specific electrophysiologic correlates of resting state fMRI networks. *NeuroImage*, 149:446–457, 2017.

[265] D. S. Bassett, P. Zurn, and J. I. Gold. On the nature and use of models in network neuroscience. *Nature Reviews Neuroscience*, pages 566–578, 2018.

[266] J. P. Boyle and R. L. Dykstra. A method for finding projections onto the intersection of convex sets in Hilbert spaces. In *Advances in order restricted statistical inference*, pages 28–47. Springer, 1986.

[267] A. M. Tuladhar, L. Snaphaan, E. Shumskaya, M. Rijpkema, G. Fernandez, D. G. Norris, and F.-E. de Leeuw. Default mode network connectivity in stroke patients. *PLoS ONE*, 8(6):e66556, 2013.

[268] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price. Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. *NeuroImage*, 12(4):466–477, 2000.

[269] M. D. Fox, D. Zhang, A. Z. Snyder, and M. E. Raichle. The global signal and observed anticorrelated resting state brain networks. *Journal of Neurophysiology*, 101(6):3270–3283, 2009.

[270] J. R. R. Tolkien. *The Hobbit, or there and back again.* George Allen & Unwin, London; United Kingdom, 1 edition, 1937.

[271] L. Cocchi, L. L. Gollo, A. Zalesky, and M. Breakspear. Criticality in the brain: A synthesis of neurobiology, models and cognition. *Progress in Neurobiology*, 158:132–152, 2017.

[272] P. Hövel, A. Viol, P. Loske, L. Merfort, and V. Vuksanović. Synchronization in functional networks of the human brain. *Journal of Nonlinear Science*, pages 1–24, 2018.

[273] Y. Qin, T. Menara, D. S. Bassett, and F. Pasqualetti. Phase-amplitude coupling in neuronal oscillator networks. *Physical Review Research*, 3(2), June 2021.

[274] E. D. Fagerholm, R. J. Moran, I. R. Violante, R. Leech, and K. J. Friston. Dynamic causal modelling of phase-amplitude interactions. *NeuroImage*, 208:116452, Mar 2020.

[275] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. Functional control of oscillator networks. *arXiv:2012.04217*, 2021. Submitted.