Functional, clustered, feedforward, and mesoscale brain networks

Thesis by

Yazan N. Billeh

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2016

(Defended March 18, 2016)

To my family

Nicola, Lina, Elias, and Omar

Acknowledgements

This PhD would not have been possible without the support of the incredible team of individuals mentioned here.

First, I am extremely thankful to my advisor, Christof Koch, who agreed to supervise me even though he was moving to Seattle at the end of my first year. Instead of simply telling me to join another lab, he decided to mentor and foster my scientific abilities and I am profoundly grateful to him. I am humbled to say that not a single email went unanswered! It is remarkable how Christof was always supportive on an academic and personal level (such as when I needed knee surgery). An amazing mentor. Many thanks Christof.

Second, I am indebted and very thankful for the support and help of Markus Meister. He provided me with much more than desk space by allowing me to feel like a member of his lab at Caltech. I was able to be a part of the intellectual environment he creates by attending and presenting in group meetings and journal clubs and his door was always open to me to ask for suggestions and help. Simply put, I could not have done it without Markus.

A key person who ignited the journey of my PhD is Costas Anastassiou. He was instrumental in my learning by taking me under his wing and showing me the ropes of computational neuroscience and Caltech. He was one of the reasons I came to Caltech and it has been a privilege to work with him. I looked forward to growing this strong relationship we formed.

I was also very fortunate to have collaborated with some brilliant scientists and I would like to acknowledge and thank Michael Schaub and Mauricio Barahona. Michael visited Caltech, from Imperial College London, for one month during my second year and we have been collaborating ever since. It is truly a pleasure to work with Michael and I know there is still more science and discoveries in our future. From Mauricio, I learnt a great deal about how to design, analyze, and present science. Thank you both.

It is with great privilege that I also got to work with many outstanding individuals at the Allen Institute for Brain Science in Seattle, Washington. I am very thankful for all the people I interacted with who shared their time to meet with me and respond to my emails. In particular, Stefan Mihalas had significant feedback and input in many of my projects, for

which I am very grateful.

Moreover, I was lucky to work with exceptional scientists at the University of Wisconsin. It was a great learning experience for me and their support and time spent are highly appreciated. Thank you to Chiara Cirelli, Giulio Tononi, Alex Rodriguez, and the members of the Center for Sleep and Consciousness.

There are many people who must be acknowledged, and although briefly mentioned here they know how important and vital they were throughout my studies. I am grateful to all members of the Koch lab for all the support, adventures, and laughs during my time at Caltech. I also am thankful to the members of the Meister lab. In addition, friends from my time at Pasadena, Michigan, London, and Amman were part of the supportive team, allowing me to write this thesis, and to all of you I am deeply thankful.

Further, I would like to acknowledge and thank my committee for support throughout the years. They helped me realize the potentials of projects, suggested classes that heavily helped my research progress, and did not make it too difficult to find a time for all of us to meet. Thank you to Markus Meister, Thanos Siapas, Richard Murray, and Christof Koch.

Finally and most importantly, I deeply thank my incredible family. They have been part of my team from before I was born, let alone Caltech. They are the superstars of the team! It is really them who made any of this possible. My thanks to them should be longer than this thesis, but I will drastically hold back. I can only dream to be as good a teammate as they are to me and hope to show them how grateful I am throughout my life. From the bottom of my heart, thank you Nicola, Lina, Elias, and Omar!

Abstract

The brain is a network spanning multiple scales from subcellular to macroscopic. In this thesis I present four projects studying brain networks at different levels of abstraction. The first involves determining a functional connectivity network based on neural spike trains and using a graph theoretical method to cluster groups of neurons into putative cell assemblies. In the second project I model neural networks at a microscopic level. Using different clustered wiring schemes, I show that almost identical spatiotemporal activity patterns can be observed, demonstrating that there is a broad neuro-architectural basis to attain structured spatiotemporal dynamics. Remarkably, irrespective of the precise topological mechanism, this behavior can be predicted by examining the spectral properties of the synaptic weight matrix. The third project introduces, via two circuit architectures, a new paradigm for feedforward processing in which inhibitory neurons have the complex and pivotal role in governing information flow in cortical network models. Finally, I analyze axonal projections in sleep deprived mice using data collected as part of the Allen Institute's Mesoscopic Connectivity Atlas. After normalizing for experimental variability, the results indicate there is no single explanatory difference in the mesoscale network between control and sleep deprived mice. Using machine learning techniques, however, animal classification could be done at levels significantly above chance. This reveals that intricate changes in connectivity do occur due to chronic sleep deprivation.

Published Content and Contributions

Billeh Y. N., Schaub M. T.*, Anastassiou C. A., Barahona M., and Koch C. (2014). "Revealing Cell Assemblies at Multiple Levels of Granularity". In: *Journal of Neuroscience Methods*, 236, pp. 92 106. *co-first authoer. DOI: 10.1016/j.jneumeth.2014.08.011

YNB participated in the conception and design of the experiments, performing the experiments, analyzing the data and the writing of the paper.

Schaub M. T., Billeh Y. N.*, Anastassiou C. A., Koch C., and Barahona M. (2015). "Emergence of Slow-Switching Assemblies in Structured Neuronal Networks". In: *PLoS Computational Biology*. e1004196. *co-first author. DOI: 10.1371/journal.pcbi.1004196

YNB participated in the conception and design of the experiments, performing the experiments, analyzing the data and the writing of the paper.

Billeh Y. N. and Schaub M. T. (submitted). "Directed information propagation driven by inhibitory interactions".

YNB participated in the conception and design of the experiments, performing the experiments, analyzing the data and the writing of the paper.

Billeh Y. N., Rodriguez A. V.*, Bellesi1 B., Bernard A., de Vivo L., Funk C. M., Harris J., Honjoh S., Mihalas S., Ng L., Koch C., Cirelli C., and Tononi G. (2016). "Effects of chronic sleep restriction during early adolescence on the adult pattern of connectivity of mouse secondary motor cortex". In: *eNeuro*. *co-first author. DOI: 10.1523/ENEURO.0053-16.2016

YNB participated in analyzing the data and the writing of the paper.

Contents

Chapter 1

Introduction

Brains can contain billions of neurons that are connected via synapses and elicit action potentials. Arguably three of the most fundamental terms in neuroscience were just stated and are thus introduced below.

1.1 Background

1.1.1 What are neurons?

The nervous system of animals, just as all other systems and organs, is composed of organic biological cells. The most studied cells, neurons, can be described analogously to a communication system that has a receiver, processor, and transmitter. Such a description is, of course, a simplification, as neurons show vast diversity with various properties [Ramón y Cajal, 1888, 1933]. For the introduction herein, however, we will view them in this most basic form (see Figure 1.1A).

The dendrites of neurons are analogous to a receiver as they are responsible for receiving information from other neurons that connect to it. Dendrites are tree-like cytoplasmic processes that have numerous architectures depending on the neuronal subtype [Snell, 2006]. The processor of a neuron is the soma (also termed cell body) as it is the center that integrates information from all inputs handled by the dendrites to determine whether or not the neuron itself will send a signal to the neurons it connects to. The soma also receives inputs from other neurons directly and contains the nucleus of the cell [Snell, 2006]. Finally, the transmitter of a neuron is its axon which is a long process that ends in axonal terminals that connect to other neurons. Signals sent along the axons are electrical impulses called action potentials – the currency of communication of the nervous system [Snell, 2006].

1.1.2 What is an action potential?

Since first being reported [Adrian and Zotterman, 1926], neuroscientist around the world now record action potentials daily. The seminal work of Alan Hodgkin and Andrew Huxley provided a leap in our knowledge of action potentials [Hodgkin and Huxley, 1952]. To understand these electrical impulses of communication, one must know that neurons at rest have a net negative charge relative to the external milieu to result in a membrane voltage of approximately -65mV [Bertil, 2001; Johnston and Wu, 1995; Kandel et al., 2013; Koch, 1999]. This potential difference arises due to the relative concentration of ionic species between the inside and outside of cells. Overall, there is a net excess of positively charged sodium ions in the extracellular space that have a net electrochemical concentration gradient driving them inside cells. On the other hand there is a higher concentration of positively charged potassium ions within the cells (relative to the extracelluar space) that have a net electrochemical concentration gradient driving them to leave neurons [Bertil, 2001; Johnston and Wu, 1995; Kandel et al., 2013; Koch, 1999]. This gradient is energetically maintained by ATP-sodium-potassium pumps that line neuronal cell membranes and pump sodium ions out of and potassium ions into cells [Johnston and Wu, 1995; Kandel et al., 2013; Koch, 1999]. As described below, it is the transient passage of sodium and potassium ions through the neuronal membrane that is responsible for the generation of action potentials. Finally, there are also large negatively charged proteins inside and chloride ions outside neurons that contribute to the overall electrochemical gradients and membrane potential though they play less of a role in the rapid dynamics of action potentials [Johnston and Wu, 1995; Koch, 1999].

The initiation of an action potential starts at the interface between the soma and axon, termed the axon hillock [Johnston and Wu, 1995; Kandel et al., 2013; Koch, 1999] . Here the resting state can be disturbed due to the passage of ions across the cell membrane throughout the soma and dendrites. If the membrane potential depolarizes (becomes positive) enough, it can reach a threshold that will trigger an action potential (Figure 1.1B) [Johnston and Wu, 1995; Kandel et al., 2013; Koch, 1999] . During an action potential, there is an inward rush of positively charged sodium ions from the extracellular space into the cell passing through rapidly opening voltage gated sodium channels which results in an increase in the membrane voltage. At a slower rate, potassium permeable voltage gated ion channels open more to increase the membrane's permeability to potassium. In addition the sodium channels start inactivating (closing) reducing sodium ion permeability. This allows more potassium ions to exit the cells (efflux) and less sodium ions to enter the cell (influx) which repolarizes the membrane potential back to the resting state. The decrease in membrane potential overshoots and becomes more negative than the normal resting state before recovering slowly

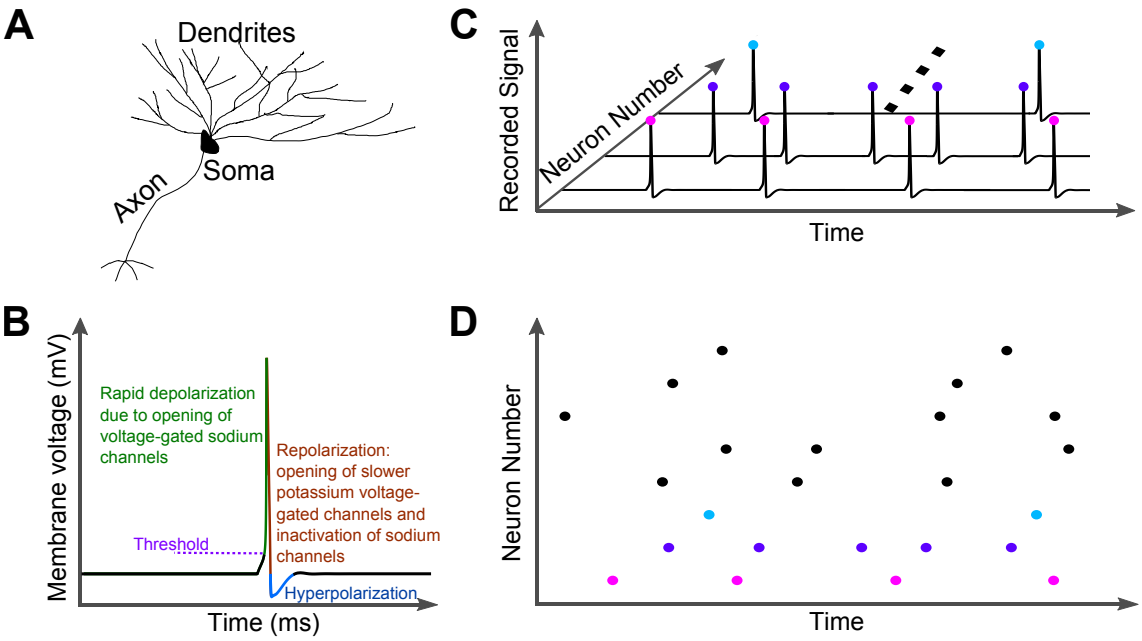

Figure 1.1: Neurons, action potentials, and raster plots. **(a)** Schematic of a neuron showing the dendrites, soma, and axon. The reader should be aware that this is just a stereotypical neuron diagram and that neurons are actually found in a variety of sizes and morphologies [Ramón y Cajal, 1888, 1933]. **(b)** Schematic of an action potential describing the different phases. Action potentials *in vivo* are in the order of 1ms in duration. **(c)** Illustration of many neurons recorded simultaneously and how action potentials can be identified for each neuron and recorded (colored dots). **(d)** Example raster plot which is a compression of the information in **(c)** by only representing the neurons and the time of action potentials. Note the y-axis can also be trial number if, for instance, a single neuron was recorded from multiple times in response to an identical stimulus.

back to baseline (a period termed hyperpolarization). This contributes to a time interval were neurons are unable to fire a second action potential called a refractory period [Johnston and Wu, 1995; Kandel et al., 2013; Koch, 1999]. Figure 1.1B shows an example simulation of an action potential. An action potential lasts for approximately 1ms *in vivo* and action potentials are self-sustaining such that, if activated at the axon hillock (interface between the soma and axon), they propagate down the whole axon to its terminals without decay of amplitude or change of shape [Johnston and Wu, 1995; Kandel et al., 2013; Koch, 1999]. It is also worth emphasizing that action potentials of a single neuron are identical in shape, termed a waveform [Gerstner et al., 2014; Kandel et al., 2013]. Thus no information is carried by the waveform of an action potential but in the rate and timing of the action potentials [Gerstner et al., 2014]. Throughout the literature and this thesis, an action potential can be referred to as a *spike* and neurons can also be said to *fire* an action potential.

1.1.3 What are synapses?

The variation in the number of neurons between animal species is vast, from the hundreds to the billions. The round worm *Caenorhabditis elegans*, a popular organism for neuroscientific study, has 302 neurons [White et al., 1986], the fruit fly, *Drosophila melanogaster*, has approximately 135,000 [Alivisatos et al., 2012], while mice and humans brains have 71 million [Herculano-Houzel et al., 2006] and 86 billion [Azevedo et al., 2009] neurons, respectively. To understand the mechanisms underpinning behavior and higher cognitive abilities that arise from these neurons, scientists also need to understand how these neurons are connected. The connection junction between two neurons is called a synapse.

It is critical to note that connectivity is not the last step in explaining the complexity of the brain. For instance, synapses can be of different types, the main division being electrical and chemical synapses [Kandel et al., 2013]. Electrical synapses refer to physical attachment between two neurons at regions termed gap junctions. These connections have cytoplasmic continuity between both cells. As such electrical signals between neurons can cross due to the electrical continuity [Kandel et al., 2013]. Chemical synapses, which are more abundant, have the presynaptic neurons release chemicals called neurotransmitters from synaptic vesicles. These neurotransmitters diffuse across the synaptic cleft, tiny space between the pre and postsynaptic neurons, and bind to postsynaptic neuron channels causing them to open and either hyperpolarize or depolarize the postsynaptic compartment via the passage of ionic species [Kandel et al., 2013]. Contrary to electrical synapses, chemical synapses are unidirectional. Frequently, neurons are categorized based on their synaptic properties [Harris and Shepherd, 2015]. One basic categorization of neurons is grouping them into excitatory and inhibitory neurons. Excitatory neurons release neurotransmitters that will depolarize neurons promoting the initiation of action potentials. Antagonistically, inhibitory neurons release neurotransmitters that hyperpolarize postsynaptic compartments and inhibit the ability of the post-synaptic neuron to send an action potential [Harris and Shepherd, 2015; Kandel et al., 2013; Roux and Buzsaki, 2015].

Given the diversity and intricateness of brains, it is not surprising that a whole field in neuroscience exists to try and uncover the complete neuronal connectivity of animals called connectomics [Behrens and Sporns, 2012]. The term connectome, in parallel to genome, refers to the complete neuronal connectivity map of organisms [Sporns et al., 2005]. As of today, the only species to have its connectome fully determined is that of the roundworm *C. elegans* [Seung, 2012; Varshney et al., 2011; White et al., 1986]. As will be discussed below, diverse technologies allow the connectome to be approximated at a spectrum of scales and this thesis focuses at three different levels.

It is not surprising that damage or abnormalities to the connectome can have drastic

consequences on information transmission and integration for an animal [Seung, 2012]. For instance, the retina connects and transmits information to higher visual brain structures via a bundle of neuronal fibers called the optic nerve [Selhorst and Chen, 2009]. If an optic nerve is damaged and no longer able to transmit information, then an animal will be blind in that eye. Other examples are Parkinson's disease which is caused by death of dopaminergic (dopamine producing) neurons in a brain structure called the substantia nigra [Davie, 2008] and Alzheimer's disease characterized by a reduction in the amount of long-range neuronal connections and overall neuronal death [Yao et al., 2010]. Other neurological conditions such as obsessive compulsive disorder and schizophrenia are also thought to be a result of abnormal wiring of the brain [Bullmore and Sporns, 2012].

At this point, I hope the reader has developed an appreciation to the complexity of neuronal networks. The connectivity is vital in determining the emergent properties of the network as discussed below.

1.1.4 What are raster plots?

There are numerous ways to collect neuronal data. Often, neuroscientist are only interested when neurons fire action potentials and this data can be represented in a raster plot as illustrated in Figures 1.1C and 1.1D. Raster plots can be thought of as compressing the data by removing the membrane voltage axis as researchers may not be interested in the waveform of spikes (identical and hence lack information) or subthreshold dynamics of neurons. Thus, time is on the x-axis, whereas the y-axis is discretized and commonly represented as neuron number (Figure 1.1D). Every neuron would be assigned a number and each time a neuron spikes, a dot (or vertical line) is used to represent that spike at a given time. The y-axis can also be trail number if the data was collected from a single neuron while a stimulus was repeated many times. Raster plots are introduced here as they are seen in more than half of the figures in this thesis.

1.2 Brain Networks

The nervous system is studied at multiple levels of granularity [Sporns, 2013]. Here, a quick introduction to the different scales is given although it must be noted that there is no absolute way of dividing the scales and the below is simply a summary of many commonly studied network scales. In all scales introduced, an analogy to airports and their connectivity is given to help the reader build intuition.

1.2.1 Subcellular networks

The discoveries made about the inner workings of neurons are simply remarkable. The complexity is appreciated by looking at the genetic expression patterns of neurons and how different genes interact (via protein production) and influence one another [Cooper-Knock et al., 2012; Tasic et al., 2016]. The genetic expression patterns govern the cell development processes in addition to a cell's overall properties and hence cell-type [Ronan et al., 2013]. Recent work has shown that the genetic level is not the only one of interest and that transcript and protein network expressions are vital [Tasic et al., 2016], the focus of the fields of trancriptomics [Cooper-Knock et al., 2012] and proteomics [Prescott et al., 2014]. There are also remarkable intricacies seen at postsynaptic membranes termed postsynaptic densities [Kennedy, 2000]. Here, the mechanisms and interactions of proteins within the cell and on the cell membrane are vast and dynamically changing at all times in response to internal and external stimuli.

Although this scale is not considered in this thesis, it is a field of neuroscience that garners a lot of attention and research. In analogy to airports, which will be done throughout this section, these networks can be thought as the internal operations of every airport individually such as the handling of luggage, gate assignment, terminal transfers, and other processes that keep the complex airport itself functional.

1.2.2 Microscopic networks

On neuronal scales, there is much interest in knowing the full connectivity at a synaptic level between all neurons as done for *C. elegans* [Seung, 2012; Varshney et al., 2011; White et al., 1986]. Current technologies are able to attain such information very slowly where animal brains can be sectioned into very thin slices ($\sim$25nm thick) and then imaged using electron microscopy [Briggman et al., 2011]. Once this is done for consecutive of slices, neurons are identified using semi-automated methods to trace the shape of every neuron in every slice, determine where they end and connect to other neurons to identify synapses [Denk et al., 2012]. This not only gives the full connectivity (microscopic network) but also the morphology of all cells.

The microscopic scale can be considered to be similar to knowing the flights of every airport in the world and how they are connected, where a connection is defined as a single flight between two airports. In this thesis, Chapter 3 models neuronal networks with full control over the synaptic connectivity of the network. The chapter investigates the crucial link between connectivity and spatio-temporal dynamics in neuronal networks. The work concentrates on a topic that has wide current interest: the emergence of dynamical cell

assemblies in networks, i.e., groups of neurons that show long-lived, coherent co-activity. Specifically, the chapter characterize the appearance of slow-switching assemblies (SSA) in leaky-integrate-and-fire (LIF) neuronal networks, i.e., a coherent increased firing in subgroups of neurons, sustained over long times and switching from group to group across the network.

Chapter 3 further shows that the ability of clustered LIF networks to support SSA activity can be determined by evaluating spectral properties of the *connectivity matrix*. I illustrate that the strength of SSA activity can be linked to the existence of a gap separating the leading eigenvalues, together with a block-localization of the associated Schur vectors of this matrix on the neuron groups acting coherently.

To understand the relevance of the connectivity eigen-gap, I consider stylized firing rate models and use the analytical insights they provide to develop LIF networks based on completely different topological and functional connectivity paradigms, which are nonetheless able to display SSA activity (e.g., SSA dynamics involving excitatory and inhibitory neurons, or SSA activity with a hierarchy of timescales).

The findings can be used to indicate if a network can support SSA activity based on spectral analysis which is relevant, for example, in the emergence of oscillations so crucial for information processing in the brain [Buzsaki and Moser, 2013]. Furthermore, this knowledge can be used to propose distinct types of topologies of relevance in neuroscience, which can display SSA dynamics. The Chapter illustrates that advancements in neuronal monitoring and connectomics go hand-in-hand to facilitate our understanding of the relationship between structure, dynamics, and function.

1.2.3 Mesoscale networks

Current technologies also allow the determination of brain coonectivity at coarser scales. Such methods are faster and have allowed the determination of the complete mouse brain mesoscopic connectivity [Kuan et al., 2015; Oh et al., 2014]. One technique involves injecting mice brains with modified adeno-associated virus (AAV) that expresses a fluorescent tracer (green fluorescent protein (GFP))[Harris et al., 2012]. The virus is taken up by projection neurons within small injection volumes and transported anterogradely all the way to their synaptic terminals [Chamberlin et al., 1998; Harris et al., 2012]. Following weeks of protein expression, this fluorescent signal is imaged (via a set of specialized two-photon microscopes) throughout the entire brain, informatically reconstructed, and mapped to a high-resolution coordinate system [Kuan et al., 2015; Oh et al., 2014]. At this scale, investigators can only identify axonal fiber projections to overall brain structures without discriminating synaptic terminals (due to limitations of the point spread function of imaging and microscopy)[Kuan

et al., 2015; Oh et al., 2014]. Using the airport analogy, this is equivalent to identifying how certain overall regions with some airports are connected to other regions without knowing which airports in particular. For instance, one may know there are flights between Southern California and the New York city area, without knowing any specific flights or which of the airports are actually connected. The strength of these connections can still be estimated as they will be proportional to the number of flights between the two regions. Similarly in mesoscale networks, the strength of connections can be estimated by the fluorescent signals at target areas [Kuan et al., 2015; Oh et al., 2014].

In Chapter 5, the effect of chronic sleep deprivation during adolescence on brain mesoscale connectivity is investigated. In this study, a control group and a sleep deprived group (during adolescence) of mice are compared. After normalizing for inherent experimental and biological noise, I observe that sleep restriction did not alter the mesoscale connectivity between secondary motor cortex (site of injection in this data set) and the rest of the brain in a unitary manner. Using a novel classification algorithm allowed the identification of the two groups significantly above chance level indicating long lasting global changes in the brain do occur due to chronic sleep deprivation in adolescence. The changes are intricate and not a simple overall effect. To further unravel the differences between both groups it may be necessary for future work to develop new analysis techniques, more sensitive recording methodologies and injection techniques, or inject in multiple brain locations. In brief, the age-old advice of not sacrificing sleep was further verified and a step was taken towards understanding how sleep deprivation affects brain connectivity.

1.2.4 Macroscopic networks

At an even coarser scale, brain connectivity between structures can be estimated by determining the projections of white matter brain tissue between different structures [Bassett et al., 2011]. Common techniques for achieving this are through the use of diffusion-based magnetic resonance imaging [Jbabdi and Johansen-Berg, 2011; Johansen-Berg, 2013; Sotiropoulos et al., 2013]. These techniques have been available for decades and work by mapping the diffusion of water in brain matter and other biological tissue [Merboldt et al., 1985; Taylor and Bushell, 1985]. The water diffusion speed is influenced by the underlying structure imposed by fibers, membranes and other biological variables which can be mapped in three dimensions [Basser et al., 1994]. In the brain, water diffusion in neuroal axons is anisotropic due to their fibrous structures [Basser et al., 2000; Conturo et al., 1999]. By measuring this anisotropy, white matter tracts can be estimated from the dominant diffusion directions [Basser et al., 2000; Conturo et al., 1999]. Such brain imaging methodologies are useful probes in understanding the causes of different diseases and aging [Alexander et al.,

2007; Bassett et al., 2011].

In the aeronautical analogy, this is an even coarser resolution than the above mesoscale connectome where for instance flight connectivity might be identified on a country/continent scale. For instance, are there direct flight connections from Europe to Australia? The work presented herein does not include modeling or analysis of macroscopic networks although this is a vibrant field of neuroscience that is essential in expanding our understanding of the brain.

1.2.5 Functional networks

Networks can also be considered on more abstract levels beyond physical connectivity. Although topology constrains the activity a network can exhibit, it is nonetheless the activity of the network that allows for our sensory and cognitive abilities. From one viewpoint, neurons may or may not be connected, but their activity can be correlated such that they appear to perform related functions. As such, based on the similarity of neuronal activity, metrics (statistical associations) can be defined between neuron pairs that identify how functionally connected they are [Bullmore and Sporns, 2009]. When this is done for all neurons, a functional network is said to be determined. Of course this is less absolute compared to actual physical connectivity as the actual metric used will alter the estimated functional network. Nonetheless, the notion is very useful and is also implemented at larger scales using technology such as functional magnetic resonance imaging (fMRI) or electroencephalography (EEG) [Betzel et al., 2012; Bullmore and Sporns, 2009; Chu et al., 2012].

Continuing with the airport analogy, functional networks are equivalent to finding airports with similar activity but are not necessarily connected to each other. This may be of interest to certain researchers as these airports may provide similar overall roles for the network at large and as such want to be identified together. An example is small airports that train pilots in addition to being strategically located for emergency landings. Other examples such as large airports might also form a group and may in fact be connected by direct flights, though in this analysis they are grouped due to similar behavior and not physical connectivity.

Estimating microscopic functional connectivity is expanding due to rapid advancements in neuronal monitoring techniques, such as calcium imaging and multi-electrode arrays, that enable neuroscientists to monitor activity from hundreds or more neurons simultaneously [Ahrens et al., 2013; Buzsaki, 2004]. This functional network view of neuroscience is considered in Chapter 2 that presents a versatile and robust method to detect cell assemblies from a novel metric for estimating functional networks. Identifying cell assemblies, groups of neurons that cooperate in some form within the brain, is critical for systems neuroscience

in order to gain insight into the way neural processing and computations are performed at the population level [Buzsaki, 2010; Hebb, 1949]. The technique works on neuronal spiking data applicable across a spectrum of relevant scales that are gathered from modern systems neuroscience experiments. Initially, a novel *biophysically inspired* measure is used to extract *directed* functional relations between *both excitatory and inhibitory neurons* based on their spiking time history. Second, the resulting induced network representation is then analyzed using a graph-theoretic community detection method to reveal groups of related neurons (cell assemblies) in the recorded time series at *all levels of granularity*, without prior knowledge of their relations or expected size. The methodology is extensively assessed through synthetic, simulated, and experimental spike-train data and highlights the advantage of the functional network perspective in finding neuronal communities based on associated activities.

1.2.6 Feedforward networks

A commonly studied architecture that can also be observed at multiple scales is feedforward networks. The topology of such networks is designed such that information is transmitted consecutively along groups (layers) [Kumar et al., 2010; Vogels et al., 2005]. Examples are seen in visual processing where one path of information propagation traverses from the retina to the dorsal lateral geniculate nucleus to the visual cortex [Grubb et al., 2003; Hendry and Reid, 2000]. Another common example is the hippocampus that receives input from the entorhinal cortex that is connected dentate gyrus to the CA3 region to the CA1 region [Llorens-Martin et al., 2014; Olsen et al., 2012]. In the airport analogy, the information can be considered as individual travels. People who are traveling around the world might be restricted from their travel agent to only go in one direction around the planet. As such, the travelers (information) have restricted paths and can only move from one region to the next in a parallel manner to feedforward networks. Scientists have investigated these networks for decades as prototypical examples of how information could be propagated within an organism. Numerous computations and theories have been developed and studied, making this a very advanced and elegant field of research [Kumar et al., 2010; Vogels et al., 2005].

Despite the wide interest in the properties and functioning of feedforward networks, however, in common models the feedforward activity is exclusively driven by excitatory units and the wiring patterns between them. Inhibitory interactions, on the contrary, play only a stabilizing role to ensure that the network dynamics do not saturate. In Chapter 4, motivated by recent discoveries of hippocampal circuitry [Nasrallah et al., 2015] and the diversity of inhibitory neurons throughout the brain [Harris and Shepherd, 2015; Roux and Buzsaki, 2015], I introduce network architectures capable of directed information transmis-

sion whose activity centrally involves *inhibitory neurons* and their connectivity structure, while *excitatory neurons remain randomly connected between each other*. Specifically, the findings point to the fact that feedforward activity observed in the brain might be caused by a much broader architectural basis than has been considered.

1.3 Recapitulation

This chapter introduced the basic cells of the brain, neurons, their connection points, synapses, and their communication currency, action potentials. The representation of neuronal data in raster plots was also presented as multiple will be seen in Chapters 2, 3, and 4. An overview was given into how the brain can be studied at multiple levels of abstraction many of which will be discussed in the following chapters. These topics were selected for the convenience of the reader to provide essential background and context for the work presented herein.

Chapter 2

Revealing cell assemblies at multiple levels of granularity

2.1 Abstract

Identifying cell assemblies, or groups of neurons that cooperate within large neural populations, is of primary importance for systems neuroscience. We introduce a simple biophysically inspired measure to extract a directed functional connectivity matrix between both excitatory and inhibitory neurons based on their spiking history. The resulting network representation is analyzed using a graph-theoretical method for community detection to reveal groups of related neurons in the recorded time-series at different levels of granularity, without *a priori* assumptions about the groups present. We assess our method using synthetic spike-trains and simulated data from leaky-integrate-and-fire networks, and exemplify its use in experimental data through the analysis of retinal ganglion cells of mouse and salamander, in which we identify groups of known functional cell types, and hippocampal recordings from rats exploring a linear track, where we detect place cells with high fidelity.

2.2 Introduction

As capabilities for parallel recordings from large neuronal populations continue to improve [Ahrens et al., 2013; Buzsaki, 2004] experimentalists are now able to probe neural population encoding in ever more detail. These experimental advances allow the study of the intricate

links between topology and dynamics of neural interactions, which underpin the functional relationships within neural populations. One such example is the activity of cell assemblies. The problem is to identify groups of neurons (termed cell assemblies) within a large number of simultaneously recorded neurons where, due to functional cooperativity, each cell in an assembly is more similar in its temporal firing behavior to members of its own group than to members of other groups. Such strongly intertwined activity patterns are believed to underpin a wide range of cognitive functions [Buzsaki, 2010; Harris, 2005; Hebb, 1949]. However, the reliable identification of cell assemblies remains challenging.

Here we introduce a technique to identify such neuron assemblies directly from multivariate spiking data, based on two steps: the definition of a simple biophysically-inspired similarity measure obtained from the observed spiking dynamics, followed by its analysis using a recent framework for multiscale community detection in weighted, directed graphs. A variety of techniques have been proposed to cluster spike-train groups to date, and have shown promising results in particular settings [Abeles and Gat, 2001; Feldt et al., 2009; Fellous et al., 2004; Gansel and Singer, 2012; Humphries, 2011; Laubach et al., 1999; Lopes-Dos-Santos et al., 2011, 2013; Peyrache et al., 2010; Quiroga and Panzeri, 2009]. In contrast to these techniques, our methodology provides a dynamics-based framework, in which both the similarity measure and the community detection method are geared towards incorporating key features of neural network dynamics. The framework is purposely designed to be simple, yet capturing a breadth of features not present concurrently in other methods.

Our similarity measure evaluates the association between neuron pairs based on their spiking history and integrates three features that are key for a network-based analysis of neuro-physiological data: (i) an intuitive biophysical picture, allowing a simple interpretation of the computed associations; (ii) a measure that is directed in time, hence asymmetric in the sense that spike-time dependent information is retained (*e.g.*, spiking of neuron A precedes that of neuron B); (iii) excitatory and inhibitory interactions are both included yet treated differently, inspired by their distinct effects on post-synaptic cells.

The detected dynamic associations are interpreted as an induced functional network, which is used to identify neuronal assemblies using a directed version of the recently introduced Markov Stability framework for community detection in graphs [Delvenne et al., 2010]. Unlike other approaches, this framework allows us to analyze directed networks and search for cell assemblies at all levels of granularity, from fine to coarse levels of resolution, extracting relevant, possibly hierarchical groupings in spike trains without *a priori* assumptions about the groups present. In the following, we present our framework and evaluate it on a series of examples, including synthetic spike-trains and leaky-integrate-and-fire network models. We also apply it to experimental datasets from retinal ganglion cells and

hippocampal pyramidal neurons.

2.3 Materials and Methods

Most existing methods to detect groups in spike-train neuronal population data are based on the following generic paradigm [Feldt et al., 2009; Fellous et al., 2004; Humphries, 2011; Lopes-Dos-Santos et al., 2011]. First, a metric is defined to quantify the relationship between all neuron pairs leading to a $N \times N$ association matrix, where N is the number of observed neurons. We call this the *functional connectivity matrix* (FCM) hereafter. Every (i, j) entry in this matrix is a non-negative number that indicates how similar the spike trains of neurons i and j are over the observed time. Second, the FCM is clustered, *i.e.*, partitioned into different groups [Aggarwal and Reddy, 2014; Fortunato, 2010; Newman, 2004].

Here we introduce a simple framework that addresses both of these steps in a consistent and integrated manner, focusing on the dynamical relations between neurons: a new directed ('causal') biophysically-inspired measure is introduced to calculate the FCM, which is then analyzed using the recently introduced dynamics-based technique of Markov Stability for community detection [Delvenne et al., 2013, 2010; Lambiotte et al., 2009; Schaub et al., 2012] to identify cell assemblies at multiple scales in the neuronal population.

The numerics are performed in MATLAB (2011b or later versions). Code implementing the algorithm for spike-train analysis is available upon request and can be found at: `github.com/CellAssembly/Detection`.

2.3.1 Biophysically-inspired causal measure of spike-train similarity

A plethora of metrics exists to describe the relationship between two signals, ranging from generic measures, such as cosine similarity and Pearson or Spearman correlation coefficients, to specialized measures designed for spike-train analysis [Fellous et al., 2004; Kreuz et al., 2013; Lyttle and Fellous, 2011; Okatan et al., 2005; Schreiber et al., 2003; van Rossum, 2001; Victor and Purpura, 1996; Vincent et al., 2012]. Although these methods can be well suited in particular contexts, they only partially account for three important features for network-driven analyses of neural recordings. First, most current metrics are based on statistical arguments lacking a simple biophysical interpretation that would allow the use of relevant biophysical characteristics of neuronal dynamics. Second, most commonly used measures are distance metrics, *i.e.,* symmetric by construction, and thus neglect spike-timing information contained in the ordering of events. Finally, to the best of our knowledge, all measures ignore whether the neurons under consideration are excitatory or inhibitory.

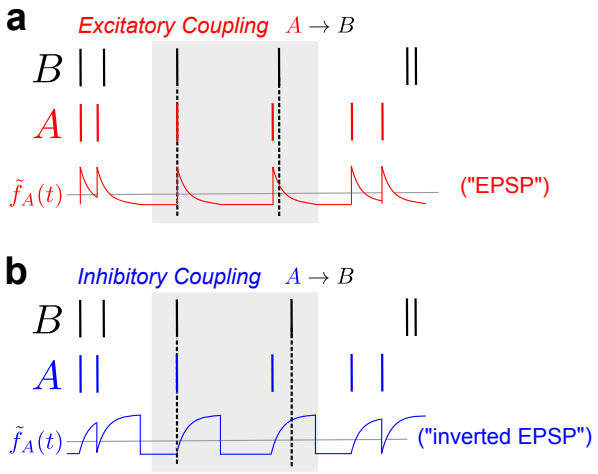

Figure 2.1: Biophysically-inspired measure of spike-train similarity leading to functional coupling between neurons. Quantification of the coupling induced by: **(a)** excitatory neuron A on neuron B and **(b)** inhibitory neuron A on neuron B. Note that both profiles shown are normalized so that the signal has zero mean (see text).

While an even finer characterization of neuronal subtypes could be of further interest, the distinction between excitatory and inhibitory neurons underpins fundamental balances in neuronal network dynamics and should be reflected in the analysis of data. Here, we propose a similarity measure that incorporates these three ingredients in a simple, intuitive form (see Figure 2.1).

Consider first an excitatory neuron A connected to neuron B. The action potentials of A induce excitatory postsynaptic potentials (EPSPs) in neuron B, increasing the likelihood of neuron B firing. These EPSPs can be, to a first approximation, modeled by an exponentially decaying time profile

$$\xi_{\text{exc}}(t) = e^{-t/\tau}$$

with synaptic time constant τ. Since detailed information about synaptic weights and membrane potentials is unavailable in neuronal population experiments, we adopt a simple strategy to compute the coupling strength S_{AB} from the observed spiking data. The general idea is that for each spiking event of neuron B (at time t_i^B), we propagate a 'virtual' EPSP from the immediately preceding spike of neuron A (at time t_i^A). We then compute all such contributions that neuron A would have made to the membrane potential of neuron B at each of its spikes, and sum them appropriately, discounting spurious effects.

More precisely, we obtain the functional connectivity from neuron A to neuron B as follows:

(i) Define the signal $f_A(t)$ that reflects the (virtual) influence of neuron A onto a potential

firing event at any other neuron taking place at time t:

$$f_A(t) = \xi_{\text{exc}}(t - t^A_{\text{last}}) = e^{-(t-t^A_{\text{last}})/\tau}, \tag{2.1}$$

where $t^A_{\text{last}} = \max_i(t^A_i | t^A_i \leq t), i = 1, \ldots, N_A$, is the time of the last preceding spike of neuron A (if there is no such spike we set $t^A_{\text{last}} = -\infty$).

(ii) It then follows that all contributions from neuron A to B can be written as the sum $\sum_{i=1}^{N_B} f_A(t^B_i)$. To gain some intuition, note that every time B fires a spike, the potential contribution to this spike by neuron A is computed by summing the values that $f_A(t)$ takes at the times of B firing, t^B_i. If neuron B always fires shortly after A spikes, the sum $\sum_i^{N_B} f_A(t^B_i)$ will be large. If neuron B fires after A but with some delay (e.g., because an integration with other neurons is required), this sum will be smaller. If neuron B never fires shortly after A, this sum will be zero.

To discount spurious correlation effects, we center and normalize the signal $f_A(t)$ first to obtain the new signal $\tilde{f}_A(t)$, which has zero mean and peak amplitude one (Figure 2.1a)

$$\tilde{f}_A(t) = \frac{f_A(t) - \langle f_A \rangle}{1 - \langle f_A \rangle}, \tag{2.2}$$

where $\langle f_A \rangle = \frac{1}{T}\int_0^T f_A(t)dt \leq 1$ is the mean over the recorded time. We then compute the effective coupling:

$$F_{AB} = \frac{1}{N_{AB}} \sum_{i=1}^{N_B} \tilde{f}_A(t^B_i), \tag{2.3}$$

and we have additionally divided by $N_{AB} = \max(N_A, N_B)$ to guarantee that the maximal coupling F_{AB} (between two identically firing neurons) is normalized to 1. The coupling between neuron A and B is then defined as the thresholded value:

$$S_{AB} = \max(F_{AB}, 0). \tag{2.4}$$

From this definition, it follows that if an action potential from neuron A is always closely followed by a spike from neuron B, this will correspond to a strong coupling S_{AB} between these neurons. Note that, in addition to being biophysically inspired, the defined measure (2.4) is non-symmetric ($S_{AB} \neq S_{BA}$).

Suppose that neuron A is known to be inhibitory. The coupling strength from neuron A onto another neuron is obtained following a similar approach (Figure 2.1b), yet recognizing that inhibitory post-synaptic potentials (IPSPs) decrease the likelihood of firing. To reflect

this influence, we adopt an 'inverted' exponential profile

$$\xi_{\text{inh}}(t) = 1 - e^{-t/\tau},$$

truncated when it reaches 99% of its steady state value. Hence, if neuron B always fires shortly after the firing of the inhibitory neuron A, it will accumulate a negative dependence from which we deduce that there is no significant inhibitory functional relation between these neurons.

The time scale τ is a parameter inspired by synaptic time constants, and can thus be adapted to reflect prior information about the recorded neurons. Although more sophisticated schemes to estimate or tune this parameter are certainly possible (*e.g.*, choosing a different τ_{exc} for excitatory and τ_{inh} for inhibitory neurons), here we follow the simplest choice $\tau_{\text{exc}} = \tau_{\text{inh}} = \tau$ throughout. The method is robust to the choice of τ: we have used $\tau = 5$ ms for the experimental data and $\tau = 3$ ms for the leaky-integrate-and-fire (LIF) simulation data, and have checked that the results remain broadly unaltered for values of τ in this range.

The main aim of our measure is simplicity, flexibility and generality, while retaining the key biophysical features outlined above. Because of its generality, highly specialized measures of spike-train associations could be tuned to outperform our simple measure for particular examples. However, it is often unknown beforehand what features of the data are of importance for the analysis and hence having such a flexible measure allows for a broad search for structure in recorded data. Once a hypothesis is formed, or particular aspects need to be investigated in more detail, more specialized association metrics could be used in conjunction with the community detection algorithm presented below. In the absence of knowledge about the specific cell types of experimentally recorded neurons we obtain the FCM using the excitatory metric. Already today, however, there are means to separate cell types (*e.g.* fast spiking interneurons) based on their electrophysiological signature [Barthó et al., 2004] and with the advancement of optical physiology and genetic tools, additional information about the cell types of the recorded cells is becoming more routine. Hence it will be possible in the future to use specialized coupling functions (instead of exponential) depending on the neuronal sub-type recorded.

2.3.2 Markov Stability for community detection at all scales

The Markov Stability method is a versatile, dynamics-based tool for multiscale community detection in networks without *a priori* assumptions about the number or size of the communities [Delvenne et al., 2013, 2010; Lambiotte et al., 2009; Schaub et al., 2012]. Here

we extend the use of Markov Stability to directed networks to find coherent groupings of neurons in the FCM created from the observed spiking data. Under our framework, we interpret the FCM as a directed network, and the graph communities revealed by our analysis correspond to groups of neurons with strong excitatory and/or inhibitory couplings extracted from the dynamics. Therefore the graph partitioning problem solved using the Markov Stability method is linked to the detection of putative cell assemblies, *i.e.*, groups of neurons with a strong dynamical influence on each other.

The main notion underpinning the Markov Stability method is the intimate relationship between structure and dynamics on a graph. A dynamics confined to the topology of a network can uncover structural features of the graph by observing how a dynamical process, such as a simple diffusion, unfolds over time. In particular, if the graph contains well defined substructures, such subgraphs will trap the diffusion flow over a significantly longer time than expected if it were to happen on an unstructured graph. This idea is readily illustrated by the example of ink diffusing in a container filled with water. If the container has no structure, the ink diffuses isotropically. If the container is compartmentalized, the ink would get transiently trapped in certain regions for longer times until it eventually becomes evenly distributed throughout. In a similar manner, by observing the dynamics of a diffusion process we can gain valuable information about the structural organization of the graph (Figure 2.2). We use this concept to define a cost function to detect significant partitions in the graph, as follows.

To make these notions precise, consider a network with a Laplacian matrix $L = D - A$, where A is the weighted adjacency matrix (A_{ij} is the weight of the *directed* link from node i to node j) and $D = \mathrm{diag}(A\mathbf{1})$ is the diagonal out-degree matrix ($\mathbf{1}$ is the vector of ones). For ease of explanation, consider first a strongly connected graph, *i.e.*, we can traverse the graph along its directed edges such that every node can be reached from any other node. On such a network, let us define a continuous diffusion process:

$$\dot{\mathbf{p}} = -\mathbf{p}\, D^{-1} L, \tag{2.5}$$

where $\mathbf{p}$ is the $1 \times N$ probability vector describing the probability of a random walker to visit different nodes over time. Note that the probability vector remains properly normalized: $\mathbf{1}^T\mathbf{p} = 1$ at all times. For an undirected connected graph, this dynamics converges to a unique stationary distribution $\boldsymbol{\pi} = \boldsymbol{d}/(\boldsymbol{d}^{\boldsymbol{T}}\mathbf{1})$. For *directed* graphs the stationary distribution has to be computed by solving $\dot{\mathbf{p}} = 0$, *i.e.*, it corresponds to the dominant left eigenvector of $D^{-1}L$. If the graph is not strongly connected (*e.g.*, if it contains a sink), the diffusion process (2.5) is generalized to include the standard random 'teleportation' term inspired

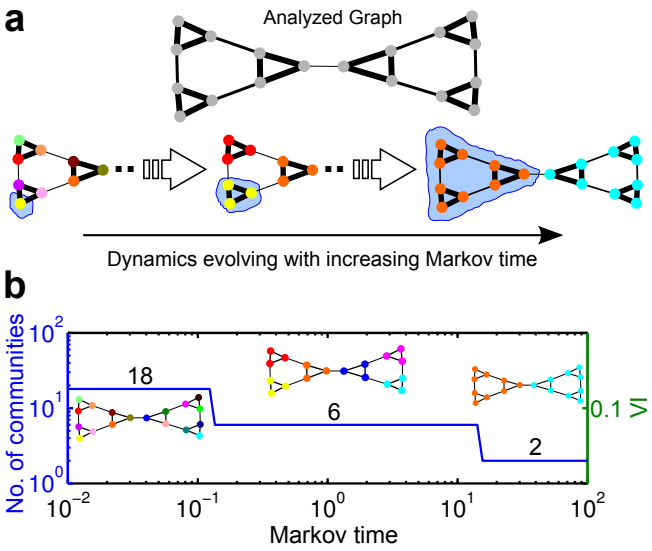

Figure 2.2: Schematic of Markov Stability method used to partition the functional network. **(a)** A diffusion process on a network can be used to reveal the structure of a graph. As the diffusion explores larger areas of the network, it enables the Markov Stability method to scan across all scales and reveal relevant partitions at different levels of granularity. **(b)** The graph analyzed has a pre-defined multi-scale community structure, given by a hierarchy of triangles. The number of communities found are plotted as a function of the Markov time (see **(a)**) long plateaus indicate well-defined partitions into 18 nodes (each node on its own), six communities (small triangular structures), and two communities (aggregated, larger triangles). Note that in this example, the variation of information (VI) is zero for all Markov times, indicating that all three partitions are relevant at different levels of resolution. We remark that the sudden steps between plateaus is a result of the specific example chosen. Since many of the weights were chosen to be identical, there is a lot of symmetry in the graph which results in many repeated eigenvalues in the stability matrix. This causes the sharp transition between plateaus during the Louvain optimization step (see text). The addition of a small amount of variation in the weights would result in a smooth transition between plateaus (provided Markov time is sampled finely enough).

by Google's page-rank algorithm [Brin and Page, 1998; Lambiotte et al., 2009; Lambiotte and Rosvall, 2012]: the random walker is transported from any node to a random node in the graph with a small, uniform probability α (set here to the commonly adopted value $\alpha = 0.15$), while in the case of a sink node, it will be teleported with unit probability. This term guarantees that the process is ergodic with a unique stationary probability distribution.

Consider a partition of this network encoded in a $N \times c$ indicator matrix H, with $H_{ij} = 1$ if node i belongs to community j. We then define the Markov Stability of the partition $r(t_M, H)$, as the probability that a random walker at stationarity starts in community i and ends up in the same community after time t_M minus the probability of such an event happening by chance, summed over all communities and nodes. In matrix terms, this may be expressed as:

$$S(t_M) = \Pi \exp(-t_M D^{-1} L) - \boldsymbol{\pi}\boldsymbol{\pi}^T$$

$$r(t_M, H) = \text{trace}\left[H^T S(t_M) H\right],$$

where $\Pi = \text{diag}(\boldsymbol{\pi})$ and t_M denotes the Markov time describing the evolution of the diffusion process. Finding a good partition (or clustering) requires the maximization of the Markov Stability in the space of possible graph partitions for a given t_M, an optimization that can be carried out with a variety of optimization heuristics. Here we use a locally greedy optimization, the so-called Louvain algorithm, which is highly efficient [Blondel et al., 2008]. In order to deal with the fact that $S(t_M)$ is in general asymmetric due to the directed nature of the graph, we use the directed notion of Markov Stability and use the Louvain algorithm to optimize $H^T \frac{1}{2}(S + S^T)H$, which is mathematically identical to optimizing $r(t_M, H)$, i.e., we still consider the *directed* network.

Our algorithm then scans across all Markov times to find the set of relevant partitions at different Markov times. With increasing Markov time, the diffusion explores larger regions of the network, resulting in a sequence of increasingly coarser partitions, each existing over a particular Markov time scale. The Markov time may thus be interpreted as a resolution (or *granularity*) parameter, and, as we sweep across resolutions, we detect communities at different levels of granularity without imposing a particular resolution *a priori*. This dynamic sweeping [Schaub et al., 2012] allows us to detect assemblies of different sizes and even hierarchical structures that would potentially go undetected if we were to use a method with a fixed intrinsic scale [Feldt et al., 2009; Fellous et al., 2004; Fortunato and Barthélemy, 2007; Humphries, 2011; Lopes-Dos-Santos et al., 2011; Newman and Girvan, 2004]. It is important to remark that the Markov time t_M used for the diffusive exploration of the network is not to be confused with the physical time of the spike-train dynamics. We remark that the time constant τ of our similarity measure is not related to the Markov time in general. The Markov time is used here as a tool to uncover the different scales in the data and should thus be seen as distinct from the biophysical (real) time.

To select meaningful partitions across levels of granularity, we use two measures of robustness. Firstly, a relevant partition should be persistent over a long Markov time-horizon, *i.e.*, it should be robust with respect to the change in Markov time and thus lead to an extended plateau in Markov time. Secondly, a relevant partition should be consistently found by the optimization algorithm, *i.e.*, it should be robust to random initializations of the Louvain optimization. In order to establish the optimization robustness, we run the Louvain algorithm 100–500 times per Markov time and compare the partitions obtained by means of the variation of information (VI) distance metric [Meila, 2003, 2007]. The variation of information can be thought of as an information-theoretic distance between two partitions that is naturally invariant to a relabeling of the groups and which has proved useful as

a standard tool to compare partitions in the context of community detection [Fortunato, 2010]. The normalized VI between two partitions $\mathcal{P}^\alpha$ and $\mathcal{P}^\beta$ is defined as [Meila, 2007]:

$$\mathrm{VI}(\mathcal{P}^\alpha, \mathcal{P}^\beta) = \frac{2\,H(\mathcal{P}^\alpha, \mathcal{P}^\beta) - H(\mathcal{P}^\alpha) - H(\mathcal{P}^\beta)}{\log N}, \tag{2.6}$$

where $H(\mathcal{P}) = -\sum_{\mathcal{C}} p(\mathcal{C}) \log p(\mathcal{C})$ is the Shannon entropy of the relative frequency $p(\mathcal{C}) = n_{\mathcal{C}}/N$ of a node belonging to community $\mathcal{C}$ in a partition $\mathcal{P}$ and $H(\mathcal{P}^\alpha, \mathcal{P}^\beta)$ is the Shannon entropy of the corresponding joint probability. We then calculate the average variation of information (VI) over all pairs in the ensemble of solutions from the optimization. When $VI \approx 0$, the solutions obtained by the different optimizations are very similar to each other indicating a robust partitioning. When $VI \approx 1$ each run of the optimization obtains a different partition, indicating a non-robust clustering. Such clear-cut communities are not always found. However, we have shown [Delmotte et al., 2011; Schaub et al., 2012] that sudden drops and dips in the VI are indicative of a clustering becoming more robust than expected for its average community size. In realistic datasets, we thus search for partitions with a long Markov time plateau and a low value (or a pronounced dip) of VI as the criterion to find meaningful partitions. An illustration of the Markov Stability framework is displayed in Figure 2.2b, where we exemplify how the graph community structure can be detected at different scales without *a priori* assumptions about the number of communities. Furthermore, our scanning across all Markov times allows for the detection of the appropriate scale for community detection, without imposing *a priori* a particular scale that might not be relevant to the analyzed data, as is implicitly done in other methods [Schaub et al., 2012].

2.3.3 Synthetic spiking data

To assess the capabilities of the framework, we generated synthetic spiking datasets with realistic statistical properties resembling those observed in experiments, yet with added temporal structure.

2.3.3.1 Synthetic data with embedded and hierarchical cell assemblies

Surrogate spike-train data were created from groups of units with variable sizes. Each group G_i was assigned a firing rate (f_i) and a level of jitter (J_i). The firing times of each *group* were drawn from a uniform distribution according to the specified firing frequency f_i, and the firing times for each *unit* were chosen from a uniform distribution with a range $\pm J_i$ around the group firing time. To account for refractory periods, we resampled if the resulting spike time conflicted with the refractory period of the unit. We used a similar scheme to generate synthetic spiking data with a hierarchical structure, but in this case each group was divided

into two subgroups: units within each subgroup always fire together, whereas between two subgroups the firing window was aligned only every second time. As before, the firing times of the individual groups were chosen randomly from a uniform distribution and were not correlated in time. This firing pattern establishes a two-level hierarchical relation between the individual units.

2.3.3.2 Synthetic data with feedforward-like firing patterns

Synthetic spiking patterns that emulate the activity of feedforward networks were created from groups that are made to spike together within a jitter window of ± 1 ms. The groups are set to spike sequentially with a delay of $\delta = 5$ ms and a repetition period of $\Delta = 20.5$ ms.

2.3.4 Simulated data from Leaky-Integrate-and-Fire Networks

We applied our algorithm to more realistic spiking computational datasets obtained by simulating neuronal networks of excitatory and inhibitory Leaky-Integrate-and-Fire (LIF) neurons [Koch, 1999].

2.3.4.1 The excitatory and inhibitory LIF units

The non-dimensionalized membrane potential $V_i(t)$ for neuron i evolved according to:

$$\frac{dV_i(t)}{dt} = \frac{\mu_i - V_i(t)}{\tau_m} + I_{\mathrm{S}}, \tag{2.7}$$

where the constant input term μ_i was chosen uniformly in the interval $[1.1, 1.2]$ for excitatory neurons and in the interval $[1, 1.05]$ for inhibitory neurons. Both excitatory and inhibitory neurons had the same firing threshold of 1 and reset potential of 0. Note that although the input term is supra-threshold, balanced inputs guaranteed that the average membrane potential remained sub-threshold [Litwin-Kumar and Doiron, 2012; van Vreeswijk and Sompolinsky, 1998]. Membrane time constants for excitatory and inhibitory neurons were $\tau_m = 15$ ms and $\tau_m = 10$ ms, respectively, and the refractory period was 5 ms for both excitatory and inhibitory neurons. The synaptic input from the network was given as:

$$I_{\mathrm{S}} = \sum_{i \leftarrow j} w_{i \leftarrow j} g_j^{E/I}(t), \tag{2.8}$$

where the $i \leftarrow j$ denotes that there is connection from neuron j to neuron i, and $w_{i \leftarrow j}$ denotes the weight of this connection (see next section for the weight settings). The synaptic inputs $g^{E/I}$ were increased step-wise instantaneously after a presynaptic spike ($g^{E/I} \rightarrow g^{E/I} + 1$)

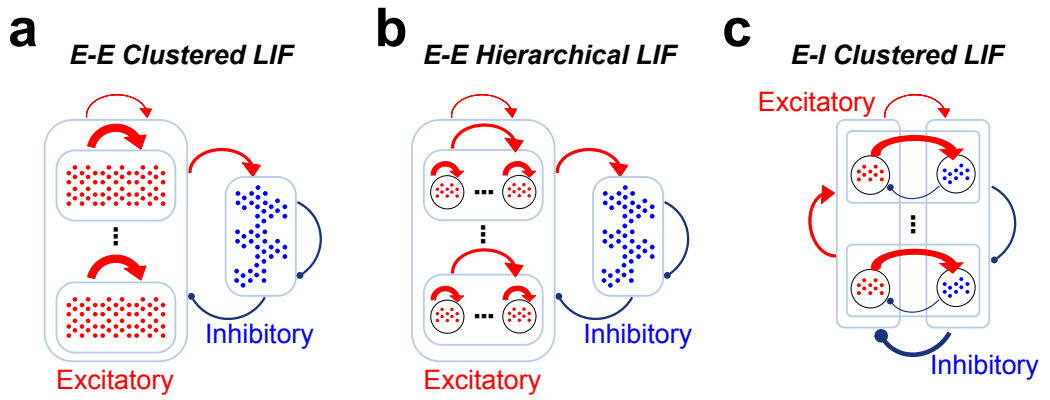

Figure 2.3: Schematic wiring diagrams of the three LIF networks used in this work: **(a)** an E-E clustered LIF network; **(b)** an E-E hierarchical LIF network; **(c)** an E-I clustered LIF network. Arrow thickness is proportional to the strength of the connection. For the parameters used in our simulations, see Table 2.1.

and then decayed exponentially according to:

$$\tau_{E/I} \frac{dg^{E/I}}{dt} = -g^{E/I}(t), \tag{2.9}$$

with time constants $\tau_E = 3$ ms for an excitatory interaction, and $\tau_I = 2$ ms if the presynaptic unit was inhibitory.

2.3.4.2 Network Topologies and Weight Matrices

LIF excitatory and inhibitory units in a proportion of $4:1$ were interconnected with three different network topologies. The resulting networks were simulated with a 0.1 ms time step. The connection probabilities and weights between the different types of neurons for these three LIF networks are shown in Table 2.1 and the schematic of the different wiring diagrams is shown in Fig. 2.3.

Network with clustered excitatory connections (E-E clustered) We first constructed a LIF network with clustered excitatory units: each excitatory neuron belongs to a group of units more strongly connected to each other than to units outside the group (Figure 2.3a). The network also included unclustered inhibitory units, which ensured that the network was balanced. These networks display temporally-structured spike-train activity [Litwin-Kumar and Doiron, 2012], and are used here as a test-bed for cell-assembly detection from spiking dynamics.

Network with hierarchical excitatory connections (E-E hierarchical) In a similar fashion, we developed a LIF network where excitatory units belonged to a hierarchy of groups

	Probabilities							
	p^{II}	p^{IE}	p^{EI}	p^{EE}	p^{EE}_{sub}	$p^{EE}_{\text{sub,sub}}$	p^{EI}_{sub}	p^{IE}_{sub}
E-E Clustered	0.5	0.5	0.5	0.167	0.5	—	—	—
E-E Hierarchical	0.5	0.5	0.5	0.15	0.3	0.99	—	—
E-I Clustered	0.5	0.454	0.526	0.2	—	—	0.263	0.90
	Weights							
	w^{II}	w^{IE}	w^{EI}	w^{EE}	w^{EE}_{sub}	$w^{EE}_{\text{sub,sub}}$	w^{EI}_{sub}	w^{IE}_{sub}
E-E Clustered	−0.04	0.01	−0.025	0.012	0.0144	—	—	—
E-E Hierarchical	−0.04	0.01	−0.03	0.012	0.012	0.014	—	—
E-I Clustered	−0.04	0.0086	−0.032	0.0155	—	—	−0.0123	0.0224

Table 2.1: Parameters for the simulated LIF networks. Connection probabilities (p^{XY}) and weights (w^{XY}) between different unit types: excitatory (E) and inhibitory (I), *e.g.*, p^{EI} is the connection probability from inhibitory to excitatory units. For the clustered networks, the average E-E connection probability was kept constant at 0.2. For a schematic representation of the wiring diagrams, see Fig. 2.3.

(Figure 2.3b). For this, we split the population of excitatory units into nested clusters, such that each group was sub-divided into smaller groups with increasing internal connectivity. The inhibitory neurons remained unclustered.

Network with excitation to inhibitory clustered feedback loops (E-I clustered) Finally, we have developed a LIF network to study the dynamical spiking patterns originated by networks in which excitatory and inhibitory neurons are co-clustered, as shown in Figure 2.3c. In this case, whereas the excitatory-to-excitatory and inhibitory-to-inhibitory couplings were kept uniform, we introduced structural features in the connections between distinct neuron types. In particular, each subset of excitatory units was more strongly connected to a subset of inhibitory units. This group of inhibitory units, in turn, had a weaker feedback to its associated excitatory neuron group, as compared to the rest of the graph. Every unit was part of one such functional group comprising both excitatory and inhibitory units.

2.3.5 Experimental data

2.3.5.1 Retinal Ganglion Cell recordings from mouse and salamander

These datasets were kindly provided by the lab of Markus Meister. Multielectrode recordings were performed as described previously [Meister et al., 1994], following protocols approved by the Institutional Animal Care and Use Committee at Harvard University and at the California Institute of Technology. Dark-adapted retina isolated from a larval tiger salamander (*Ambystoma tigrinum*) or adult mouse (*Mus musculus*; C57BL/6) was placed on a flat array of 61 extracellular electrodes with the ganglion cell side down. The salamander retina was superfused with oxygenated Ringer's medium (in mM: NaCl, 110; $NaHCO_3$, 22;

KCl, 2.5; $MgCl_2$, 1.6; $CaCl_2$, 1; and D-glucose, 10; equilibrated with 95% O_2 and 5% CO_2 gas) at room temperature. The mouse retina was perfused with oxygenated Ame's medium (Sigma-Aldrich; A1420) at 37°C.

Recordings were made with a custom-made amplifier and sampled at 10 kHz. Spike sorting was performed offline by analyzing the shape of action potentials on different electrodes [Gollisch and Meister, 2008; Pouzat et al., 2002]. The spike-triggered averages (STAs) and receptive fields of the salamander retinal ganglion cells (RGCs) were determined by reverse correlation to a checkerboard stimulus flickering with intensities drawn from a normal distribution. Singular-value decomposition of the spatio-temporal receptive field allowed the extraction of the temporal filter of every RGC receptive field [Gollisch and Meister, 2008].

2.3.5.2 Hippocampal CA1 and CA3 recordings from rats under a spatiotemporal task

We analyzed spike trains obtained by Diba and Buzsaki [2007] from hippocampal neurons of rats moving along a linear track implanted with silicon probe electrodes along CA1 and CA3 pyramidal cell layers in left dorsal hippocampus.

2.3.6 Performance of the method and comparisons to other techniques

In those examples where the results could be compared against a ground truth, the performance of the method was determined by the percentage of correctly classified neurons (hit rate) relative to the true membership in the data.

We have compared the performance of our methodology with two other popular community detection techniques: modularity optimization (two variants) using the code provided and explained in [Humphries, 2011], and standard agglomerative hierarchical clustering using the nearest distance linkage criterion as implemented in MATLAB.

2.4 Results

2.4.1 Assessing the algorithm with synthetic datasets

We first tested our method on synthetic spike-train datasets to evaluate its performance and to showcase its distinct capabilities compared to other methodologies.

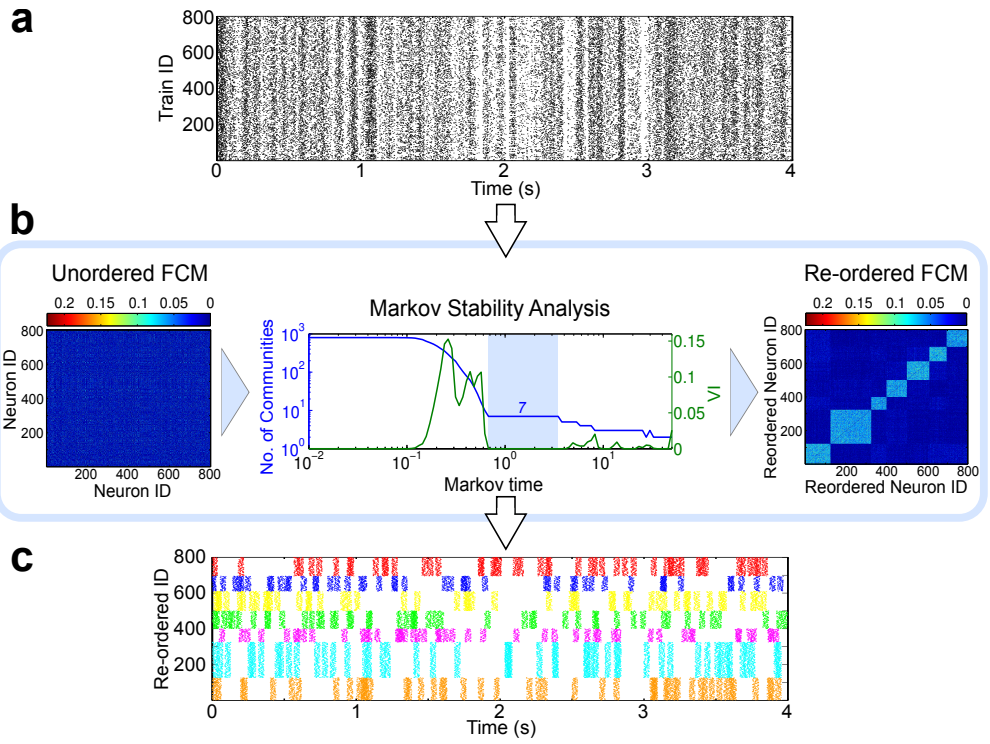

Figure 2.4: Markov Stability analysis of a synthetic data set. **(a)** Unsorted raster plot of a population of 800 spike-trains obtained from 7 groups of different sizes. Each 'cell assembly' fires at different times with varying amounts of jitter. **(b)** FCM from the unsorted spike train rastergram, followed by the Markov Stability plot and the FCM reordered according to the partition into 7 groups obtained by the algorithm. Note the long plateau (blue shaded) around Markov time $t_M = 1$ with $VI = 0$, indicating the presence of a robust partition with 7 groups. At later Markov times, the algorithm detects other robust coarser partitions corresponding to aggregates of the seven groups with similar firing patterns. **(c)** Color-coded raster plot reordered according to the partition obtained in **(b)**.

2.4.1.1 Analysis of synthetic data with embedded cell-assemblies

As a first illustration, Figure 2.4 shows the application of our method to a synthetic spike dataset with inherent group structure (see Materials and Methods). A population of 800 units was divided into 7 differently sized groups comprising 75 to 200 units. The average spiking frequency for all groups was 12 Hz with ±20 ms jitter around the uniformly chosen firing times within the total length of 4 s.

Figure 2.4a displays the raster data prior to clustering. Figure 2.4b shows the functional connectivity matrix (FCM) calculated from the spike trains used in the Markov Stability analysis, leading to the identification of a robust seven-community partition and a reordered FCM. The raster plot reordered according to the communities identified by our algorithm is shown in Figure 2.4c. The detected partition into 7 groups corresponds to an extended plateau in Markov time (from $t_M = 0.68$ to $t_M = 3.47$) with $VI = 0$, in which all the

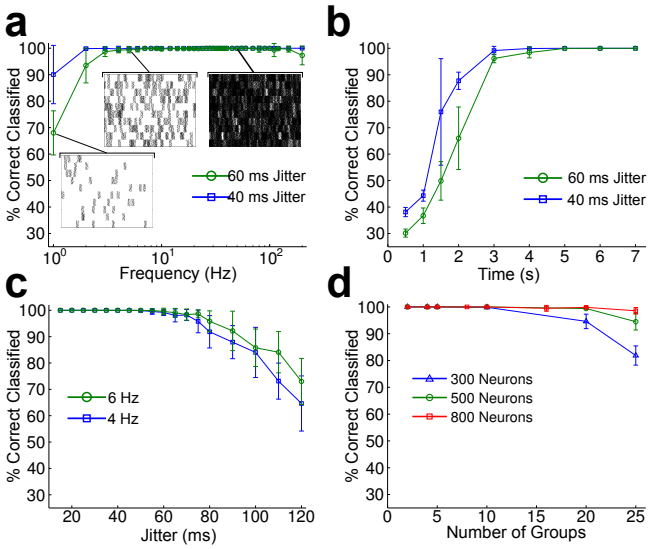

Figure 2.5: Assessing the performance of the clustering algorithm using synthetic data. **(a)** At very low firing frequencies, the classification performance is low due to a small number of spikes per neuron. Performance quickly improves with increasing firing frequency. **(b)** The classification performance improves as the duration of the recording increases. **(c)** As the jitter increases, the classification performance degrades. **(d)** The classification performance degrades mildly as the number of groups to be detected increases.

neurons were correctly clustered. Note that the algorithm detects other partitions with relatively long plateaux in Markov time, although their variation of information is non-zero. In particular, a relatively robust partition into three clusters between $t_M = 8.21$ and $t_M = 25.12$ is detected corresponding to a coarser grouping of the seven groups embedded in our data.

To assess the performance and robustness of the procedure, we determined the percentage of correctly classified neurons under a variety of noise conditions, different amounts of data, and other sources of variability. Figures 2.5a,b demonstrate the accuracy of classification for 500 units with 10 cell assemblies when the number of observed spiking events is varied. In Figure 2.5a, spiking datasets of a fixed length of 4 s are analyzed as the firing frequency is increased. As expected, the performance degrades at very low and very high spike frequencies when the number of firing events is either too low or too large to distinguish the groupings (see insets). This effect can be reduced with increasing spike-train lengths, as shown in Figure 2.5b where the performance improves as we increase the duration of the recording for a fixed firing rate (4 Hz). For short recordings, spurious correlations in the firing events degrade the performance, which consistently improves as the duration increases, reaching 100% accuracy for recordings of length above 4 s. To assess the effect of jitter, we checked that the above examples (Figure 2.5a-b) show similar behavior for different amounts of jitter ($\pm$40 ms, $\pm$60 ms). The performance degrades only when the jitter is

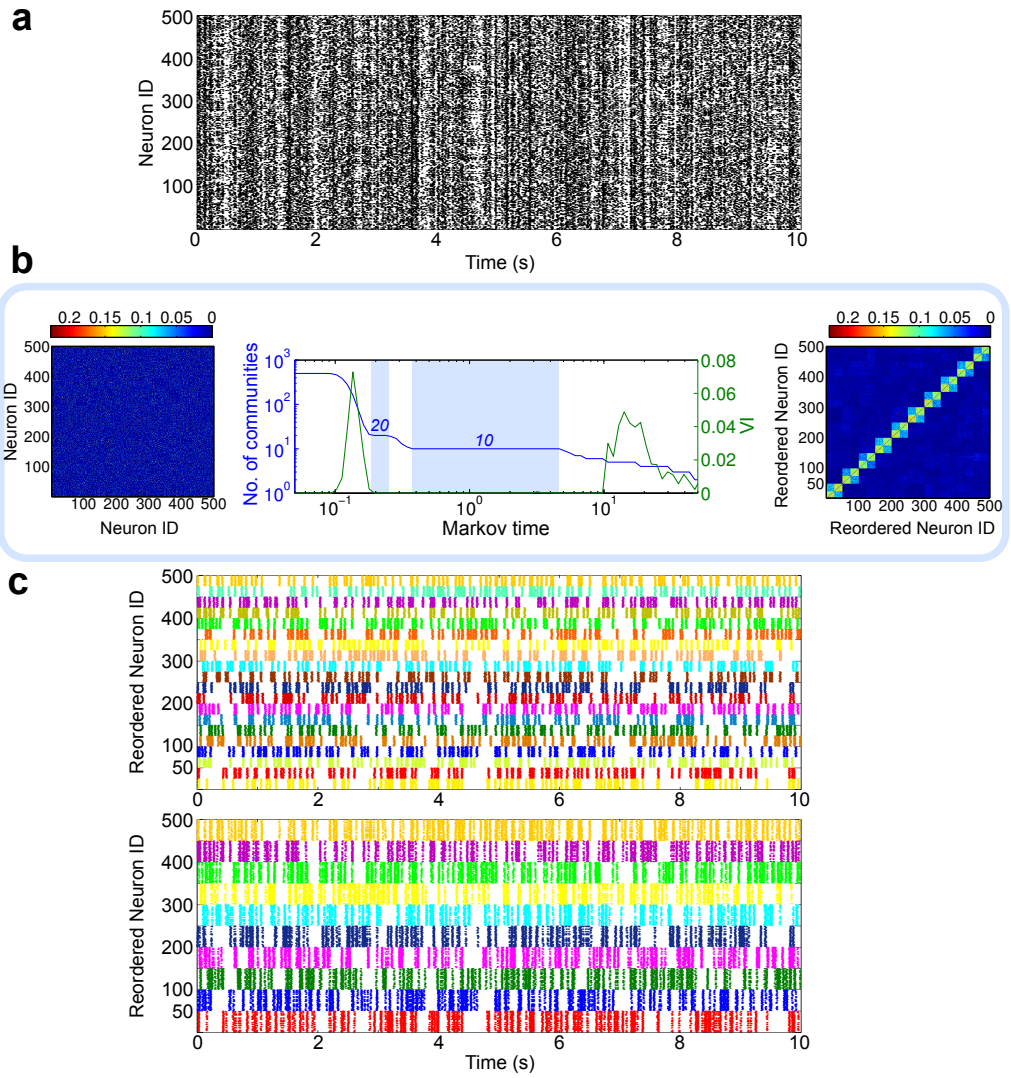

Figure 2.6: Detecting hierarchically structured spike train communities in synthetically generated data. Synthetic data of 500 units clustered into 10 groups with 2 subgroups each (20 subgroups in total). **(a)** Unsorted raster plot of data. **(b)** Markov Stability analysis of the associated FCM. Clear plateaus indicate the presence of robust partitions into 20 and 10 communities, with classification accuracy of 100% in both cases. **(c)** Sorted raster plots for the finer (20 groups, top panel) and coarser (10 groups, bottom panel) partitions revealing the hierarchical organization in the data.

increased strongly (Figure 2.5c). Finally, we assessed the sensitivity of the method and its ability to detect an increasing number of groups in a population of given size (Figure 2.5d). As expected, the accuracy of the classification drops, but only mildly, as we increase the number of groups to be detected. Note that this is due partly to an entropic effect: a correct assignment among a larger number of groups conveys more information than a correct decision between fewer groups. Hence, the decrease in performance is even less dramatic if corrected for this effect.

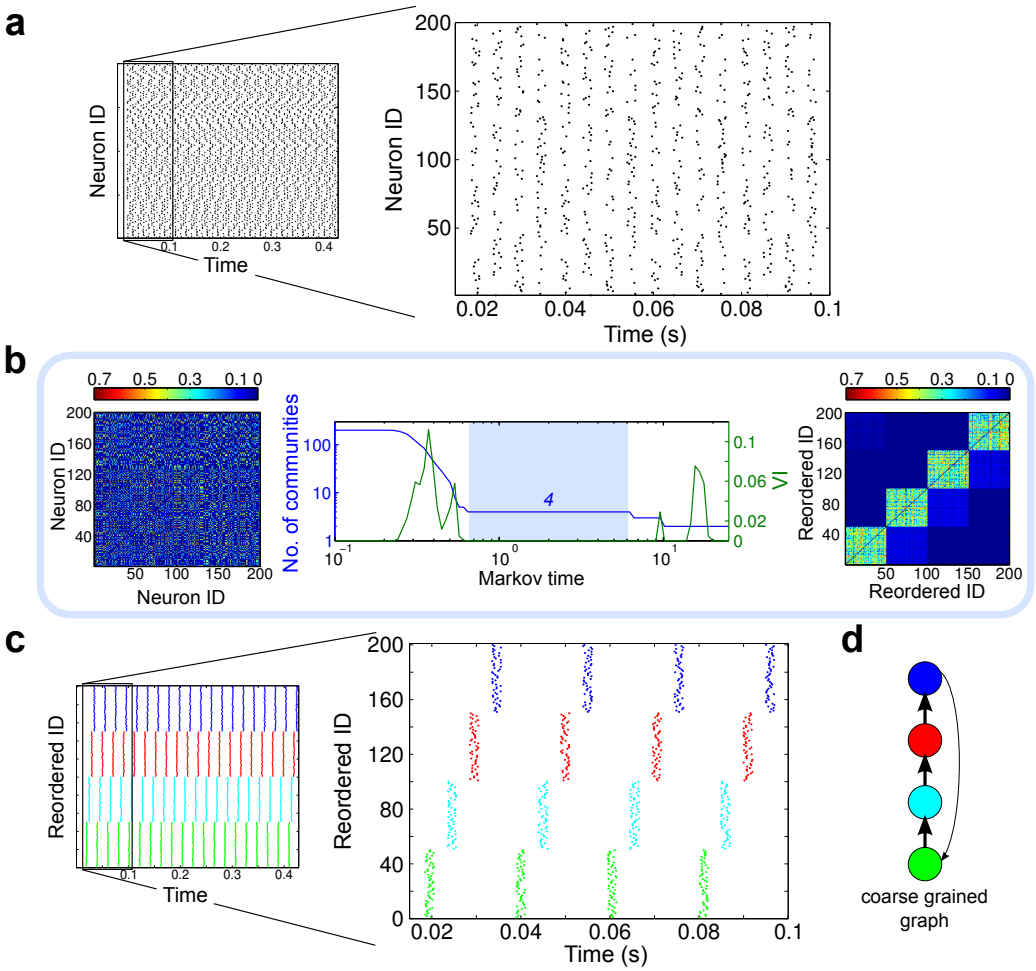

Figure 2.7: Analysis of feedforward-like firing patterns. **(a)** Unsorted raster plot of the synthetic data and zoom-in. **(b)** Markov Stability analysis of the FCM identifies a robust partition into 4 groups, with 100% classification accuracy. Note that the FCM is asymmetric, thus revealing the directionality of the data. **(c)** Color-coded raster plot and zoom-in color-coded according to the partition found reveals the feed-forward functional relationship in the data. **(d)** Coarse-grained representation of the functional connectivity network found from the clustering. For a second example, see supplementary information.

2.4.1.2 Analysis of synthetic hierarchical spiking patterns

Hierarchical neuronal connectivity [Ambrosingerson et al., 1990; McGinley and Westbrook, 2013; Savic et al., 2000] can lead to spiking dynamics with temporal structure at different scales. One advantage of using Markov Stability is its ability to detect hierarchical structure in data without *a priori* knowledge of such relations. To showcase this capability, we created synthetic data sets with embedded hierarchical relationships (see Materials and Methods). Figure 2.6 illustrates the analysis of the spiking dynamics from 500 units, which are split into 10 groups of co-firing units with each group further sub-divided into two subgroups that fire together more frequently. This results in a hierarchical organization of $20 \rightarrow 10$ subgroups.

Using the same methodology as above, our analysis reveals two extended plateaus with $VI = 0$, for 20 and 10 groups (Figure 2.6b). The sorted raster plots for the 20 and 10 groups, shown in Figures 2.6c, correspond to 100% correct classification. As we will demonstrate below in the context of LIF networks, this consistent multi-scale detection of cell assemblies is a distinct feature of our methodology, which is not present in many other methods which only detect groupings at a particular level of granularity [Fortunato, 2010].

2.4.1.3 Analysis of synthetic feedforward spiking patterns

To highlight the capability of our framework to deal with directed dynamical patterns, we show how feedforward-like functional patterns in the data lead to a pronouncedly asymmetric FCM, which can then be analyzed with Markov Stability. Synthetic spiking patterns were generated in which four groups of 50 neurons (with jitter) spiked 20 times, emulating synchronous activity in feedforward networks (see Materials and Methods). As shown in Figure 2.7, our method is able to detect feedforward patterns between cell assemblies: the corresponding Markov Stability plot shows a robust and extended plateau with four communities with 100% classification accuracy revealing an effective coarse-grained description of a functional feedforward network.

This is an instance in which the directed nature of our FCM, together with the fact that Markov Stability can detect communities in directed networks, leads to the detection of cell assemblies with directed, causal relationships. Indeed, there are instances of directed functional couplings [Rosvall and Bergstrom, 2008] in which using symmetric measures will lead to different cell assemblies to those obtained if directionality is taken into account. Hence for some networks, directionality is absolutely essential for proper clustering as we consider with a second example.

Figure 2.8a shows the wiring diagram of a network which, if analyzed with a symmetric similarity measure, lead to different cell assemblies that miss the causality/direction of the connections. This network may produce firing patterns as shown in Fig. 2.8b.

There are two possible outcomes when analyzing the raster plot

- When this spike-trains are analyzed using a symmetric measure, leading to a symmetric FCM, we get a partition into two groups: $\{1, 4\}$ and $\{2, 3\}$, which reflect the strength of the connections but not the causality of the dynamics (see Figure 2.8c).

- When the spike-trains are analyzed using the true, directed FCM, we find the assemblies based on flow, which illustrates the fact that clearly different clusterings can arise when taking into account directionality (see Figure 2.8d).

It is important to remark that only by using a directed FCM we are able to reveal both

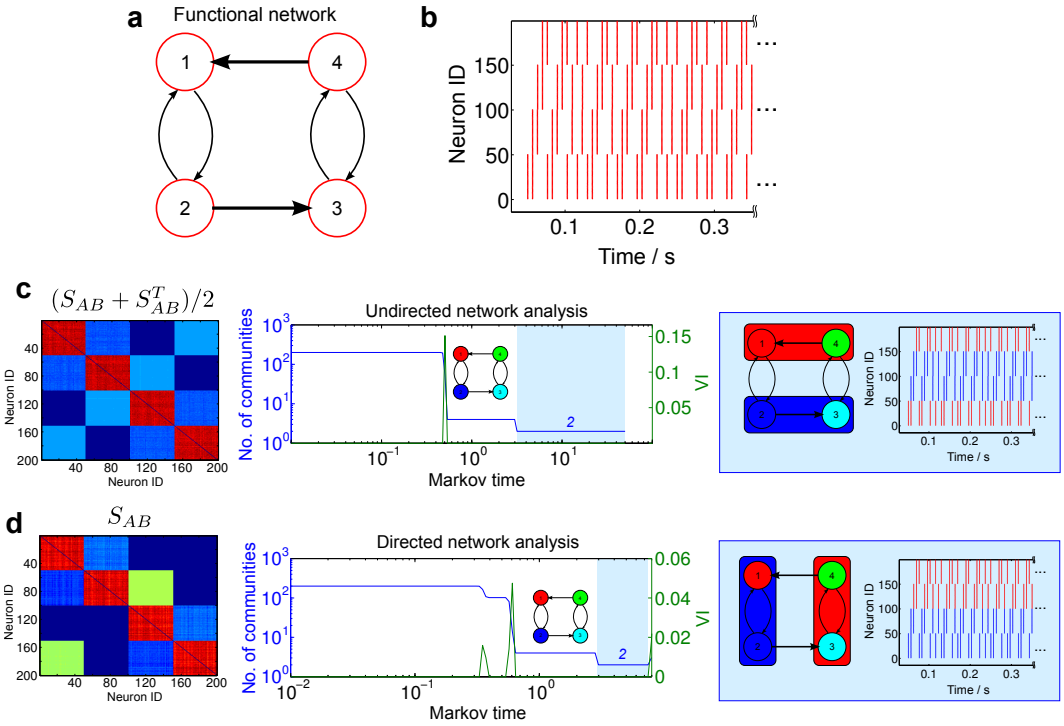

Figure 2.8: Comparison of how directionality can affect clustering results. **a-b** A functional network as depicted in **a** can be emulated by synthetic data by varying the group firing sequences and delay between individual group firings (**b**). Ignoring directionality (using the symmetrized FCM) leads to finding assemblies based on the strength between the groups (see **c**). Taking into account directionality reveals a grouping based on flow (see **d**).

cases: once we use a symmetric FCM the directions of the links are lost, and thus we can only recover the partition based on strength. Within our framework, the importance of directionality can be tested by including or disregarding directionality in the analysis and comparing the outcomes.

2.4.2 Detecting cell assemblies in simulated networks

Beyond purely synthetic datasets, we now consider three examples of simulated dynamics of LIF networks, which exhibit a range of features of relevance in realistic neural networks. LIF networks provide a simple, controlled testbed to assess our framework on network dynamics broadly used in computational systems neuroscience.

2.4.2.1 Cell assemblies in LIF networks with clustered excitatory connections

It was demonstrated recently that balanced LIF networks with clustered excitatory connections can display network dynamics in which the clustered neurons spike in a coordinated manner over long timescales [Litwin-Kumar and Doiron, 2012]. We implemented such a

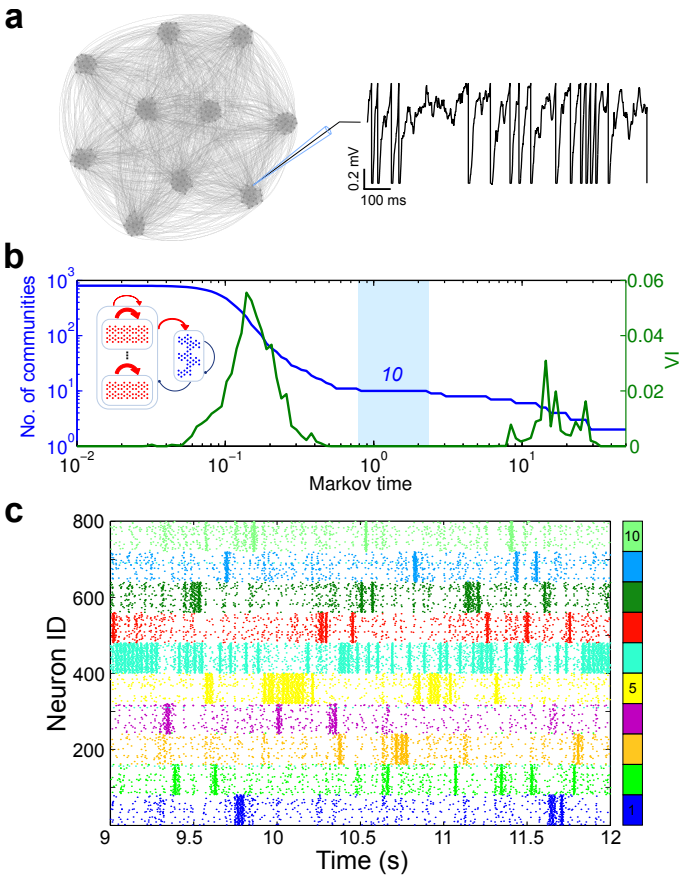

Figure 2.9: Detecting cell assemblies in spiking data from E-E clustered LIF networks. **a** Schematic of the excitatory connectivity of the LIF network. The 800 excitatory units were split into 10 groups such that the intra-group connection probability and synaptic strength were larger than the inter-group values. The network was balanced with 200 unclustered inhibitory units. An example of the simulated membrane potential traces for an excitatory unit is also shown. **b** Markov Stability analysis of the corresponding FCM. There is a clear plateau with $VI = 0$ for a split into 10 groups (blue shaded). Inset: Schematic of network topology. **c** Color coded raster plot according to the partition obtained. Units are ordered consecutively according to their grouping in the underlying LIF topology. The correct grouping is also indicated by the colored band on the side. These cell assemblies exhibit clear bands of activity. Only 5 neurons were misclassified relative to the imposed structure (99.5%).

LIF-network to determine if our framework was able to recover the underlying *structural* connectivity directly from the observed spiking dynamics.

Figure 2.9a depicts a schematic of the structural connectivity imposed on the simulated E-E clustered network: 800 excitatory units were split into 10 groups such that the connection probability and synaptic strengths within each group were larger than the inter-group values (see Materials and Methods). The network was balanced by 200 inhibitory units that were uniformly connected (*i.e.,* unstructured). The spiking dynamics of the excitatory units were then analyzed using our framework: the spike trains of the 800 excitatory units

were used to generate the FCM, and Markov Stability revealed a clear plateau with $VI = 0$ corresponding to a partition into 10 groups with a classification performance of 99.5% with respect to the embedded structural groups.

The results of our method contrast with other commonly used methods. For instance, applying hierarchical clustering to the FCM achieves only a classification performance of at best 22%, and does not provide a clear criterion for determining the number of groups present. Similarly, we apply Modularity optimization using the two versions implemented by [Humphries, 2011] in conjunction with our FCM matrix. In this case, Modularity imposes an intrinsic scale leading to the identification of 8-13 groups, with a classification performance of 49-68% depending on which of the two versions of the optimization is used. Let us remark that we also applied hierarchical clustering directly to the time-series. However, it leads to similarly poor performance.

As explained above, an advantage of Markov Stability is that it does not impose *a priori* the scale or the number of clusters to be detected. Instead, the method scans across all scales and extracts robust, meaningful partitions at different levels of granularity, thus revealing potentially relevant partitions in the data. In clear-cut cases, such as the simple synthetic datasets studied above, the method reveals unequivocally the partitions embedded in the data. In general, however, and especially for noisy data, it is not expected that a unique partition is found. Rather, a set of candidate partitions will emerge. This is observed in the analysis of this LIF network. Figure 2.9b shows that the partition into 10 communities is the clearest choice for this dataset (longest plateau in Markov time with $VI = 0$). However, other good candidate partitions include the following: one into 11 communities, which is similar to the grouping into 10 albeit with an additional split for a small group of neurons, and a partition into 8 communities obtained by the merging of 4 of the groups into 2 groups. Our methodology provides candidate partitions at different levels of resolution based on their robustness (i.e., long plateaux in Markov time, drops in VI) as a guide for the analysis, which can then be complemented with further biological knowledge.

2.4.2.2 Cell assemblies at multiple levels of granularity in hierarchical LIF networks

To further test the multiscale capabilities of our method, we evaluated how the algorithm would perform on the analysis of a LIF network with a hierarchical structure. The population of 800 excitatory units was divided into 10 groups with 2 subgroups each, *i.e.,* 20 groups in two levels of a hierarchy (Figure 2.10a). As before, the network was balanced with 200 unclustered inhibitory units. The Markov Stability analysis displayed in Figure 2.10 reveals two robust partitions: one into 21 communities and one into 10 communities

with classification accuracies of 99.9% and 100%, respectively, and which correspond to the two levels of resolution. Our method also finds other suitable but less robust candidate partitions, *e.g.*, one into 9 communities obtained by merging 2 of the top-level communities into a single group, or one into 4 groups obtained by similar mergers.

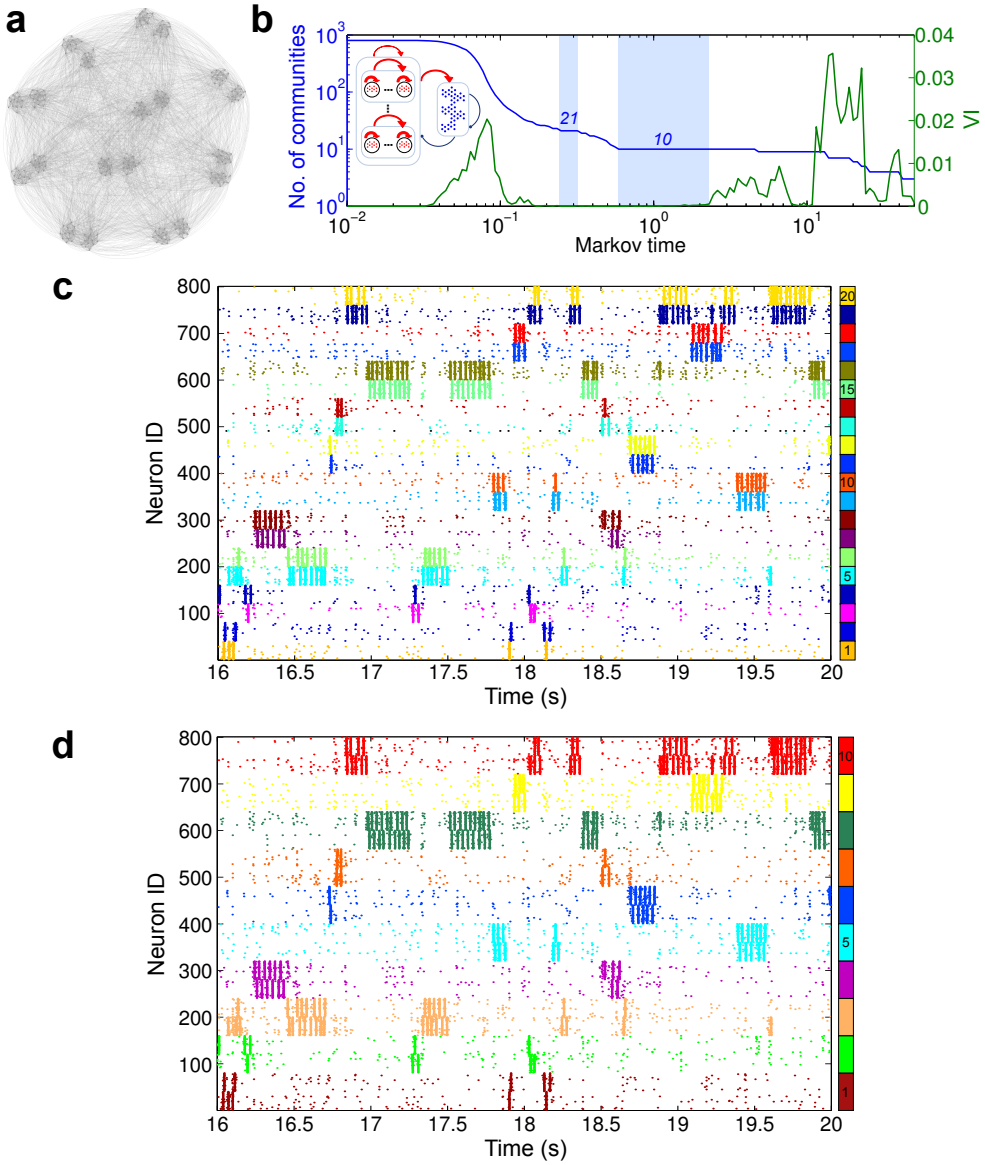

Figure 2.10: Detecting cell assemblies in spiking data from E-E hierarchically clustered LIF networks. **(a)** Schematic of the network with excitatory units split into 10 groups which were further sub-divided in 2 subgroups each. **(b)** Markov Stability plot of the analysis reveals two robust partitions with 21 and 10 communities. Inset: Schematic of network topology. **(c)** Color coded raster plot of the clustering into 21 communities with classification accuracy of 99.9% (one neuron was misclassified). Units are ordered consecutively according to their grouping in the underlying LIF topology. The correct grouping is also indicated by the colored band on the side. **(d)** Color coded raster plot of the 10 community clustering with an accuracy 100%.

As stated above, other commonly used methods are unable to detect these multiple levels of granularity. For instance, Modularity (using again the implementation of [Humphries, 2011]) finds a partition at one particular scale (10 communities, 41-80% accuracy for the two versions) and cannot detect the presence of the finer grouping. The application of hierarchical clustering leads to an agglomerative tree with no better accuracy than 41% at any level of granularity and no clear criterion to detect the number of communities present.

2.4.2.3 Mixed cell assemblies with excitatory and inhibitory units in LIF networks

Hitherto we have only considered the clustering of excitatory units. However, functional groups of neurons may contain both excitatory and inhibitory neurons [Buzsaki, 2010]. The definition of our spike-train similarity allows for the detection of such relationships by incorporating the biophysical effect of both excitatory and inhibitory neurons on their postsynaptic neurons (*i.e.,* EPSPs *vs.* IPSPs; see Figure 2.1A). To determine how our method would perform in a context where mixed functional groups are present, we created a LIF network with an embedded structure between excitatory and inhibitory units. As shown in Figure 2.6A, the coupling between alike neuron types is uniform but we create preferential coupling between subsets of excitatory and inhibitory neurons, *i.e.*, each subset of excitatory units is preferentially connected to a subset of inhibitory units (relative to all other inhibitory ones) and, in turn, this subset of inhibitory units feeds back weakly to their corresponding subset of excitatory units (relative to all other excitatory units). Our simulated LIF network included 10 such groups with 80 excitatory neurons and 20 inhibitory neurons per group for a total of 1000 neurons.

The analysis of the dynamics of this LIF network is presented in Figure 2.11. As indicated by our color coding, we find a robust partition into 10 communities that comprise a combination of both neuron types with a 91.4% of correctly classified cells according to the embedded structure. Having the ability to account for the role of the inhibitory neurons within a cell assembly may provide a key difference to find a meaningful interpretation of the data. In our LIF network simulations, we observed that not accounting for cell type differences may result in a drop of up to 20% in classification performance (data not shown).

2.4.3 Applying the algorithm to experimental data

2.4.3.1 Detecting distinct Retinal Ganglion Cells in mouse data

As a simple first check of our framework when applied to experimental data, we tested that our algorithm could detect distinct mouse retinal ganglion cells (RGCs) from spike-train

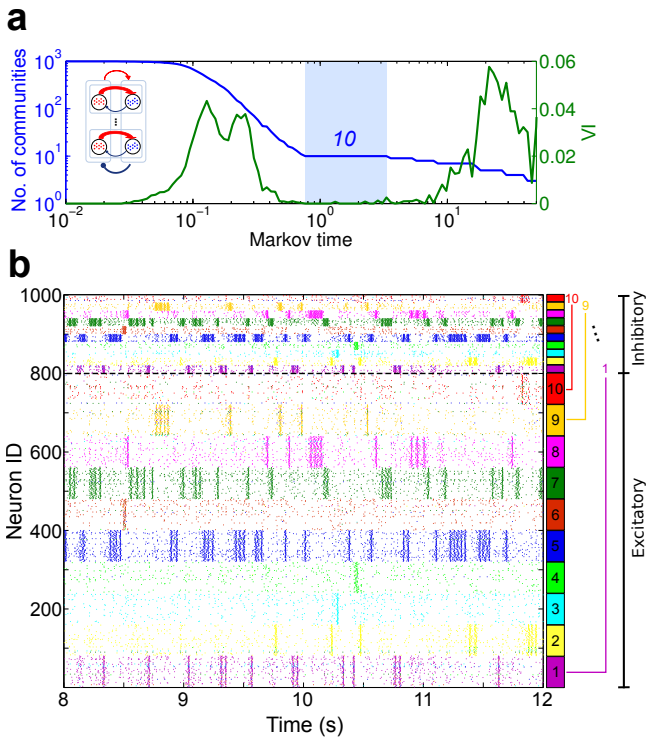

Figure 2.11: Detecting cell assemblies in spiking data from E-I clusterd LIF networks. The network has functional groups comprising excitatory and inhibitory neurons. **(a)** Stability plot of the clustering analysis. Note the clear plateau for 10 communities. Inset: Schematic of network topology. **(b)** Color-coded raster plot according to the obtained partition. Note that each group contains excitatory and inhibitory units, as indicated by the color-coded band on the right side which displays the true structural grouping. For a 20 s simulation, the classification rate was 91.4%.

data. Extracellular recordings of a flattened mouse retina were performed while a full field black and white flicker stimulus was repeated 52 times for approximately 28 s (see Materials and Methods). Data was collected from three different cells that were reliably identified and spiked consistently for every stimulus repetition. Spike-triggered averages (STAs) and spike-triggered covariances (STCs) [Schwartz et al., 2006] of the three cells were used to characterize the neurons as an ON-cell, an OFF-cell, and a noisy OFF-cell (i.e., an OFF-cell with high trial-to-trial variability).

To test our algorithm, the 52 repetitions from the three neurons were compiled and shuffled randomly into a composite raster plot. The algorithm was then applied to this raster plot so as to find relevant groupings in the 156 spike-trains. Our algorithm (Fig. 2.12) reveals a robust partition into three groups corresponding to spike-trains from each cell (98.7% correctly identified with their original cell). Interestingly, a further plateau at longer Markov times can be seen in Fig. 2.12A, corresponding to a 2-way partition, in which the two OFF cells are grouped together and the ON cell is separate. The performance accuracy

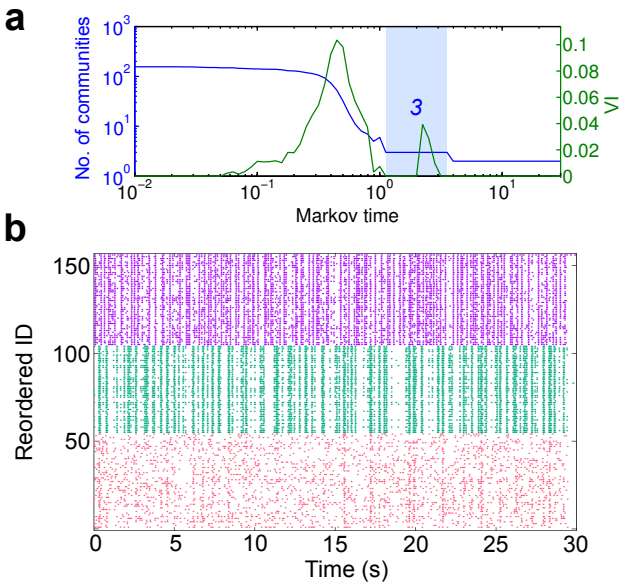

Figure 2.12: Detecting cells in a set of extracellular recordings from three mouse retinal ganglion cells (RGCs) to a full field stimulus (52 repetitions for 156 spike trains in total). **(a)** The Markov Stability plot obtained from all the spike trains reveals the presence of a robust partition into three groups. **(b)** Raster data color coded and reordered according to the communities found. The classification was 98.7% accurate.

is still 98.7 %.

Although used here as a check of our method on real data, this analysis also illustrates how our method is able to extract valuable information from the data at multiple levels of resolution directly from the spike-trains. In contrast, Modularity finds here only a partition into two groups with 60.3-98.1% correctly identified, a classification performance similar to hierarchical clustering (67%).

2.4.3.2 Detecting classes of Retinal Ganglion Cells in salamander data

Next we analyzed a dataset of extracellular simultaneous recordings from multiple RGCs from the salamander retina of three different animals exposed to the same stimulus. The applied stimulus (i.e., random flickering bars on a screen) entailed both time and spatial components. Approximately 50 neurons were recorded simultaneously from each animal for a total of 141 neurons which were combined into a single raster plot.

Upon application of our algorithm to the spike-train data (Figure 2.13), a robust partition into two groups was observed. To check if this grouping was meaningful, we studied *a posteriori* the STAs of the recorded neurons, which characterized them as mixed groups of ON and OFF cells, yet with different temporal characteristics. Therefore the two communities found do not correspond to a pure separation into ON and OFF cells but rather to transient and sustained RGC populations responding to fluctuating light intensities on differ-

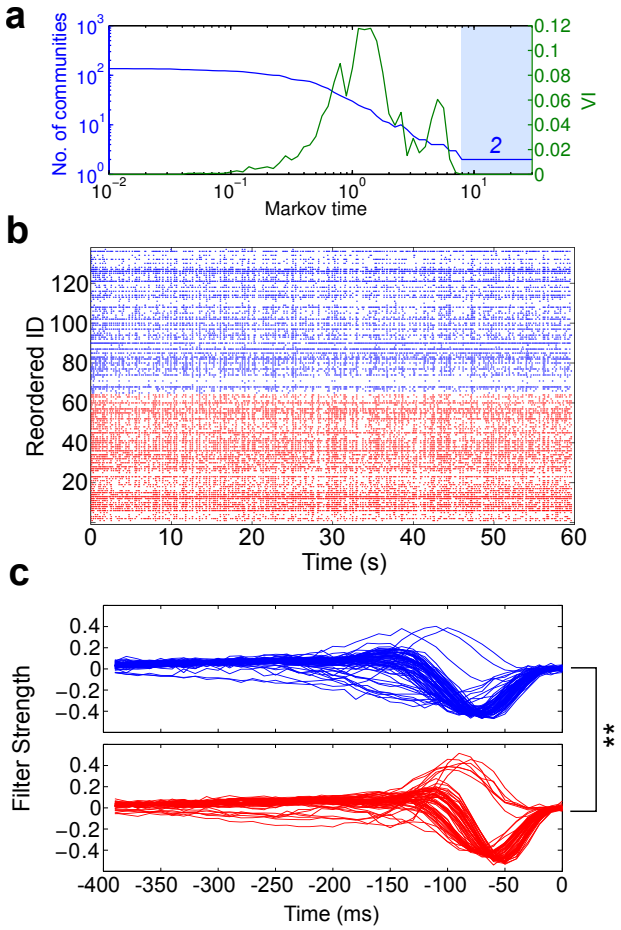

Figure 2.13: Detecting cell assemblies in spike train recordings from salamander retina RGCs (141 neurons simultaneously recorded with extracellular electrodes). **(a)** The Markov Stability plot obtained from the recordings reveals the presence of a robust partition into two groups. **(b)** Raster data color coded and reordered according to the communities obtained. **(c)** Spike triggered average (STA) responses from all neurons recorded. Each line is an STA for a different cell. The two panels correspond to the two communities (colored as in **(b)**) and correspond to transient and sustained RGCs. The full width half maximum of the sustained RGCs (upper panel) is 112.2 ± 9.1 ms while for the transient RGCs (lower panel) is 83.0 ± 6.4 ms (mean $\pm$ sem). The difference between the temporal characteristics of the two panels is statistically significant. ($p < 0.01$; see text for details).

ent time-scales [Awatramani and Slaughter, 2000]. Note the stimulus was random flickering *bars* and hence the ON and OFF cells can receive uncorrelated stimuli and therefore the algorithm could not pick up this most distinct difference as in the previous example. Nevertheless, independent statistical confirmation was obtained by checking that the distribution of the full-width-half-maxima of both populations was significantly different between groups (Kolmogorov-Smirnov test to check normality, Wilcoxon rank sum test with $p < 0.01$).

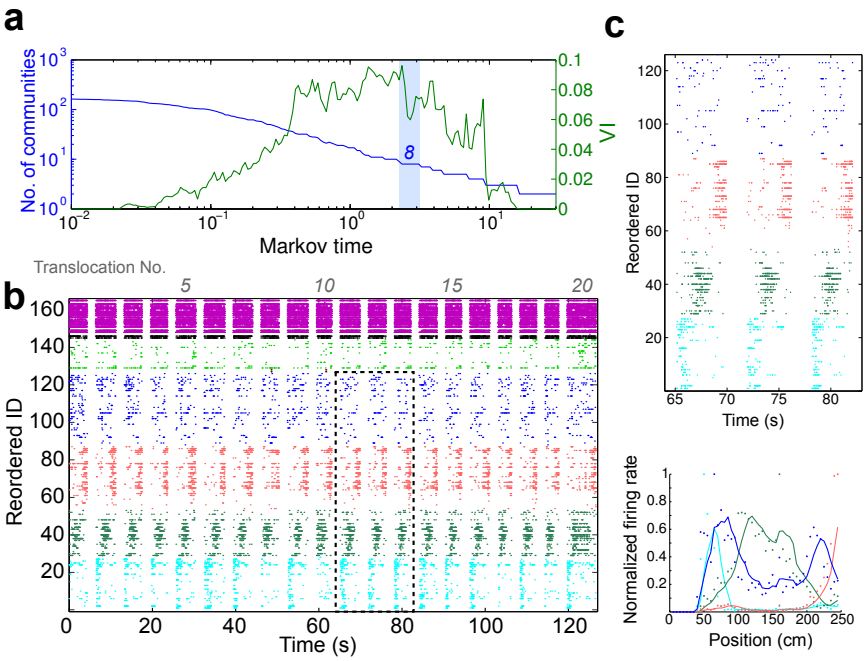

Figure 2.14: Detection of cell assemblies in recordings from rat hippocampal neurons transversing along a linear track. **(a)** The Markov Stability plot from the spike-trains shows a small plateau with dip in VI corresponding to a partition into 8 communities. **(b)** The raster data clustered according to the 8 communities reveals two groups containing the inhibitory neurons (purple and black), and 4 communities with structured time firings (cyan, dark green, red, blue) containing putative place field neuronal groups. These four communities contained 100.0% of the place cells identified by an independent identification method. The columnar gaps separate different translocation events. **(c)** Blow-up of the raster plot corresponding to the four place cell communities (top panel) and normalized firing rate of these four groups as a function of distance along the linear track. Each group favors firing at specific positions along the track.

2.4.4 Detecting Hippocampal Place Cells in rat recordings

Finally, the algorithm was applied to CA1 and CA3 hippocampal recordings from a rat moving in a linear track for a water reward [Diba and Buzsaki, 2007]. The analyzed data contains 165 neurons which were recorded simultaneously during a series of translocations in which the rat was always traveling in one direction. The spike trains of the twenty translocations were then spliced together into a raster plot.

Figure 2.14 presents the results of our analysis. The FCMs calculated for each translocation were averaged to obtain the FCM. The Markov Stability analysis finds a stable bi-partition, yet one of the groups comprised only two neurons of no apparent biological relevance, and this partition was not considered further.

An additional long plateau was marked by a dip of VI at $t_M = 2.7$ corresponding to a partition into 8 communities. We found that two of those communities contain all 18 inhibitory neurons (as identified via spike-sorting in [Diba and Buzsaki, 2007]): the purple

group (16 inhibitory neurons) and the black group with the remaining two. In addition, four of the other communities exhibit structured firing at particular times for every translocation: the first group (cyan) is active at the start of the translocation followed by the dark green group and the red group, while the dark blue group, though less salient, corresponds to cells with firings in between those groups. Such groups with localized firing patterns are good candidates to include place cells so we checked *a posteriori* the normalized firing rate of these communities as a function of position. Figure 2.14C shows that the assemblies found spike at different positions along the linear track, indicating that place cells are being identified. To validate our results, we compared to the results obtained with a place cell detection technique and found that these four cell populations account for 100.0% of the place cells [Taxidis et al., 2015]. Notably, our method only used the spike-trains to detect these cells and was able to also extract the inhibitory neurons.

2.5 Discussion

We introduce here a versatile technique to detect cell assemblies directly from spike-train data. The method uses biophysically inspired notions to create a functional connectivity matrix reflecting neuron-to-neuron relationships extracted from the observed spike dynamics. Groups of neurons are obtained from the functional connectivity matrix using a graph theoretical method for community detection, which scans partitions across all resolutions and detects relevant groupings without prescribing the level of granularity or the number and size of groups *a priori*. In contrast to most other methods, our technique is able to extract hierarchical structure in recorded data; incorporates functional differences between excitatory and inhibitory neurons; and can detect the directionality emanating from feedforward connectivity. All these are vital aspects enabling novel types of analyses in recorded datasets. As the method relies only on spike timings, it can be applied to both electrophysiological and optical recordings. We tested the performance of the method on a variety of synthetic data, where we showcased its ability to extract clustered and hierarchical assemblies, in contrast to standard methodologies, such as hierarchical clustering or Modularity which have an inherently lower performance in finding such hierarchical ensembles. We further confirmed that the directed nature of our technique allows the inference of feedforward connectivity, minimizing information loss when going from spiking data with relevant temporal ordering to a FCM. This may open the possibility to the inference of underlying anatomical connectivities, as well as gaining insight about feedforward connections from recorded neuronal network datasets.

We applied the framework to the analysis of spike-train simulations from several LIF

network topologies (E-E clustered, E-E hierarchical and E-I clustered organizations), which result in temporally-structured network activity. Our technique was able to identify the hierarchical structure in such simulated data without the need to rerun the analysis adjusting the parameter settings of the spike-train similarity measure. This capability could be of interest to detect hierarchical neuronal connectivity in real systems [Ambrosingerson et al., 1990; McGinley and Westbrook, 2013; Savic et al., 2000], as our integrated approach can identify structure at different scales without imposing strong assumptions *a priori*. In addition, the method was able to detect clusters that included both excitatory and inhibitory neurons in LIF networks. Taking into account the functional differences of these neuron types is a distinctive feature of our methodology. If all neurons are treated equally, irrespective of their cell type, as is commonly done, a simple split between excitatory and inhibitory neurons is often observed. This effect is essentially due to the strong difference in firing statistics between excitatory and inhibitory neurons, although inhibitory neurons may also exhibit a range of firing-frequencies [Isaacson and Scanziani, 2011; Markram et al., 2004]. Being able to extract such structural information, and to distinguish between excitatory and inhibitory interactions where applicable, is of interest for the understanding of the functional role of cell assemblies, as these groups are likely to include both excitatory and inhibitory neurons.

We additionally showed how our framework is able to recover biological information in three sets of experimental data from mouse RGCs, salamander RGCs, and rat hippocampal data, highlighting the versatility of the method. In the mouse RGC data, the method can assign repeated trials of the same stimulus recorded from three different neurons to the three cells without *a priori* information. Salamander RGCs recorded simultaneously in response to a stimulus of randomly flickering bars were clustered into two groups displaying distinct temporal characteristics in their responses, *i.e.*, transient and sustained RGCs could be distinguished. In the hippocampal recordings of a rat translocating along a linear track, the method was able to identify all the inhibitory neurons as well as four assemblies containing all the place cells associated with specific spatial information. An advantage of our technique in this context is that it only requires spike-timing data and can thus provide complementary information and cross-validation for other techniques currently used for place cell detection. Furthermore, the algorithm could be optimized for place cell detection to carry out additional analyses, *e.g.*, examining the spiking data of rats asleep after translocation sessions; studying the effects of varying the track length; the conjunct analysis of group firings with the recorded local field potentials; or the analysis of spatially-induced firing patterns of inhibitory neurons, among others.

While previous clustering methodologies have shown good results for particular applica-

tions, we aimed here for a method with the versatility to account for the wide differences in neuronal data while simultaneously providing a simple, interpretable approach to detect cell assemblies from spiking data. Hence our main focus was on the conceptual and generic aspects of our dynamics-based framework, rather than on fine-tuning the technical details towards more specialized applications. Further refinements of the method are possible (or indeed desirable for specific datasets) and could lead to improved performance. For instance, future work could be aimed at a more explicit use of neuron firing statistics, including varying firing frequencies such as during burst periods, or at the construction of functional coupling measures with more specialized biophysical couplings, including different neuron types. Other possible extensions of the FCM similarity measure include enforcing sparsity constraints on the couplings, possibly coupled with more refined statistical assessments of the importance of individual couplings. A particularly interesting question for future work will be to consider how cell assemblies (and underlying neuronal networks) can change their group structure over time via different mechanisms such as synaptic plasticity, and how this relates to learning. In addition, it will be interesting to study the relationship between the neuronal time scales and the Markov times that appear as optimal in our community detection.

In passing we note that in order to distinguish neuron types, all recorded neurons should ideally receive the same inputs. For example, in visual experiments a full-field spatially homogeneous stimulus should be strongly preferred over a noisy checker stimulus, as otherwise neurons of the same type may have uncorrelated firing due to uncorrelated inputs. This point is applicable to all clustering algorithms and not just the one presented here. A further possible application is clustering resting state activity recordings that assume there are no (time varying) stimuli that are driving neurons.

As modern neuronal recording techniques enable simultaneous recordings of ever increasing numbers of neurons, approaching nearly entire brains [Ahrens et al., 2013], techniques for detecting spike-train communities, such as the one proposed here, will become a vital tool to provide insight into such complex data. Collecting meaningful data from systems neuroscience experiments is the key requirement, yet being able to provide concise, intelligible representations of these recordings is just as critical in order to identify and comprehend the spatio-temporal information encoded in the data.

Chapter 3

Emergence of slow-switching assemblies in structured neuronal networks

Some or all of the work presented in this chapter has been published [Schaub et al., 2015].

3.1 Abstract

Unraveling the interplay between connectivity and spatio-temporal dynamics in neuronal networks is a key step to advance our understanding of neuronal information processing. Here we investigate how particular features of network connectivity underpin the propensity of neural networks to generate slow-switching assembly (SSA) dynamics, i.e., sustained epochs of increased firing within assemblies of neurons which transition slowly between different assemblies throughout the network. We show that the emergence of SSA activity is linked to spectral properties of the asymmetric synaptic weight matrix. In particular, the leading eigenvalues that dictate the slow dynamics exhibit a gap with respect to the bulk of the spectrum, and the associated Schur vectors exhibit a measure of block-localization on groups of neurons, thus resulting in coherent dynamical activity on those groups. Through simple rate models, we gain analytical understanding of the origin and importance of the spectral gap, and use these insights to develop new network topologies with alternative connectivity paradigms which also display SSA activity. Specifically, SSA dynamics involving excitatory and inhibitory neurons can be achieved by modifying the connectivity patterns between both types of neurons. We also show that SSA activity can occur at multiple timescales reflecting a hierarchy in the connectivity and demonstrate the emergence of SSA

in small-world like networks. Our work provides a step towards understanding how network structure (uncovered through advancements in neuroanatomy and connectomics) can have an impact on spatio-temporal neural activity and constrain the resulting dynamics.

3.2 Introduction

Neuronal ensembles exhibit a broad repertoire of activity patterns. Such dynamics are governed by a time-evolving network of synaptic connections with an intricate, yet structured, organization. Due to the advancement of connectomics, our knowledge about such networks is rapidly growing, and increasingly detailed maps of neuronal wiring are becoming available. In parallel, modern recording techniques, such as calcium imaging and multi-electrode arrays, allow neuroscientists to monitor the activity from thousands of neurons simultaneously, with recordings from entire brains at single neuron resolution becoming technologically feasible [Ahrens et al., 2013; Buzsaki, 2004; Du et al., 2011]. The observed dynamics of neural networks exhibit an interplay of structured spatio-temporal scales, which underpin a wide range of cognitive functions [Buzsaki, 2010].

The idea that neuronal group activity induced by network structure is at the core of neural computation dates back at least to the work of Hebb [Hebb, 1949], who hypothesized that the transient activity of groups of neurons (so called *cell assemblies*) is the currency of information processing [Buzsaki, 2010; Harris, 2005]. This notion is supported by recent experiments showing that reciprocal connections between neurons occur above chance level [Perin et al., 2011; Song et al., 2005], especially if neurons receive common inputs [Otsuka and Kawaguchi, 2011; Yoshimura et al., 2005]. In the case of the visual system, for instance, excitatory neurons with similar response features tend to be more connected to each other [Harris and Mrsic-Flogel, 2013; Ko et al., 2011]. Moreover, studies have demonstrated that neurons exhibit layer-specific connectivities within rodent sensory cortex [Lefort et al., 2009] and neocortex [Yassin et al., 2010]. In addition, organized architectures have been observed to occur at multiple hierarchical scales [Felleman and Van Essen, 1991; McGinley and Westbrook, 2013; Savic et al., 2000; Shimono and Beggs, 2014] and in non-mammalian organisms [Ito et al., 2013]. These findings suggest that cortical regions contain well-defined subnetworks. However, the underlying question is whether given a network topology, we can predict the potential of the network to sustain structured spatio-temporal activity. Such questions are not only of interest for network dynamics, but also have implications for memory formation and learning, since neural networks undergo topological changes over time due to plasticity [Hyman et al., 2003; J Pavlides et al., 1988].

Recently, it has been shown computationally (see e.g. Ref. [Litwin-Kumar and Doiron,

2012]) that leaky-integrate-and-fire (LIF) networks with equal excitatory and inhibitory connection net strengths (*i.e.* balanced [van Vreeswijk and Sompolinsky, 1998]) yet with clustered excitatory connections, can exhibit prolonged heightened group activity, with the activity transitioning between groups in the network (Fig. 3.1A). Here we characterize the emergence of such slow-switching segregated dynamics in balanced LIF networks as a result of the network connectivity. Specifically, we find that the spectral properties of the synaptic weight matrix (i.e., the existence of an eigenvalue gap and a block-localized dominant subspace) provide a criterion to predict the appearance of such activity in the network. We then use simple linear rate models to gain insight into the mechanisms underpinning the origin of such dynamics in structurally clustered LIF networks. Using these insights, we construct novel LIF topologies that display slow-switching group activity with distinct properties: involving both inhibitory and excitatory neurons, exhibiting multiple slow time-scales, as well as demonstrating the possibility of such dynamics in networks with no obvious clustered connectivity, such as small-worlds. Finally, we discuss briefly possible implications of the different wiring schemes for neural computation.

3.3 Results

3.3.1 Slow-switching assemblies in LIF networks with clustered excitatory connections: spectral insights

Clustered excitatory topologies in a balanced LIF network can lead to dynamics in which localized high activity states transition between assemblies of neurons within the network [Litwin-Kumar and Doiron, 2012]. This is illustrated in Figure 3.1A, where the dynamics of an unstructured and a balanced clustered network with 20 groups are shown side by side (see Materials and Methods for a description of the networks). Hereafter, we will refer to such activity as *slow-switching assembly* (SSA) dynamics. Visually, SSA dynamics manifests itself as bands of increased activity in the raster plots, and can be statistically quantified from the resulting spike-train dynamics *a posteriori* (see Eq. (3.18) in Materials and Methods). Ideally, however, we would like to establish *a priori*, solely from the given connectivity, the possibility of such dynamical patterns emerging.

The full dynamics of LIF networks are notoriously difficult to analyze due to their inherent non-linearity and hence an exact analytical treatment of the dynamical evolution for an arbitrary clustered topology is essentially intractable. However, two concepts from spectral graph theory and linear systems provide valuable insights: (i) for symmetric, non-negative connectivity matrices, it can be shown that a modular network structure implies a gap in

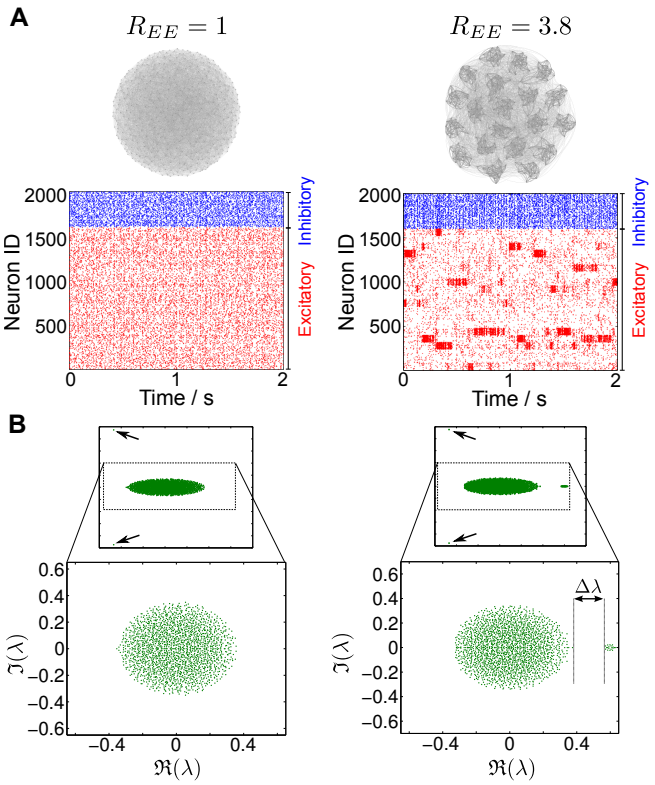

Figure 3.1: Network dynamics and eigenvalue spectra of two LIF networks: one with uniform synaptic connections (left), and one with 20 groups of clustered excitatory connections (right). To create the clustered network, excitatory neurons were partitioned into groups (with in-group connection probability p_{in}^{EE} and out-group connection probability $p_{\text{out}}^{EE} < p_{\text{in}}^{EE}$) while keeping the average connectivity constant (see Materials and Methods and Ref. [Litwin-Kumar and Doiron, 2012]). The ratio $R_{EE} = p_{\text{in}}^{EE}/p_{\text{out}}^{EE}$ controls the strength of the excitatory clustering. **A** Visualization of the network topologies (top) and exemplars of raster plots (bottom). The dynamics of the clustered network exhibit the banded structure associated with slow-switching group activity. The magnitude of the activity can be characterized statistically *a posteriori* from the data through the spike-rate variability metric $\widehat{S}$, defined in Materials and Methods Eq. (3.18), as discussed in the text. In this case, the unclustered network has $\widehat{S} = 0.035$ while the clustered network has a much larger value $\widehat{S} = 8.23$. **B** Eigenvalue spectra of the network weight matrices W. The weighted connectivity matrix of the clustered network exhibits a clear eigengap $\Delta\lambda$ separating the 19 eigenvalues with largest real parts from the cloud of eigenvalues in the bulk. There is no such eigengap for the unclustered network. As indicated by the two arrows, both matrices have a pair of complex conjugate eigenvalues associated with the (damped) global activation modes of the networks characteristic of balanced networks (see text and Ref. [Murphy and Miller, 2009]).

the spectrum of the graph (i.e., in the set of eigenvalues of the weight matrix) [von Luxburg, 2007; Zhang et al., 2013], as well as the block-localization of the associated eigenvectors on the modules of the network (noting that isolated eigenvalues may also be the result of other features [Nadakuditi and Newman, 2013]); (ii) for linear systems, a gap in the spectrum of the graph results in a *separation of time scales* in the dynamical process [Galán, 2008;

Murphy and Miller, 2009; Simon and Ando, 1961]. This relation between the modular structure, the eigenvalues and associated eigenvectors, and linear network dynamics can be used to discover modular structures in graphs from a dynamical perspective [Billeh et al., 2014; Delvenne et al., 2013, 2010; Schaub et al., 2012].

In fact, the weighted connectivity matrices of unclustered and clustered LIF networks display different spectral characteristics [Litwin-Kumar and Doiron, 2012]. Figure 3.1B shows the spectra of two networks with 1600 excitatory and 400 inhibitory neurons with unclustered (left) and clustered (right) topologies. In the unclustered case, we find the expected circular distribution of eigenvalues, which follows from the properties of random graphs [Rajan and Abbott, 2006; Sommers et al., 1988], although the presence of groups of neurons with different cardinalities and variances means that the eigenvalue distribution is not completely uniform on the circle [Rajan and Abbott, 2006]. Note also that the balanced construction of the LIF network, with a marginally larger inhibitory input for each neuron in order to keep the network stable, leads to the existence of one pair of complex conjugate eigenvalues which lies separate from the main bulk of the spectrum (black arrows in Figure 3.1B). This eigenvalue pair is associated with the global activation mode of the network (as explained below in Section 3.3.2 and in Ref. [Murphy and Miller, 2009]). As shown in Figure 3.1A, this unclustered network exhibits the expected asynchronous, unstructured neuronal spiking dynamics.

In contrast, the LIF network with clustered excitatory neurons displays banded SSA dynamics. Spectrally, its weight matrix exhibits a clear gap $\Delta\lambda$ along the real axis of its spectrum (Figure 3.1B, right) and, as shown below in Section 3.3.2.3 (Fig. 3.6), the associated Schur vectors also exhibit a measure of structural block-localization. We remark that, in general, this should not be expected *a priori*, since the weight matrices are *asymmetric* and include both positive (excitatory) and negative (inhibitory) couplings.

To ascertain the dynamical relevance of the spectral gap, we examined the relation between the clustering strength in the network, defined as the ratio of probabilities of connections inside and outside the neural assemblies ($R_{EE} = p_{\text{in}}^{EE}/p_{\text{out}}^{EE}$), the spectral gap ($\Delta\lambda$), and the magnitude of the numerically observed SSA dynamics. To quantify the assembly spike-rate variability, we have defined two complementary metrics. First, the metric $\widehat{S}$ measures the heterogeneity in the firing rates of the putative cell-assemblies averaged over the simulation (see Eq. (3.18) in Materials and Methods). A large value of $\widehat{S}$ indicates that the average firing rates of the assemblies are diverse, whereas a low $\widehat{S}$ indicates that all groups have very similar firing rates at all times and no group shows elevated firing. Under SSA dynamics, the variability of firing rates across groups increases in time as the assemblies transition between high and low firing rates. As discussed below in Section 3.3.2.4, it is

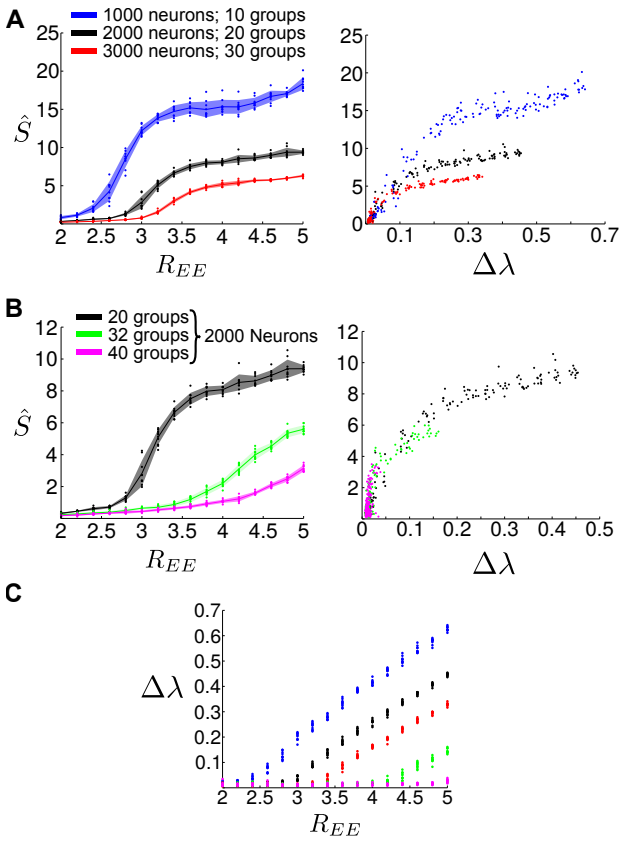

Figure 3.2: The relationship of observed SSA dynamics with the structural connectivity clustering of the LIF network, R_{EE}, and the spectral gap, $\Delta\lambda$. The presence of SSA dynamics is quantified through the spike-rate variability metric ($\widehat{S}$), which measures the variance of the firing rates of the assemblies normalized by a randomly shuffled bootstrap: $\widehat{S}$ increases with increasing SSA activity and $\widehat{S} \approx 0$ for completely asynchronous activity (as in the unclustered case in Figure 3.1A) (see Materials and Methods). **A** Spike-rate variability is plotted as a function of R_{EE} (left) and $\Delta\lambda$ (right) for different network sizes (dots: raw data from simulations, line: mean, shading: standard deviation). Above a certain clustering threshold, SSA emerges and increases as R_{EE} grows; the intensity of the SSA dynamics is in line with the presence of an eigenvalue gap $\Delta\lambda$ in the weight matrix. **B** Spike-rate variability as a function of R_{EE} (left) and $\Delta\lambda$ (right) for a network of 2000 neurons with different numbers of clusters, yielding qualitatively similar results (dots: raw data from simulations, line: mean, shading: standard deviation). **C** Relationship between the clustering strength R_{EE} and the spectral gap $\Delta\lambda$. Observe that R_{EE} is not sufficient to determine $\Delta\lambda$, i.e. $\Delta\lambda$ is influenced by other aspects such as the network size and number of groups.

possible that the heightened firing activity is localized in a particular assembly and does not switch between assemblies. This non-switching dynamics where only a group of neurons exhibits elevated firing over the entire simulation time can be detected using the second variability measure $\widehat{S}_T$, defined in Eq. (3.20). For SSA dynamics, both $\widehat{S}$ and $\widehat{S}_T$ give similar results (see Figure 3.3). In the rest of the chapter, we focus on SSA dynamics and mainly use $\widehat{S}$, except when discussing the transition to non-switching behavior in Section 3.3.2.4.

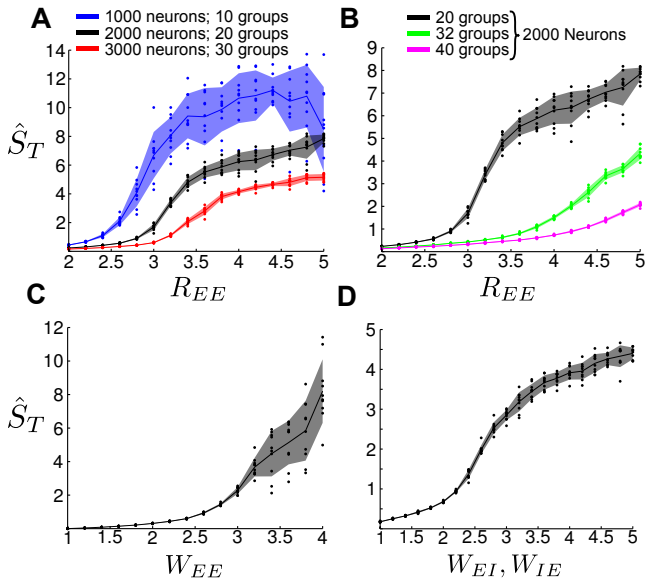

Figure 3.3: Spike rate variability across time, as a function of topological changes of the network. **A** Clustering Strength R_{EE} vs spike rate variability across time $\hat{S}_T$ for varying network sizes with a fixed assembly size. Compare to Fig. 2A of the main text. **B** Clustering Strength R_{EE} vs spike rate variability across time $\hat{S}_T$ for a fixed network size with varying assembly sizes. Compare to Fig. 2B of the main text. **C** Weight Clustering W_{EE} vs spike rate variability across time $\hat{S}_T$ for a LIF network with 2000 neurons (20 groups; topological clustering $R_{EE} = 1$). **D** Spike rate variability across time $\hat{S}_T$ as a function of $W_{EI} = W_{IE}$ for a network with excitatory to inhibitory feedback loops. In all plots dots denote raw data from simulations; line: mean, shading: standard deviation.

Definitions of these quantities are given in Materials and Methods.

Figure 3.2 shows that SSA becomes observable above a threshold of the clustering strength R_{EE}, although the dependence of this threshold on the size of the network and number of clusters does not seem to follow an obvious pattern. On the other hand, our numerics show that the spectral gap provides a more direct indicator of the presence of SSA dynamics in the network. The relationship between R_{EE} and $\Delta\lambda$ (Figure 3.2C) shows that knowing R_{EE} is not sufficient to determine $\Delta\lambda$ as the spectral gap depends on other factors such as the network size and the number of groups in the network. Together with the examination of the structure of the associated orthogonal subspace (see Section 3.3.2.3), such spectral characterization may be used to establish the potential of networks to sustain SSA activity, even if there is no obviously clustered network model known *a priori*.

The observed spectral properties of clustered LIF networks lead us to consider a stylized linear rate model for neuronal activity in the next section, as a means to gain insights into the defining factors of the wiring diagram leading to SSA.

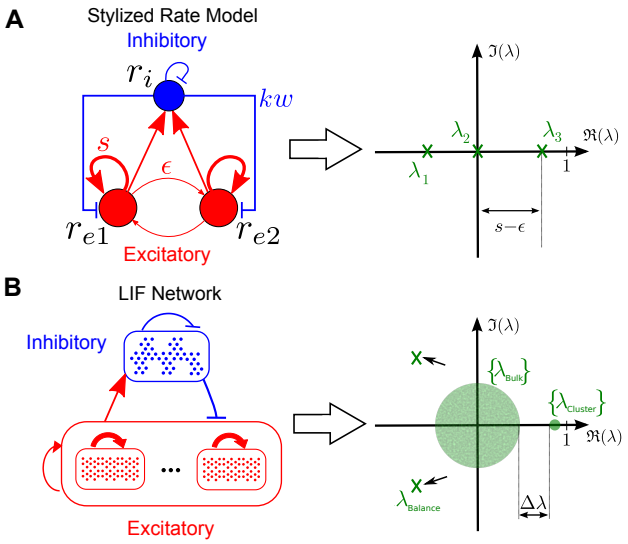

Figure 3.4: Structure of the eigenspectra of stylized and realistic networks that can exhibit SSA activity. **A** Schematic of the stylized model of a clustered network (3.1)–(3.2) and illustration of the corresponding eigenvalues (3.6). **B** Schematic of a more realistic clustered LIF network with the same wiring scheme on many groups and multiple neurons per group with an illustration of the associated spectrum.

3.3.2 Linear rate models and slow localized activity in clustered LIF networks

3.3.2.1 A stylized linear rate model for networks with clustered excitatory neurons

Motivated by the above findings, we proceeded to investigate how much of the observed LIF dynamics can be described by simple rate models. From above, we expect that the spectral properties of the network lead to SSA dynamics that are coherent over, and transitioning between, clusters. Hence we consider simple rate models on coarse-grained networks where each node describes the behavior of a cluster. As will become clear below, this reduced description allows us to identify the mechanism underpinning the onset of SSA activity without having to consider a plethora of statistical and technical detail. Furthermore, our main conclusions can be easily extended to multiple groups of clustered excitatory neurons, or to scenarios with probabilistic connectivity between the nodes (see, *e.g.*, Ref.[Murphy and Miller, 2009]).

To explore this idea in the simplest setting, consider a linear firing rate model with three groups: two groups of clustered excitatory neurons and a group of inhibitory neurons with uniform coupling. This wiring scenario is abstracted in the form of a three-node network (Figure 3.4A), where each node represents a group of neurons, and a simple rate model is

given by the following equations:

$$\tau \frac{d\mathbf{r}}{dt} = -\mathbf{r} + W\mathbf{r} + \boldsymbol{\xi} = -(I - W)\mathbf{r} + \boldsymbol{\xi}. \tag{3.1}$$

Here $\mathbf{r} = (r_{e1}, r_{e2}, r_i)^T$ are the firing rates of the neuron groups, W is the synaptic connectivity weight matrix, and $\boldsymbol{\xi}$ is a random input to the network. The firing rates are measured relative to some baseline activity, and may be positive or negative.

To describe this topology with two excitatory clusters we define a connectivity matrix W of the form:

$$W = \begin{pmatrix} s & \epsilon & -kw \\ \epsilon & s & -kw \\ w/2 & w/2 & -kw \end{pmatrix}, \tag{3.2}$$

$$s > \epsilon > 0, \quad k \geq 1, \quad w = (s + \epsilon). \tag{3.3}$$

The first inequality in (3.3) guarantees that the *coupling strength* is positive: $s - \epsilon > 0$, i.e., the cross-coupling between the excitatory groups is weaker than the intra-group coupling, as would be expected from the notion of a cluster. The second inequality in (3.3) ensures a balanced network (while keeping $k \approx 1$), such that each group has at least the same amount of inhibitory and excitatory input: $\sum_j W_{ij} = (1 - k)w \leq 0$, $\forall i$.

To understand the dynamical behavior of this system, we evaluate the spectral properties of the connectivity matrix W [Galán, 2008; Goldman, 2009; Murphy and Miller, 2009]. Note that we are dealing with a non-normal system, in which the eigenvectors may not be orthogonal and thus will provide a possibly misleading description of the dynamics [Trefethen and Embree, 2005]. We therefore consider the Schur decomposition of W, which yields an orthonormal basis of the system [Goldman, 2009; Murphy and Miller, 2009], i.e., it finds an orthogonal matrix U and an upper triangular matrix Q such that $W = UQU^T$. Here U contains an orthonormal basis while Q contains the eigenvalues of W on the diagonal and non-zero elements above the diagonal accounting for effective feedforward connections between the different modes [Goldman, 2009; Murphy and Miller, 2009].

For the matrix (3.2), the associated upper triangular Schur form Q is:

$$Q = \begin{pmatrix} -w^+ & w_{ff} & 0 \\ & 0 & 0 \\ & & s - \epsilon \end{pmatrix}, \tag{3.4}$$

$$w^+ = w(k - 1), \quad w_{ff} = (k + 1/2)(w + \epsilon),$$

and the matrix containing the orthonormal (Schur) basis is $U = \left(\mathbf{u}_1\, \mathbf{u}_2\, \mathbf{u}_3\right)$ given by

$$\mathbf{u_1} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{u_2} = \sqrt{\frac{2}{3}} \begin{pmatrix} 0.5 \\ 0.5 \\ -1 \end{pmatrix}, \mathbf{u_3} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}. \tag{3.5}$$

Each of the Schur vectors $\mathbf{u_i}$ represents a pattern of rates, a particular mode of firing on the network. The first mode $\mathbf{u_1}$ is associated with an overall mean firing pattern, while the second mode $\mathbf{u_2}$ corresponds to a relative difference in firing between inhibitory and excitatory groups. Note that in both cases the firing patterns of the two groups of excitatory neurons are aligned. In contrast, the third mode $\mathbf{u}_3$ describes an antagonistic activity localized on the excitatory neuron groups (i.e., when the firing rate of one excitatory group increases the other decreases, and vice versa) with the inhibitory node remaining unaffected at a baseline firing rate, precisely in line with the SSA behavior observed in the full LIF network (Fig. 3.1A). Observe also that, while the first two modes are coupled via an off-diagonal term in Q (i.e., an effective feedforward connection [Goldman, 2009; Murphy and Miller, 2009]), the switching behavior described by the third mode is uncoupled from the other system modes.

The structure of the Schur decomposition provides us with insight into the dynamics of the model (3.1). From (3.4), the eigenvalues of W are:

$$\lambda_1 = -w^+ \leq \lambda_2 = 0 < \lambda_3 = s - \epsilon, \tag{3.6}$$

whence the eigenvalues governing (3.1) are $\eta_1 = -1 - w^+$, $\eta_2 = -1$ and $\eta_3 = -1 + (s - \epsilon)$. For the system to be linearly stable, we require $\eta_3 < 0$, which implies that the clustering strength in a stable system is constrained to be

$$0 < s - \epsilon < 1. \tag{3.7}$$

Without an external input, the three modes decay exponentially with time constants $\tau_i = -1/\eta_i = 1/(1 - \lambda_i)$. Therefore, it is possible to slow down the time scale associated with the SSA mode $\mathbf{u}_3$ by increasing the clustering strength $s - \epsilon \to 1$. The perturbations introduced by the random input induce mode-mixing and, in particular, break the symmetry between the firing rates of the two groups of excitatory neurons. These perturbations excite the SSA mode $\mathbf{u_3}$, which will decay with time scale $1/(1 - \lambda_3)$ towards the solution with uniform rates on all groups ($\mathbf{r}^* = 0$).

In conclusion, the stronger the clustering (i.e., the larger $s - \epsilon < 1$), the slower the

decrease of the asymmetric transients, and the more prominent SSA dynamics becomes. Note that this slow time scale in the activity patterns is directly associated with the eigengap $\lambda_3 - \lambda_2 = s - \epsilon$ (Fig. 3.4), which is dictated by the clustering strength, in agreement with our observations relating the presence of SSA with $\Delta\lambda$ in Figure 3.2. The insights gained from this simple model have testable implications for LIF networks and lead to further ideas for neuro-physiological network wiring structures, which we explore in the following sections.

3.3.2.2 The linear rate model and the full dynamics of clustered LIF networks

Our rate model can be extended to networks with c groups and the overall structure of the eigenspectrum of W will essentially retain its features (Figure 3.4; see [Murphy and Miller, 2009] for a related discussion). First, there will be a 'negative' eigenvalue (λ_1 in (3.6) related to the pair λ_{Balance} in Fig. 3.4B), which reflects the fact that the network is balanced and stable, *i.e.* the associated global activity mode decays. Second, there will be a bulk of 'small' eigenvalues centered around the origin (λ_2 in (3.6) and in general the set $\{\lambda_{\text{Bulk}}\}$ in Fig. 3.4B), which stem from the random connectivity present in the network—a random matrix has a circular eigenvalue distribution around the origin). Finally, there will be an eigenvalue gap separating a small set of $c-1$ 'large positive' eigenvalues (λ_3 in (3.6) and $\{\lambda_{\text{Cluster}}\}$ in Fig. 3.4B) from the bulk of the spectrum. These are the 'slow' eigenvalues associated with an orthogonal Schur subspace with a block structure coherent with the clusters, and are thus responsible for the appearance of a new time scale in the system originating the observed SSA dynamics. This structure of the eigenvalues can be seen in the spectrum of the LIF network in Figure 3.1B.

Inspired by the analysis of the linear rate model, we investigate how the results translate to the full nonlinear LIF dynamics. First, it is clear from above that only the difference between the effective intra- and inter-group coupling strengths is essential for the observed slow activity. Such difference can be the result of clustered topological connectivity as in Ref. [Litwin-Kumar and Doiron, 2012], but can be equally achieved keeping the topology uniformly connected and tuning the synaptic weights. To assess this possibility, we conducted numerical simulations of LIF networks with uniform connection probabilities between excitatory neurons, yet with larger intra-group synaptic weights while keeping the average weight constant (see Materials and Methods). Figure 3.5 shows that tuning the weights towards a clustered structure (as given by the weight ratio $W_{EE} = w_{\text{in}}^{EE}/w_{\text{out}}^{EE}$) also leads to the emergence of SSA dynamics linked to a spectral gap. Although unsurprising from a linear systems perspective, this result confirms the applicability of the spectral characterization to nonlinear LIF networks, and specifically establishes its relevance for SSA dynamics in networks where only synaptic weights can be tuned.

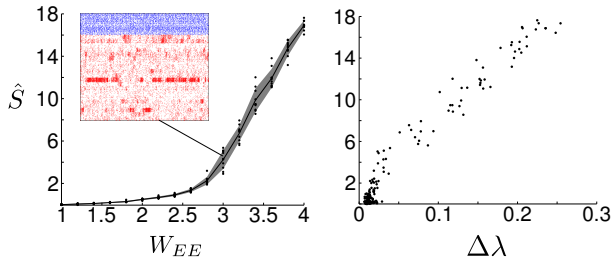

Figure 3.5: SSA dynamics can be achieved by changing solely the synaptic strengths even when the topological connections are kept uniform. To create clustered networks, the ratio $W_{EE} = w_{\text{in}}^{EE}/w_{\text{out}}^{EE}$ was varied, where w_{in}^{EE} and w_{out}^{EE} refer to the in-group and out-group synaptic weights, respectively. The spike-rate variability $\widehat{S}$ measuring the intensity of SSA dynamics is shown as a function of W_{EE} (left - dots: raw data from simulations, line: mean, shading: standard deviation) and the spectral gap $\Delta\lambda$ (right). Note that the connection probability is uniform for all the simulations, *i.e.* $R_{EE} = 1$, and only the clustering of the weights is varied.

3.3.2.3 The block-localization of the dominant linear subspace

We now consider a critical factor in the emergence of SSA, namely, that the leading Schur vectors of the weight matrix of LIF networks exhibit a measure of consistent block localization on the groups of neurons which then exhibit cell assembly dynamics (Figure 3.6). Due to the non-normality outlined above [Goldman, 2009; Murphy and Miller, 2009], we consider orthogonal Schur vectors and not eigenvectors. A typical Schur vector corresponding to one of the eigenvalues in the leading group $\{\lambda_{\text{Cluster}}\}$ exhibits a coherent pattern on each of the groups of the network, such that all neurons within a particular group are driven uniformly. This block-uniformity results in the grouped dynamics that characterizes SSA. Representative patterns are shown in Figure 3.6A. In contrast, the Schur vectors from the bulk of the spectrum $\{\lambda_{\text{Bulk}}\}$ show no specific pattern of localization on any group of neurons. Note that the Schur vectors show no localization pattern on the inhibitory neurons either, in line with our analysis. The block-localization of the leading Schur vectors constitutes an additional spectral characterization to assess the emergence of SSA dynamics from the weight matrix alone.

In Figure 3.6, we quantify how the patterns observed in the SSA dynamics align with the dominant Schur vectors of the weight matrix. To do this, we first performed a principal component analysis (PCA) of the simulated firing rates of the full LIF network and extracted the top activation patterns $\{\mathbf{p_1}, \ldots, \mathbf{p_c}\}$ corresponding to the first principal components (see Figure 3.6B and Materials and Methods). To measure the alignment of the data patterns $\{\mathbf{p_i}\}$ with the dominant Schur vectors $\{\mathbf{u_1}, \ldots, \mathbf{u_c}\}$ of the weight matrix W corresponding to the leading eigenvalues above the gap, we computed the *first principal angle* θ between these two subspaces: the smaller the angle, the closer the alignment [Golub and Van Loan,

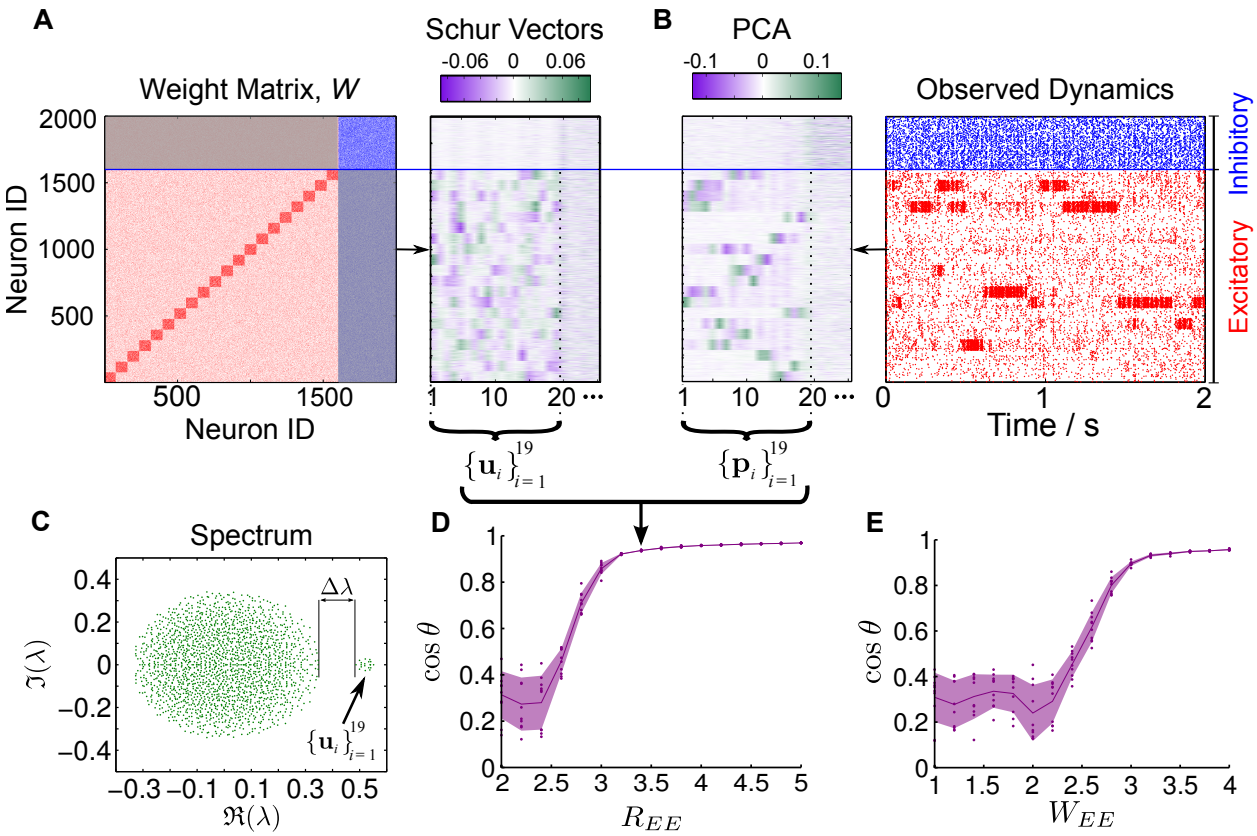

Figure 3.6: Unraveling the association between the leading Schur vectors of the weight matrix and the observed SSA activity. **A** Illustration of a weight matrix of the $N = 2000$ neuron LIF network with $c = 20$ groups of excitatory neurons ($R_{EE} = 3.4$) shown together with the heatmap of the real parts of its first 25 Schur vectors. The first 19 Schur vectors exhibit block-uniform patterns relatively constant within each of the 20 groups of neurons, whereas no such pattern is observed for the other Schur vectors. Note that there is no pattern discernible over the inhibitory neurons. The leading Schur vectors correspond to the cloud of 'slow' eigenvalues above the gap, as indicated in **C**, and span a patterned dominant subspace that induces grouped dynamics in the network. **B** A long simulation (80s) of the LIF network dynamics (only the first 2s shown) was analyzed using PCA and the first 25 principal components (PCs) are shown. Reflecting the banded structure of the simulated dynamics, the leading PCs also show a block-patterned structure consistent with the neuronal groups. **C** On the spectrum of W, we indicate the group of leading eigenvalues above the gap associated with the dominant subspace. **D** The alignment between the dominant Schur subspace of the W matrix and the subspace of the strongest principal components is measured by the first principal angle θ (3.21). Above a threshold of the clustering strength R_{EE}, both subspaces become highly aligned in line with the observations in Fig. 3.2 (dots: raw data from simulations; line: mean; shading: standard deviation). **E** The same effect is observed when the clustering is introduced in the weights by varying W_{EE}, as in Fig. 3.5.

1996; Stewart, 2001]. Figure 3.6D shows that the dominant subspace of the Schur vectors of the weight matrix (W) are indeed aligned with the observed patterns and consistent with the embedded groups of neurons in the network. Finally, Figure 3.6E shows again that

tuning the synaptic weights to cluster the network while keeping the topology uniformly connected renders a similar outcome.

3.3.2.4 Increasing the clustering beyond the linearly stable regime: one dominant assembly

Until now, our linear analysis has concentrated on the relevant regime for SSA, where the system is linearly stable (3.7): $\lambda_3 = s - \epsilon < 1$. However, it is possible to increase the clustering strength $(s - \epsilon)$, so that the intra-cluster connections are much stronger than the inter-cluster connections, and the linear system has an unstable eigenvalue. For the LIF network, if we increase R_{EE} or W_{EE} so that $\lambda_{\max} \in \{\lambda_{\text{Cluster}}\} \geq 1$, then the associated localized firing mode, once excited, will activate itself. This regime can thus lead to a 'winner takes all' solution, where one cell assembly dominates the firing of the whole network and suppresses all other neurons [Rutishauser et al., 2011].

From the perspective of the observed dynamics of the LIF network, an increase of the clustering strength (R_{EE} or W_{EE}) initially leads to the development of the eigengap for $\{\lambda_{\text{Cluster}}\}$ and, the emergence of SSA dynamics. However, when the clustering strength becomes too large (and the largest eigenvalue goes beyond 1), a single assembly starts to dominate the firing pattern and the firing variability is reduced across time, an effect which is captured by a strong reduction in $\widehat{S}_T$, as illustrated in Figure 3.7. This behavior, which has been observed previously [Litwin-Kumar and Doiron, 2012], is thus also closely linked to the spectral properties of the system. Note that the linear condition $\lambda_{\max} \geq 1$ is only indicative: the non-linearity and boundedness of LIF dynamics can lead to (non-linearly stable) SSA dynamics beyond this condition.

3.3.3 Beyond clustered excitatory neurons: SSA dynamics involving both excitatory and inhibitory neurons.

So far we showed that a linear rate model can provide valuable insights into the mechanisms underpinning SSA dynamics in networks with clustered excitatory neurons. The crucial point for the emergence of SSA is the separation of time scales, dictated by the splitting of the leading eigenvalues of the weight matrix, together with the block-localization of the associated dominant subspace. These spectral properties can be introduced into LIF networks by entirely different synaptic couplings, and we now consider a mechanism in which the inhibitory neurons are involved more centrally in generating SSA dynamics.

Consider the wiring diagrams presented in Figure 3.8A-B, representing the simple rate model and its corresponding LIF schematic. In this case, the wiring does not involve clus-

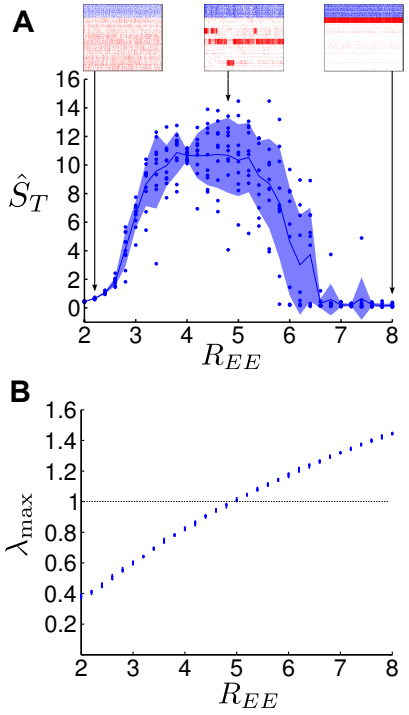

Figure 3.7: Effect of large clustering strength on SSA activity. **A** Large values of the clustering R_{EE} lead to linear instability of the SSA dynamics and localization of the activity on one assembly. As measured by the spike rate variability across time ($\hat{S}_T$), the increase of R_{EE} leads to SSA (signalled by the increased value of $\hat{S}_T$). If the clustering increases further, $\hat{S}_T$ decreases, as the dynamics becomes dominated by one assembly only. The dots denote raw data from simulations (line: mean; shading: standard deviation). Inset: examples of raster plots for three data points in the three regimes. The analysis corresponds to a clustered LIF network of 1000 neurons. **B** Plot of the eigenvalue with the largest real component $\lambda_{\max}$ against the clustering strength R_{EE}. The linear condition $\lambda_{\max} > 1$ is a good indicator of the dynamics becoming dominated by one cell assembly.

tered coupling between groups of excitatory neurons, but rather relies on preferential connectivity patterns *between* inhibitory and excitatory neuron groups only. Each group of excitatory neurons activates preferentially an associated group of inhibitory neurons and, in turn, this group of inhibitory neurons feeds back more weakly to its associated group of excitatory neurons (see Materials and Methods for a full description of the network). Therefore the effective functional circuitry consists of both excitatory and inhibitory neurons embedded in a feedback loop and, as a consequence, the inhibitory neurons play an integral role in generating the spatio-temporal dynamics and display SSA dynamics too, as we show below.

This wiring mechanism was suggested by the stylized firing rate model with two coupled pairs of inhibitory/excitatory feedback loops in Figure 3.8A. This system is described by Eq. (3.1) with a vector of firing rates for the four groups $\mathbf{r} = (r_{e1}, r_{e2}, r_{i1}, r_{i2})^T$ and a

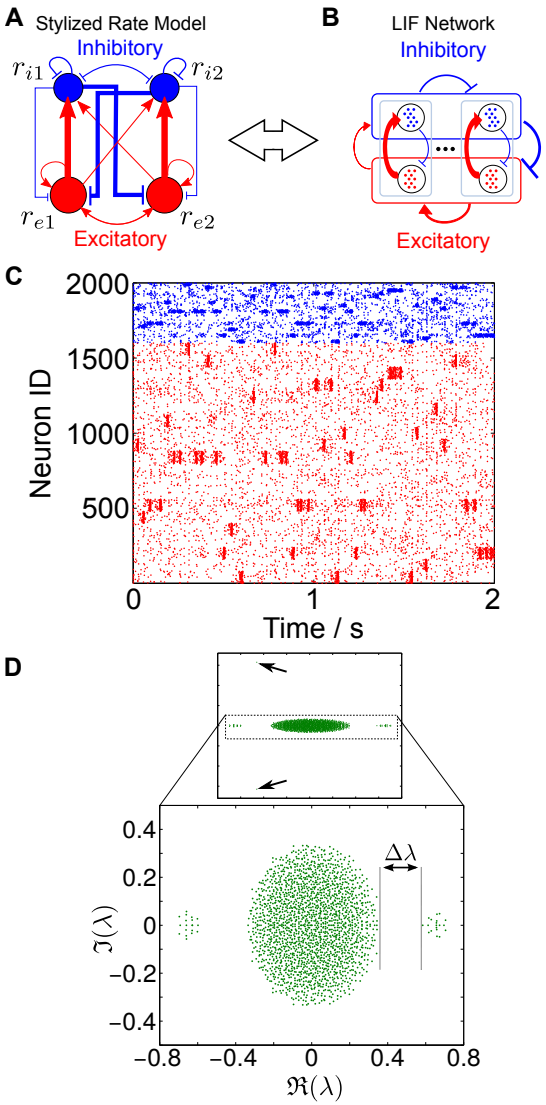

Figure 3.8: SSA dynamics can result from excitatory-to-inhibitory feedback loops. **A** Schematic of a stylized network model with two excitatory-to-inhibitory feedback loops corresponding to the model (3.8). **B** Illustration of the corresponding LIF network with such a co-clustered feedback mechanism between neuron types. **C** Raster plot of the dynamics of a co-clustered LIF network with $N = 2000$ neurons and $c = 20$ pairs with excitatory-to-inhibitory feedback. Note that the inhibitory neurons also exhibit SSA dynamics here. **D** Spectrum of the weight matrix exhibiting a spectral gap both on the left and right hand side of the bulk of the spectrum (see text for details).

synaptic weight matrix defined as:

$$W = \begin{pmatrix} w & w & -k\epsilon & -ks \\ w & w & -ks & -k\epsilon \\ s & \epsilon & -kw & -kw \\ \epsilon & s & -kw & -kw \end{pmatrix}, \tag{3.8}$$

$$s > \epsilon, \quad k \geq 1, \quad w = (s+\epsilon)/2. \tag{3.9}$$

Similarly to (3.2), the system is balanced and $s-\epsilon$ captures the clustering strength within the excitatory-inhibitory pairs. The Schur decomposition of W leads to the following Schur form

$$Q = \begin{pmatrix} -w^+ & w_{ff} & & \\ & 0 & & \\ & & \sqrt{k}(s-\epsilon) & w_{ff,2} \\ & & & -\sqrt{k}(s-\epsilon) \end{pmatrix}, \tag{3.10}$$

$$w^+ = (k-1)(s+\epsilon),$$

$$w_{ff} = (k+1)(s+\epsilon), \quad w_{ff,2} = -(k-1)(s-\epsilon),$$

where the eigenvalues of W are on the diagonal. When the leak term is considered, the largest eigenvalue of (3.1) is $(-1+\sqrt{k}(s-\epsilon))$. Hence, to keep the system stable we need to constrain

$$0 < \sqrt{k}\,(s-\epsilon) < 1, \tag{3.11}$$

and the spectral gap is again controlled by $s-\epsilon$. In the associated orthonormal Schur basis, the first two modes of the firing rate dynamics

$$\mathbf{u_1} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{u_2} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \tag{3.12}$$

are again global 'sum' and 'difference' modes, which interact via a balanced amplification mechanism [Murphy and Miller, 2009]. However, there are also two localized modes $\mathbf{u}_3$ and $\mathbf{u}_4$ that can lead to slow structured activity:

$$\mathbf{u_3} = \frac{1}{\sqrt{2k+2}}\begin{pmatrix} \sqrt{k} \\ -\sqrt{k} \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{u_4} = \frac{1}{\sqrt{2k+2}}\begin{pmatrix} 1 \\ -1 \\ -\sqrt{k} \\ \sqrt{k} \end{pmatrix}. \tag{3.13}$$

Of these, $\mathbf{u_3}$ is associated with the largest eigenvalue and describes the slow(est) dynamics. This mode corresponds to a firing pattern of correlated activity within the pairs (r_{e1}, r_{i1}) and (r_{e2}, r_{i2}), and anti-correlated activity across the pairs. As before, this analysis extends to networks with more groups and/or stochastic coupling between groups (for related arguments see supplementary material of Ref. [Murphy and Miller, 2009]).

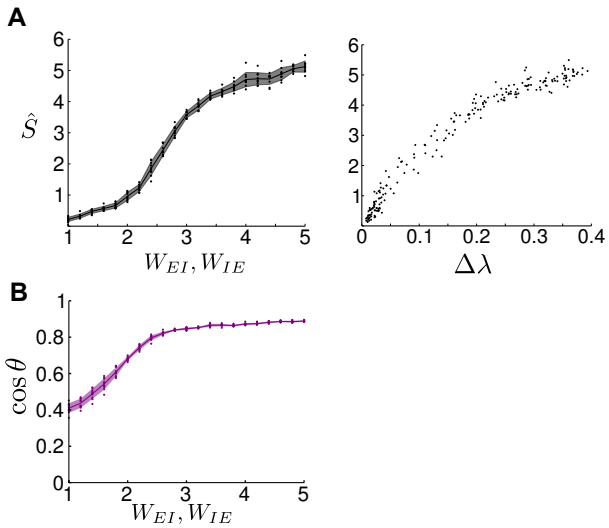

Figure 3.9: The presence of SSA dynamics in LIF networks with excitatory to inhibitory feedback loops. **A** The spike-rate variability, which measures the intensity of SSA dynamics, as a function of $W_{EI} = W_{IE}$ (left; dots: raw data from simulations, line: mean, shading: standard deviation) and $\Delta\lambda$ (right). See text and Fig. 3.5 for further details. **B** The first principal angle between the subspaces of the principal firing patterns and the dominant Schur vectors of the weight matrix W show high alignment. See text and Fig. 3.6 for further details.

In Figure 3.8C-D, we show that the implementation of this wiring mechanism into a full LIF network displays SSA dynamics with a spectral gap, as predicted by our simple rate model, and the paired excitatory/inhibitory neurons act as a functional circuit displaying synchronous firing behavior. Importantly, note that, in contrast to the excitatory clustered scenario, there are no groups with dense reciprocal couplings in this network topology. Our numerics also show that varying the strength of the functional grouping ($W_{EI} = W_{IE}$) leads to the emergence of SSA dynamics, linked to the presence of the spectral gap $\Delta\lambda$ (Fig.3.9A), and to the alignment of the leading Schur vectors with the 'correct' functional circuits, i.e., pairs of excitatory and inhibitory groups together (Figure 3.9B).

It is interesting to note that in this topology there is also a group of eigenvalues bounded away in the negative direction (see Figure 3.8D). These modes are the quickest decaying and correspond to 'anti-correlated' firing states, in which excitatory neurons act in synchrony with the inhibitory groups to which they are not functionally associated. Such fast decay reinforces the survival of synchrony within the functional groups in the network. Interestingly, these modes were already present in our linear rate model: $\mathbf{u_4}$ in (3.13) shows the same firing pattern with the fastest eigenvalue $\lambda_4 = -\sqrt{k}(s - \epsilon)$.

3.3.4 SSA dynamics in LIF networks with alternative topologies

Although perhaps the most intuitive way of generating SSA dynamics follows from clustering the connectivity of the LIF network, our analysis above suggests that alternative types of structural organization support SSA dynamics in LIF networks, as long as they are characterized by a gap in their eigenvalue spectrum and a measure of block-localization of the Schur vectors. More generally, one could conjecture that any low rank perturbation of the weight matrix of a balanced network which leads to an eigenvalue gap might be a valid candidate for a mechanism to generate SSA in neuronal dynamics. A few such network architectures are worth highlighting, as they have been considered of particular relevance in a neuroscience context.

3.3.4.1 SSA dynamics in networks with small-world organization

The first noteworthy example is the broad class of small-world like networks [Watts and Strogatz, 1998], since small-world organizations have been observed in many structural and functional neurophysiological networks [Bullmore and Sporns, 2009]. While many modular networks can have the small-world property [Meunier et al., 2010], here we focus on the archetypal small-world structure à la Watts and Strogatz [Watts and Strogatz, 1998], which does not display a distinct modular organization but instead may be seen as a mixture of a k-nearest neighbor ring lattice and a random graph. It is well known that such small-world networks possess distinctive spectral properties which have important implications for dynamical processes, such as synchronization [Barahona and Pecora, 2002]. As Figure 3.10A shows, small-world like LIF networks (see Materials and Methods for details of the construction) can indeed display SSA dynamics. In line with our arguments above, the weight matrix has a leading group of eigenvalues separated from the bulk, which dictate the slow dynamics, although in this case the gap is not as large and the slow eigenvalues are not as tightly bunched (Figure 3.10B). These separated eigenvalues correspond to a subspace spanned by eigenmodes with a slow spatial variation along the 'backbone' ring, and which lead to the localization and subsequent switching of the neural activity between subgroups of neurons. The principal angle between the firing patterns and the dominant Schur vectors (Figure 3.10C) indicates that the firing patterns are again dictated by the spectral properties of the underlying small-world topology. Note, however, that due to the lack of hard boundaries in the spectral groupings, both for the eigenvalue structure and the smoothly-varying dominant Schur vectors, the SSA dynamics (Figure 3.10A), although present, is not as distinctly localized as in the examples above. This behavior is effectively a result of the heavy overlap induced by the underlying lattice-like pristine world in the

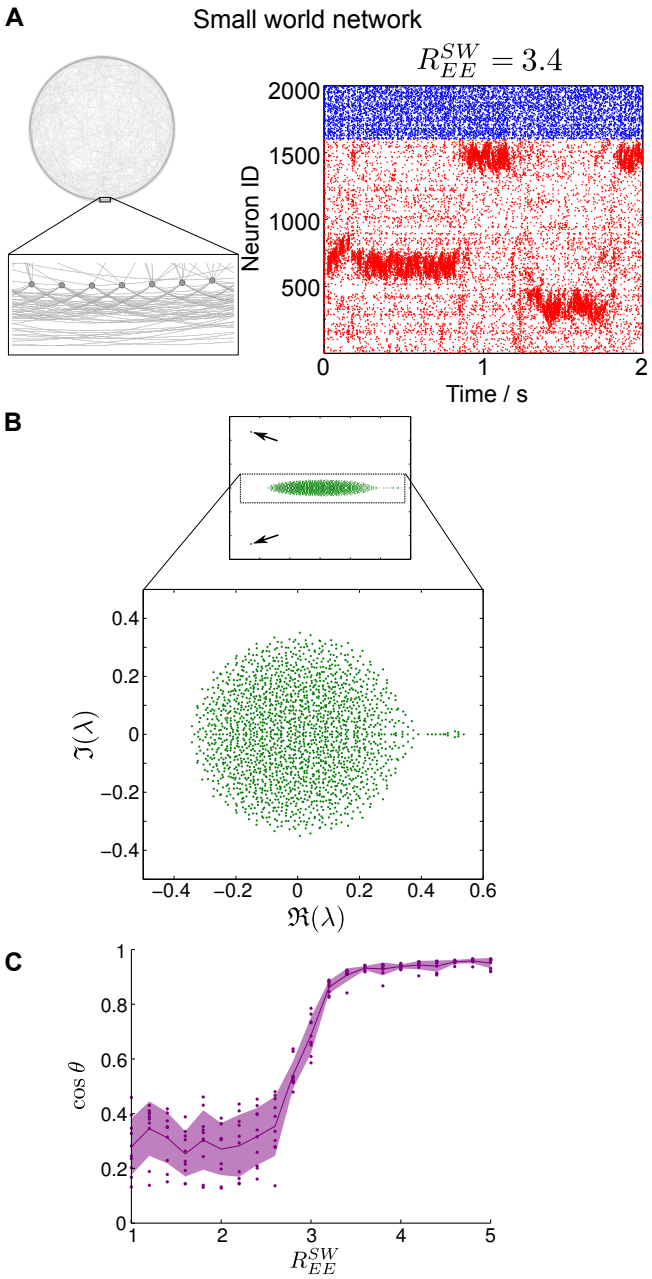

Figure 3.10: Small-world networks can exhibit SSA activity. **A** Small-world network and the resulting raster plot of its LIF network simulation. SSA activity is observed, yet with less distinct boundaries when compared to the clustered case. **B** Spectrum of the weight matrix of the small-world network showing a small eigenvalue gap. **C** The principal angle between the dominant subspace spanned by principal components of the firing rates and the dominant Schur vectors of the weight matrix shows that, as the modularity of the small-world rewiring becomes larger, there is alignment between the observed dynamics and the slow directions of the weight matrix (dots: raw data from simulations; line: mean; shading: standard deviation).

small-world construction [Barahona and Pecora, 2002].

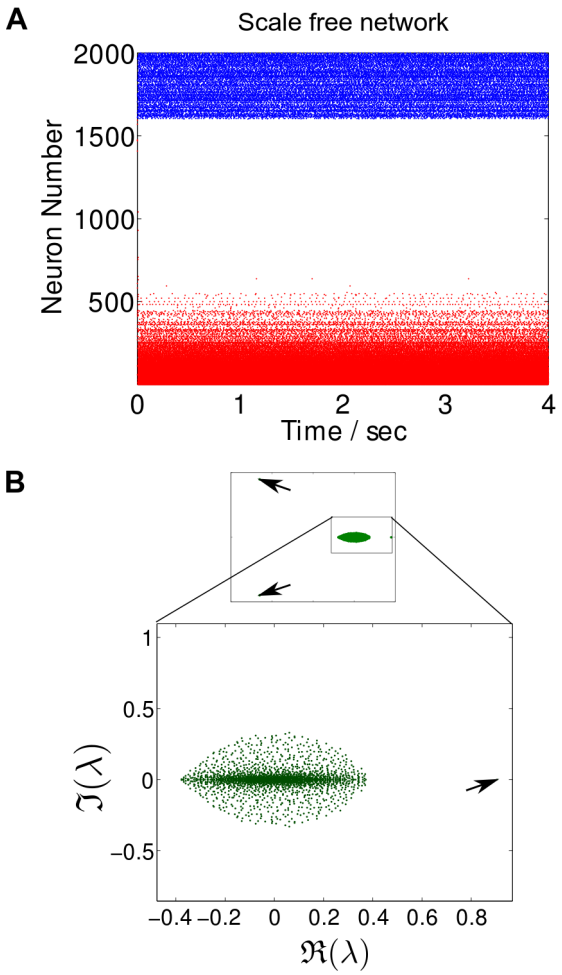

Figure 3.11: Network dynamics for a scale-free network. **A** Raster plot for an example scale-free network. Excitatory neurons are ordered according to their (expected) degree: high (bottom) to low (top). Note that the spiking activity is largely concentrated around the hub (which due to assortativity is connected to other nodes of high degree), and there are no distinguishable groups or switching behavior. **B** Spectrum of the simulated scale-free network. There is only one eigenvalue separated from the main bulk of the spectrum due to the presence of the large hub in the network.

3.3.4.2 Lack of SSA dynamics in scale-free networks

Another class of networks that has attracted tremendous interest is so-called scale-free networks [Barabsi and Albert, 1999]. These networks are characterized by a fat-tailed degree distribution, a fact that has been observed empirically for a variety of networks. As shown for undirected random graphs, the spectrum of scale-free networks may have a spectral gap, due to the presence of hubs with very large degrees [Nadakuditi and Newman, 2013]. However, the associated eigenvectors are usually localized around the hubs, and as such are not able to trigger a consistent pattern of localized activity across a group of nodes. In our simulations of scale-free LIF networks, we did not observe SSA dynamics:

the dynamics were essentially concentrated around the hub with no transitions in time (*i.e.*, only a core of neurons around the hub had consistently elevated firing rates), and hence no SSA. Indeed, the networks had only one large eigenvalue separated from the bulk (due to the high degree of the main hub) and its associated Schur vector was highly localized around the hub [Nadakuditi and Newman, 2013]. An example of this hub-centric dynamics can be found in Figure 3.11.

3.3.4.3 SSA dynamics with multiple time scales in LIF networks

It is also possible to enforce hierarchical arrangements of the functional wiring mechanisms discussed so far, thus allowing for a variety of combinatorial arrangements. Depending on the number of hierarchical layers, this enables the introduction of *multiple time scales* in the spatio-temporal segregated activity. As an illustrative example, Figure 3.12 shows the dynamics of a LIF network with a two-level hierarchy of clustered excitatory neurons. This network exhibits SSA with two 'slow' time scales: a very slow switching between the groups of the top level of the hierarchy, and the not-so-slow pulsation of activity between the subgroups within each of the large groups, thus reflecting the second level of the hierarchy. We remark that one could have employed a combination of different 'functional circuits' in the individual layers of the hierarchy leading to potentially different behaviors. As real neural networks display multiple layers of hierarchical organization [Ambrosingerson et al., 1990; Felleman and Van Essen, 1991; McGinley and Westbrook, 2013; Savic et al., 2000; Shimono and Beggs, 2014], understanding such schemes is of great importance for neural computation, and will be addressed in future work.

3.4 Material and Methods

Simulations were performed in MATLAB (2012b or later) and code implementing the networks and analysis can be found at: `github.com/CellAssembly/SSA-Dynamics`

3.4.1 Leaky integrate-and-fire networks

We simulated leaky-integrate-and-fire (LIF) networks where the non-dimensionalized membrane potential of each neuron $(V_i(t), i = 1, \ldots, N)$ was modeled by:

$$\frac{dV_i(t)}{dt} = \frac{1}{\tau_m}(\mu_i - V_i(t)) + \sum_j W_{ij} g_j^{E/I}(t), \tag{3.14}$$

with a firing threshold of 1 and a reset potential of 0. The constant input terms μ_i were chosen uniformly in the interval $[1.1, 1.2]$ for excitatory neurons, and in the interval $[1, 1.05]$

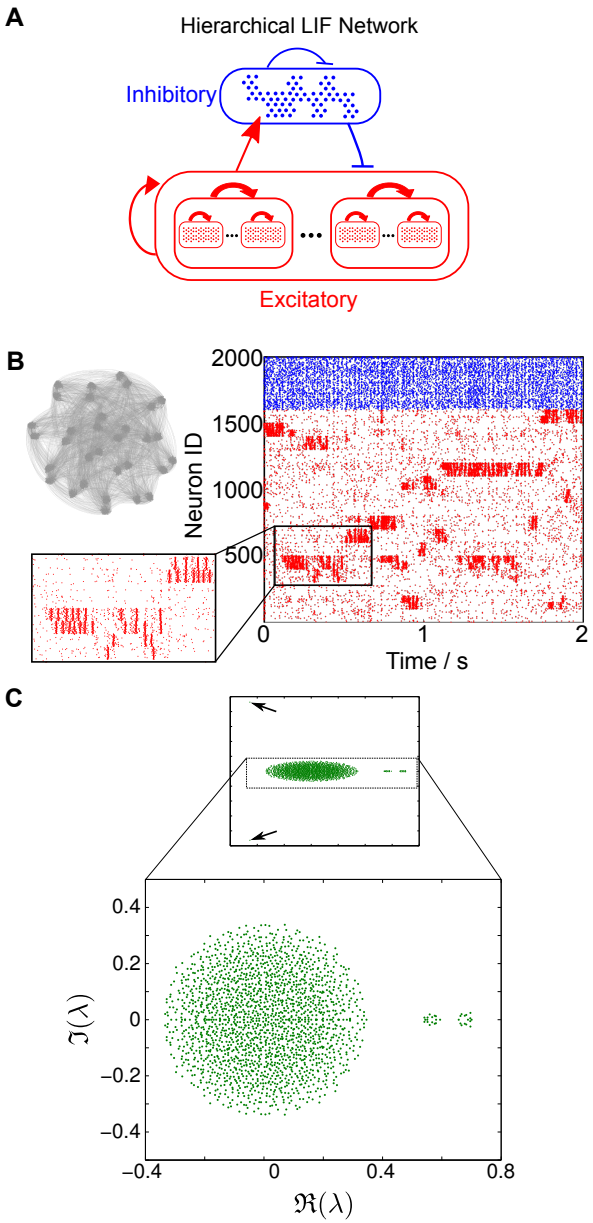

Figure 3.12: SSA activity in hierarchical LIF networks. **A** Schematic of a hierarchically clustered LIF network. Each excitatory group consists of two subgroups. **B** A hierarchical modular LIF network with $N = 2000$ neurons and the resulting raster plot of its simulation. SSA activity is observed at two time scales, corresponding to the two hierarchical levels embedded in the network structure: slow switching between the large groups and faster switching between the inner subgroups (see inset). **C** The spectrum of the weight matrix of this hierarchical network exhibits two eigenvalue gaps corresponding to the two slow time-scales the network can support.

for inhibitory neurons. The membrane time constants for excitatory and inhibitory neurons were set to $\tau_m = 15$ ms and $\tau_m = 10$ ms, respectively. The refractory period was fixed at 5 ms for both excitatory and inhibitory neurons. Note that although the constant input

term is supra-threshold, balanced inputs guaranteed that the average membrane potential is sub-threshold [Litwin-Kumar and Doiron, 2012].

The network dynamics is captured by the sum in (3.14), which describes the input to neuron i from all other neurons in the network. The topology of the network is encoded by the weight matrix W, i.e., W_{ij} denotes the weight of the connection from neuron j to neuron i, where W_{ij} is zero if there is no connection. Synaptic inputs are modeled by $g_j^{E/I}(t)$, which is increased step-wise instantaneously after a presynaptic spike of neuron j ($g_j^{E/I} \to g_j^{E/I} + 1$) and then decays exponentially according to:

$$\tau_{E/I} \frac{dg_j^{E/I}}{dt} = -g_j^{E/I}(t), \tag{3.15}$$

with time constants $\tau_E = 3$ ms for an excitatory interaction, and $\tau_I = 2$ ms if the presynaptic neuron is inhibitory.

Equation (3.14) can be rewritten in matrix notation as:

$$\frac{d\mathbf{V}(t)}{dt} = T^{-1}[\boldsymbol{\mu} - \mathbf{V}] + W\mathbf{g}(t), \tag{3.16}$$

where $T = \mathrm{diag}(\tau_i)$ and $\mathbf{V} = [V_1, \ldots, V_N]^T$ and the vectors $\mathbf{V}, \boldsymbol{\mu}$, and $\mathbf{g}$ are $N \times 1$ vectors. The spectral analyses in the main text (eigenvalues and Schur vectors) refer to the weight matrix W. Different network topologies correspond to different weight matrices W, as explained below.

In all simulations, the ratio of excitatory to inhibitory neurons was fixed to be 4 : 1. Unless otherwise stated, the networks comprise $N = 2000$ units (1600 excitatory, 400 inhibitory) and were simulated over 20 seconds to calculate the statistics reported. The time step for all simulations was 0.1 ms.

3.4.2 Network Topologies and Weight matrices

Different network topologies were simulated using different W matrices but maintaining a general balanced network. If not indicated differently, all parameters below correspond to the case $N = 2000$.

3.4.2.1 Networks with unclustered balanced connections

Unless otherwise stated, unclustered connections between neurons were drawn uniformly at random according to the probabilities $p^{EE} = 0.2$ and $p^{EI} = p^{IE} = p^{II} = 0.5$, where the first superscript denotes the destination and the second superscript denotes the origin of the synaptic connection, and E, I stand for an excitatory or inhibitory neuron, respectively.

Synaptic weights between inhibitory neurons had always the weight $w^{II} = -0.0297$. The other weight parameters were set to $w^{EI} = -0.0297$, $w^{IE} = 0.0074$ and $w^{EE} = 0.0156$, except where indicated differently in the scenarios described below.

3.4.2.2 Networks with clustered excitatory-to-excitatory connections

First, LIF networks with clustered excitatory neurons were constructed according to the protocol in Ref. [Litwin-Kumar and Doiron, 2012]. The connection probabilities between excitatory neurons were changed so that neurons had more connections within its group than to neurons outside the group. The average number of connections was kept the same as for an unclustered balanced network ($p^{EE} = 0.2$) while varying the ratio $R_{EE} = p_{in}/p_{out}$, where p_{in} and p_{out} are the probabilities of connectivity within a group and between groups, respectively. $R_{EE} = 1$ corresponds to the unclustered balanced network.

Second, we generated LIF networks where the connectivity between all excitatory neurons was uniform ($p^{EE} = 0.2$, $R_{EE} = 1$) but the weights were varied by changing the ratio $W_{EE} = w_{in}^{EE}/w_{out}^{EE}$, where w_{in}^{EE} refers to synaptic weights within the same group and w_{out}^{EE} refers to synaptic weights to neurons in other groups. The average synaptic weight between excitatory neurons was kept at $w^{EE} = 0.0156$, as for the unclustered case. Strictly speaking, this scheme does not correspond to a clustering in a topological sense, but rather to a clustering of synaptic weights between excitatory neurons. However, in terms of the dynamics, both E-E clustered network variants (adjusting probabilities or weights) lead to effectively the same results, as discussed in the text. Unless otherwise stated, the excitatory neurons were divided into 20 groups of 80 neurons each.

As the weights in balanced networks should scale with $1/\sqrt{N}$ [Litwin-Kumar and Doiron, 2012; van Vreeswijk and Sompolinsky, 1998], when varying the network size (see Fig. 3.2A) the parameter settings were kept the same but all weights were scaled accordingly, e.g., for $N = 1000$ the weights were multiplied by $\sqrt{2}$ compared to those of the network with $N = 2000$.

3.4.2.3 Networks with excitatory-to-inhibitory feedback loop

To construct LIF networks with co-clustered excitatory and inhibitory neurons, we kept uniform excitatory-to-excitatory and inhibitory-to-inhibitory couplings. The structure was imposed by varying the ratios $W_{IE} = w_{in}^{IE}/w_{out}^{IE}$, $W_{EI} = (w_{in}^{EI}/w_{out}^{EI})^{-1}$, or the analogous ratios of connections probabilities R_{IE}, R_{EI}. The subscript 'in' indicates connections within the excitatory/inhibitory pair. Note that W_{EI}, R_{EI} are defined by the inverse in/out ratio, since they describe an inhibitory effect. To modulate the co-clustered excitatory-to-inhibitory network dynamics, we vary the ratios W_{IE}, W_{EI}, while keeping the average

weights constant. For instance, in the example shown in Figure 3.9, we kept $R_{EI} = R_{IE} = 2$ while simultaneously varying the ratios $W_{IE} = W_{EI}$ between 1 and 5. For all networks, we divided the network into 20 groups with 80 excitatory and 20 inhibitory neurons each.

3.4.2.4 Networks with small-world connectivity

Networks with small-world connectivity between excitatory neurons were constructed as follows. Excitatory neurons ($N = 1600$) were connected within a 40-nearest neighbor ring with probability p_{in}^{EE}, i.e., neurons i was connected to neuron j with probability p_{in}^{EE} if $|i - j| \leq 40$ subject to periodic boundary conditions. The connection probability outside this neighborhood was set to p_{out}^{EE}. As in previous models, the average connectivity was kept constant at $p^{EE} = 0.2$ while varying the ratio $R_{EE}^{SW} = p_{in}^{EE}/p_{out}^{EE}$. Hence, by increasing R_{EE}^{SW} we increase the amount of 'backbone' connectivity through the (stochastic) 40 nearest-neighbor cycle, while decreasing the number of connections to neurons elsewhere in the network. This construction may thus also be interpreted as an overlapping clustering and is effectively a variant of the small-world scheme introduced by Watts and Strogatz [Watts and Strogatz, 1998].

3.4.2.5 Networks with scale-free connectivity

LIF networks with a scale-free topology for the excitatory connections were created by a preferential attachment process [Barabsi and Albert, 1999] using algorithm 5 in [Batagelj and Brandes, 2005] with parameter $d = 64$. The weight matrix was created iteratively by adding one neuron at a time, such that the probability that it is connected to an existing node is proportional to the degree of that node (at the current state), *i.e.*, the earlier a node is added, the larger its final degree will be. The resulting degree of the network tends to a power-law distribution [Barabsi and Albert, 1999]. The value $d = 64$ was chosen to avoid saturation of the dynamics of the network, and this leads to an excitatory-excitatory matrix with 35% of the edges of our other examples.

3.4.2.6 Network with hierarchical excitatory connections

We simulated LIF networks where excitatory neurons belonged to a hierarchy of groups (Figure 3.12) by dividing the excitatory neurons into nested clusters. We implemented two layers in the hierarchy: the excitatory neurons were divided into 16 groups, and each of these groups was then subdivided into 2 more groups to give a total of 32 subgroups. This was achieved by varying the ratios $R_{EE}^{top} = p_{\text{group}}^{EE}/p_{out}^{EE}$, $R_{EE}^{sub} = p_{\text{subgroup}}^{EE}/p_{\text{group}}^{EE}$ while keeping the average connectivity between excitatory neurons constant. The example in Figure 3.12

corresponds to $R_{EE}^{top} = 1.45$ and $R_{EE}^{sub} = 3.7$. The weight within the subgroups was set at $w_{sub}^{EE} = 0.0163$. All other connections remained unchanged.

3.4.3 Quantifying SSA dynamics from spike-train LIF simulations

To evaluate the extent to which a network displays slow-switching assembly dynamics, we used the following two spike-rate variability measures.

Given a partition of the neurons into c groups, we compute the average spiking frequency $f_i(t)$ of the neurons in each cluster $i \in \{1, \ldots, c\}$ over non-overlapping windows of 100ms. For each time window, we obtain the firing rate vector $\mathbf{f}(t) = [f_1(t), \ldots, f_c(t)]^T$ of which we compute the standard deviation $\sigma(t)$. The standard deviations are then averaged over the duration of the simulation:

$$S = \frac{1}{T} \sum_{t=1}^{T} \sigma(t), \tag{3.17}$$

where T is the total number of time windows in the simulation. We then obtain a boot-strapped expectation $\langle S^{\text{shuff}} \rangle$ computed by reshuffling neurons at random into groups of the same sizes as those in the partition. The spike rate variability score is then defined as:

$$\widehat{S} = S - \langle S^{\text{shuff}} \rangle, \tag{3.18}$$

where the average $\langle S^{\text{shuff}} \rangle$ is obtained over 10 random reshufflings of the neurons. If the network fires uniformly (with no localized patterns), $\widehat{S}$ is low; whereas $\widehat{S}$ increases if the network displays heterogeneous activity aligned with the partition under investigation.

As explained in the text, we introduced a second spike-train variability measure to quantify the variations of the group firing-rates across time. This complementary measure allows us to discern scenarios in which there is a group of neurons dominating the firing (thus leading to a large variation across groups), but no switching between groups. To quantify these effects we use an analogous measure to $\widehat{S}$ above.

Given a partition of the neurons into c groups, we compute the average spiking frequency $f_i(t)$ of the neurons in each cluster $i \in \{1, \ldots, c\}$ over non-overlapping windows of 100ms. For each group i, we compute the standard deviation $\sigma_T(i)$ of the vector of coarse-grained firing rates across time $\mathbf{f_i} = [f_i(t_1), \ldots, f_i(t_N)]$, where t_k stands for the kth time bin and $i = 1, \ldots, c$. We then average over groups to obtain:

$$S_T = \frac{1}{c} \sum_{i=1}^{c} \sigma_T(i). \tag{3.19}$$

As above, we obtain a boot-strapped expectation $\langle S_T^{\text{shuff}} \rangle$ by reshuffling neurons at random

into groups of the same sizes as those in the partition. The spike rate variability score over time is then defined as:

$$\widehat{S}_T = S_T - \langle S_T^{\text{shuff}} \rangle, \tag{3.20}$$

where the average $\langle S_T^{\text{shuff}} \rangle$ is obtained over 10 random reshufflings of the neurons. Again, if the network fires homogeneously (with no localized patterns in time), $\widehat{S}_T$ is low, whereas $\widehat{S}_T$ increases if the network displays heterogeneous firing rates over time.

In Figure 3.3, we show the behavior of $\widehat{S}_T$ for the examples discussed above. As expected, for all cases where SSA dynamics is present, both measures $\widehat{S}$ and $\widehat{S}_T$ behave consistently. On the other hand, $\widehat{S}_T$ detects the end of the SSA region when, through increased clustering, the dynamics gets localized on one cell assembly, as shown in Figure 3.7.

3.4.4 Measuring alignment of LIF dynamics with the Schur vectors of the weight matrix: the principal angle

First, we find the dominant firing patterns in the network dynamics. We perform simulations of the LIF network and obtain the firing rates of every neuron in 250 ms bins to generate an $N \times T$ matrix, where N is the number of neurons and T is the number of bins. On this matrix, we perform a *principal component analysis* (PCA) and select the first $c-1$ principal components $\{\mathbf{p_i}\}_{i=1}^{c-1}$. This set of N-dimensional vectors captures most of the variability observed in the simulated dynamics.

We then assess how aligned the $c-1$ principal components are with the dominant Schur vectors of the weight matrix of the network. More precisely, we compute the Schur decomposition of the weight matrix W; keep the $c-1$ dominant Schur vectors $\{\mathbf{u_i}\}_{i=1}^{c-1}$ associated with the eigenvalues with largest real part; and then compute the (first) principal or canonical angle θ between the subspaces spanned by the two sets of vectors $\mathcal{P} = \text{span}\{\mathbf{p_i}\}$ and $\mathcal{U} = \text{span}\{\mathbf{u_i}\}$ [Golub and Van Loan, 1996; Stewart, 2001]:

$$\cos(\theta) = \max \left\{ \frac{\mathbf{u}^T \mathbf{p}}{\|\mathbf{u}\| \|\mathbf{p}\|} \middle| \, \mathbf{u} \in \mathcal{U} \;\; \mathbf{p} \in \mathcal{P} \right\}, \tag{3.21}$$

The first principal angle measures how 'close' the observed firing patterns are to the dominant modes computed solely from the weight matrix. If $\cos(\theta) \approx 1$, there is a large alignment between the span of both subspaces.

3.5 Discussion

Answering the question of how the wiring of neural circuits governs neural dynamics is a key step in understanding neuronal computations. Here, we showed how knowledge of certain spectral features of the matrix of neuronal connectivity can help understand what dynamics a network can support.

Herein we focused on LIF networks exhibiting slow-switching assembly (SSA) dynamics, where groups of neurons show a sustained and coherent increase in their firing rates, with slow switching between epochs of localized firing in different groups across the network. We found that the presence of the slow switching time scale is reflected in the spectral properties of the synaptic weight matrix: a gap separating the leading eigenvalues together with a block-localization of the associated Schur vectors on groups of neurons is a key indicator of the presence of SSA dynamics in the network. In line with this observation, multiple gaps in the eigenvalue spectrum are indicative of further time scales in the network dynamics (see Figure 3.12). Moreover, when the leading eigenvalue becomes larger than a critical value, the SSA dynamics becomes localized on one of the assemblies (see Figure 3.7).

First, we revisited the case of balanced LIF networks with clustered excitatory neurons [Litwin-Kumar and Doiron, 2012] and observed that only when there is an eigenvalue gap and Schur block-localization does the network display SSA dynamics. Further analytical understanding from a stylized firing-rate model allowed us to determine that the clustering strength drives the development of the eigenvalue gap responsible for the slow switching between localized firing modes. We remark that clustered excitatory connectivity leads to a Hebbian amplification regime [Murphy and Miller, 2009]: the more stable the pattern formation, the slower the dynamics of any banded activity. Fast switching between well-defined, stable patterns is thus only achievable for strong input changes.

As suggested from our spectral characterization, and confirmed by simulations, we showed that SSA dynamics can be achieved through the structured tuning of synaptic strengths, rather than through the clustered rewiring of the connections. While the equivalence between topological and weight organizations is clear for a linear system, such a direct correspondence is not guaranteed *a priori* for a non-linear system. In fact, the clustering of weights appears to have a slightly larger influence compared to topological organization in our simulations of non-linear LIF systems. More importantly, this dynamical equivalence between clustered topology and clustered synaptic weights has potential ramifications for learning and synaptic plasticity; it indicates that the alteration of the weight structure can lead to the emergence of grouped activity without the need for structural rewiring of physical synaptic connections, thus suggesting a cost-effective adaptation to stochastically

encode different firing patterns [Laughlin and Sejnowski, 2003].

Within LIF networks with clustered excitatory connectivity [Litwin-Kumar and Doiron, 2012], inhibitory neurons play only a balancing background role, whereas experimental studies have revealed a vast diversity of interneuron subtypes that play key roles in computation [Kepecs and Fishell, 2014; Roux and Buzsaki, 2015]. We thus investigated mechanisms in which inhibitory neurons have an active, functional role in generating SSA dynamics. We demonstrated both analytically and via numerical simulations how such functional circuits can be constructed with the help of a feedback loop between excitatory and inhibitory neurons, so that inhibitory neurons themselves exhibit SSA dynamics and become an integral functional part of the dynamics. As fewer inhibitory neurons are present, adapting the weights according to this functional co-clustering scheme may also provide a cost effective alternative to generate SSA activity in neuronal networks.

The emergence of SSA dynamics can be achieved not only by clustering excitatory neurons but also by shaping the network structure in several ways. For instance, a small-world network topology of the excitatory neurons is also able to support SSA dynamics due to the spectral properties of small-worlds, which induce a separation of the slow eigenvalues and the localization and switching of firing activity associated with slowly spatially-varying dominant Schur vectors. In contrast, scale-free networks lack the block-localization of the Schur vectors on multiple groups which is necessary to consistently induce SSA dynamics in LIF networks. Further topologies will be explored in the future. In particular, inhibitory neurons that inhibit other inhibitory neurons (so-called disinhibition patterns [Kepecs and Fishell, 2014; Lee et al., 2013; Pfeffer et al., 2013; Pi et al., 2013; Silberberg and Markram, 2007; Xu et al., 2013]) provide an interesting example. A different avenue might be provided by balanced amplification mechanisms [Murphy and Miller, 2009], which could be potentially used for creating grouped activity, yet without the introduction of a slow time scale. Other interesting questions in this respect are: which wiring mechanisms provide the most economical, or evolutionarily fit, variant [Laughlin and Sejnowski, 2003; Niven and Laughlin, 2008] to induce a given dynamics; and how does the observed diversity of interneurons (and the multiple roles they can play) relate to this dynamical picture.

Our work emphasizes the importance of the spectral characterization of the weight matrix for the dynamics taking place on these topologies. Although spectral properties are also key in characterizing the dynamical response of networks operating at criticality [Larremore et al., 2011], our observations here correspond to a different phenomenon. The emergence of a dominant assembly when $\lambda_{\max} \gtrsim 1$ leads to a saturation of the network and a decrease of its dynamical heterogeneity, in contrast to systems of coupled non-modular excitatory units at criticality, which maximize their dynamic range for $\lambda_{\max} = 1$ [Beggs and Plenz, 2003;

Larremore et al., 2011].

Our work links up with experimental findings that cortical states can arise spontaneously and exhibit switching behavior. Recordings from anesthetized cat visual cortex have revealed activity states that dynamically switch, and these states match closely to recorded orientation maps [Kenet et al., 2003]. Such cortical states likely arise due to similarly tuned neurons having higher intra-cortical connectivity [Harris and Mrsic-Flogel, 2013; Ko et al., 2011], as is also predicted through models of orientation maps [Ben-Yishai et al., 1995; Ernst et al., 2001; Somers et al., 1995]. These observations are similar to our theoretical outcomes which indicate that dynamic transitioning between different network states can be driven and sustained based only on the underlying network topology. Other experiments have identified states in CA3 network activity of the hippocampus [Sasaki et al., 2007] where active cell assemblies exist for tens of seconds before sharply transitioning to a new state. However, these were metastable states (hence unlikely to be reactivated) and included a core population of cells consistently active in all states. While such complex dynamics are beyond the simple models presented here, it may be feasible to model such metastable dynamics through more elaborate network topologies, which may include synaptic plasticity [Litwin-Kumar and Doiron, 2014]. Finally, let us remark that our reduced model descriptions effectively focused on networks implementing a *rate-based* coding mechanism and do not include spike-timing of cell assemblies, which has been identified to be an important component in neuronal computation, e.g. in rodents [Foster and Wilson, 2006; Harris et al., 2003], monkeys [Riehle et al., 1997], songbirds [Hahnloser et al., 2002], and grasshoppers [Rokem et al., 2006]. Whether our results can be translated to a time-coding regime would be an interesting question for future work.

As connectomics continues to advance the mapping of connections and synaptic strengths in wide areas of the brain, experimentally obtained weight matrices can be analyzed spectrally as described above to determine if SSA dynamics can be supported by the networks under study. While we focused here on the simplest mechanisms underlying SSA activity, future work could also consider the construction of biophysically realistic models that take into account the growing literature on the distribution of synaptic strengths, synaptic contacts, firing rates, and other relevant cortical parameters that tend to show a lognormal distribution [Buzsaki and Mizuseki, 2014; Hromadka et al., 2008; Oh et al., 2014; Song et al., 2005].

Although knowledge of the network structure (connectomics) is not sufficient to predict the dynamics a network circuit will display (for instance, the network dynamics can be dominated completely by a strong input), it can still give valuable insights on the firing patterns the network can support. Conversely, the observation of neuronal dynamics alone

may not be sufficient to understand neural computation in detail, since different network topologies can yield similar dynamics. Hence, our work hints at how connectomics and neuronal dynamics data can provide complementary and intertwined routes for systems neuroscientists to study the computational principles implemented by the brain.

Finally, it is important to remark that in order to get a fuller picture of the relation between structure and dynamical network properties many other factors not considered here will be of interest, including the distribution of inputs, the precise location of synapses on the post-synaptic neuron, and the plasticity rules that govern the evolution of the network over time [Clopath et al., 2010]. Linking the insight gained from our work to experimental observations or to highly detailed computational models [Ahrens et al., 2013; Reimann et al., 2013] would thus be a fruitful next step in order to bridge the gap between neuronal structure, dynamics, and ultimately function.

Chapter 4

Feedforward architectures driven by inhibitory interactions

4.1 Abstract

Directed information transmission is a paramount requirement for many social, physical, and biological systems. For neural systems, scientists have studied this problem under the paradigm of feedforward networks for decades. Here we illustrate that one can construct such networks even if the connectivity between the excitatory units in the system remains random. This is achieved by endowing inhibitory nodes with a more active role in the network. Our findings demonstrate that feedforward activity can be caused by a much broader network-architectural basis than often assumed.

4.2 Introduction

The ability to reliably propagate signals in a targeted manner is essential for the operation of many natural systems. A prototypical model for such a targeted information transmission within a neural substrate are feedforward networks, which have been considered in the literature for decades. In these models, the basic paradigm is to group nodes (neurons) into separate layers, each of which receives excitatory input from the preceding layer, and projects excitatory connections to the subsequent layer [Kumar et al., 2010; Vogels et al., 2005]. The thus established forward-directed excitatory pathways guide the activity sequentially through the layers. Inhibitory units (neurons) play merely a balancing role: they ensure that the network remains stable, either separately for every layer or globally. A large number of variations of this scheme have been considered, such as embedding feedforward architectures in randomly connected networks to examine their effect on the overall network dynamics and signal propagation [Kumar et al., 2008; Mehring et al., 2003]. Furthermore,

feedforward networks have been shown to propagate firing rates [van Rossum et al., 2002; Vogels and Abbott, 2005], synchrony/pulse packets [Aertsen et al., 1996; Cateau and Fukai, 2001; Diesmann et al., 1999; Gewaltig et al., 2001; Kistler and Gerstner, 2002; Litvak et al., 2003], combinations of firing rates and synchronous spiking [Kumar et al., 2010], and are even able to gate activity transmission [Vogels and Abbott, 2009].

Interestingly, recent work in neuroanatomy has revealed an enormous diversity of inhibitory neurons [Bortone et al., 2014; Harris and Shepherd, 2015; Huang, 2014; Isaacson and Scanziani, 2011; Kepecs and Fishell, 2014; Klausberger and Somogyi, 2008; Olsen et al., 2012; Roux and Buzsaki, 2015; Taniguchi, 2014]. Moreover, specific plasticity rules for different subtypes of inhibitory neurons [Chen et al., 2015] add further to their diverse and heterogeneous connectivity profiles. In this light it would be strongly surprising if inhibitory neurons serve the cortex only in a homogeneous, passive role when it comes to information propagation, as is assumed in most feedforward networks. While some studies exist that specifically account for intra- and interlayer inhibition [Aviel et al., 2003; Kremkow et al., 2010; Mehring et al., 2003; Tetzlaff et al., 2003; Vogels and Abbott, 2005], information is still mediated by a cascade of excitatory neurons. However, as experimental evidence suggests, the nervous system likely uses a combination of methods to transmit information [Reyes, 2003; Vincent et al., 2012]. An important natural question is thus if one can construct networks in which there are no preferred excitatory-to-excitatory pathways, but *inhibitory neurons* play a pivotal role for the propagation of activity between layers. In the following we demonstrate, via numerical simulations and brief analytical considerations, that such feedforward processing is indeed possible.

4.3 Materials and Methods

To illustrate our ideas, we have used leaky-integrate-and-fire (LIF) networks, stylized models of neural networks, which act like pulse-coupled oscillators. Using a time step of 0.1ms we numerically integrated the non-dimensionalized membrane potential of each neuron, which evolved according to:

$$\frac{dV_i(t)}{dt} = \frac{1}{\tau_m}(\mu_i - V_i(t)) + \sum_j W_{ij} g_j^{E/I}(t), \tag{4.1}$$

with a firing threshold of 1 and a reset potential of 0. All networks comprised $N = 2000$ units, with an excitatory to inhibitory neuron ratio of $4:1$ (1600 excitatory, 400 inhibitory). The input terms μ_i were chosen uniformly in the interval $[1.1, 1.2]$ for excitatory neurons, and in the interval $[1, 1.05]$ for inhibitory neurons. Membrane time constants for excitatory

and inhibitory neurons were set to $\tau_m = 15$ ms and $\tau_m = 10$ ms, respectively, and the refractory period was fixed at 5 ms for both excitatory and inhibitory neurons. Note that although the constant input term is supra-threshold, balanced inputs guaranteed an average sub-threshold membrane potential [Litwin-Kumar and Doiron, 2012; Schaub et al., 2015].

In the model, the network coupling is captured by the sum in (4.1), which describes the input to neuron i from all other neurons in the network. Here W_{ij} denotes the weight of the connection from neuron j to neuron i ($W_{ij} = 0$ if there is no connection). After a presynaptic spike of neuron j, the synaptic inputs $g_j^{E/I}(t)$ are increased step-wise ($g_j^{E/I} \rightarrow g_j^{E/I} + 1$) instantaneously, and then decay exponentially according to:

$$\tau_{E/I} \frac{dg_j^{E/I}}{dt} = -g_j^{E/I}(t), \tag{4.2}$$

with time constants $\tau_E = 3$ ms for an excitatory interaction, and $\tau_I = 2$ ms if the presynaptic neuron is inhibitory. For all networks described in the following, the total connection-strength per neuron was kept equivalent to an unstructured, balanced network displaying asynchronous activity, with $p_{EI} = p_{IE} = p_{II} = 0.5$, $p_{EE} = 0.2$, $w_{EI} = w_{II} = -0.042$, $w_{IE} = 0.0115$, and $w_{EE} = 0.022$. Here, p and w stand for the connection probability and connection weight, respectively. The first subscript denotes the destination and the second superscript denotes the origin of the synaptic connection, and E,I stand for an excitatory or inhibitory neuron, respectively. All simulations were performed in MATLAB (2012b or later) and the code can be found at: `github.com/CellAssembly/inhibitory-feedforward`.

4.4 Results

4.4.1 Cross-coupled feedforward networks

In a recent experimental study of the hippocampus [Nasrallah et al., 2015] it was demonstrated that excitatory neurons in the brain region CA3 are unable to drive excitatory neurons in area CA2 due to strong feedforward inhibition. However, when this strong inhibition of CA2 is alleviated, CA3 can indeed excite CA2 excitatory cells to elicit action potentials [Nasrallah et al., 2015] (see Figure 4.1a). An interesting feature of this finding is that the directionality in the interaction between CA2 and CA3 appears to be dictated by the connections between excitatory and inhibitory neurons, rather than a consequence of unidirectional excitatory connections targeting CA2. Indeed, excitatory connections between CA2 and CA3 are reciprocal [Kohara et al., 2014; Llorens-Martin et al., 2014] – yet there is still a *directed* propagation towards CA2 [Nasrallah et al., 2015]. Stated differently, the targeted activation of CA3 is controlled by an excitatory-inhibitory-excitatory pathway.

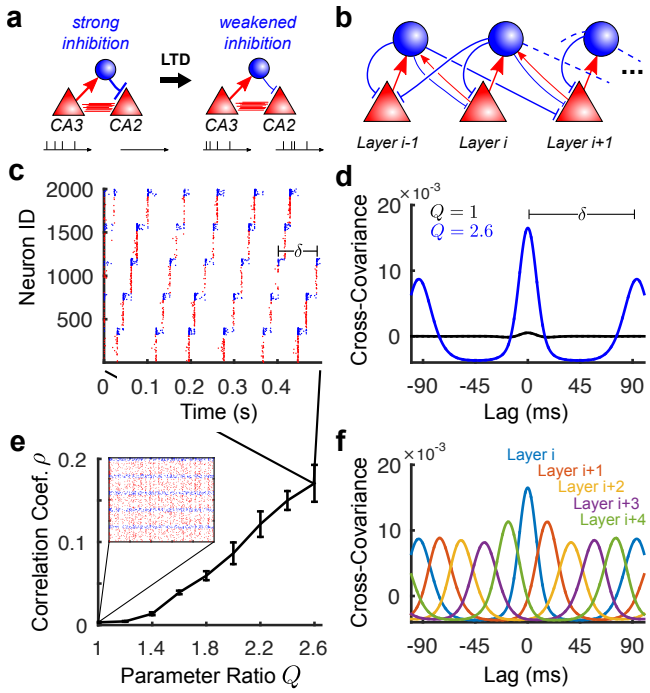

Figure 4.1: **Cross coupled feedforward networks**. **(a)** Schematic of the finding of Ref. [Nasrallah et al., 2015]. Inhibitory long term depresssion (LTD) lowers the feedforward inhibition of CA2, allowing information transfer from CA3 to CA2. **(b)** Schematic of proposed network architecture. Note that only connections that are not uniformly distributed in the network are displayed for visual clarity. The feedback between excitatory and inhibitory neurons drive the feedforward activity, while connections between alike neurons remain uniform (see text). **(c)** Example raster plot with 5 groups that show the propagation of activity between layers. Observe that inhibitory neurons also show feedforward propagating activity. **(d)** Group-averaged cross-covariance for parameter ratio values $Q = 1$ and $Q = 2.6$. For $Q = 1$ (no feedforward structure) the firing is clearly not synchonous and does not display any pattern. In contrast, with imposed cross-coupled feedforward structure ($Q = 2.6$), there is a peak indicating strong synchronous firing inside each layer. The second peak at time $\pm\delta$ indicated the periodically repeating firing pattern (see also **(c)**). **(e)** The average Pearson correlation coefficient within layers as a function of Q. The larger the feedforward ratio, Q, the greater the correlation of firing within layers. Error bars correspond to standard deviation. **(f)** Average cross-covariance of the neural firing patterns in layer i with neurons in all other layers j for $Q = 2.6$. Observe that the time lag of the peaks are arranged consecutively, illustrating the orderly feedforward progression between groups.

We sought to leverage this targeted activation mechanism by cascading this connection motif. The result is a circuit with *uniform* connectivity among excitatory neurons with *no preferred direction*, which is nevertheless able to propagate feedforward activity due to the specific cross coupling between the excitatory and inhibitory neurons. We term this circuitry a cross-coupled feedforward network (ccFFN). A schematic can be found in Figure 4.1b.

The behavior of ccFFNs can be explained by the following rationale: (i) the excitatory neurons in each layer are more strongly coupled to the group of inhibitory neurons in their own layer relative to other inhibitory neurons; (ii) an activity increase of such an excitatory

group thus triggers elevated activity in the corresponding inhibitory neurons; (iii) this inhibitory group of neurons targets the subsequent layer of excitatory neurons more weakly relative to other excitatory neurons; (iv) the reduced inhibition (relative) of the subsequent excitatory group leads to increased excitatory activity in the subsequent layer, while the activity in the initial layer returns to baseline; (v) by cascading this cross-coupling motif, elevated activity of excitatory neuron groups propagates through the circuit.

A simulation of a network with such a ccFFN topology is shown in Figure 4.1c. Note that, to eliminate transient effects from particular driving inputs we connected the last with the first layer, thus establishing a circular pathway with a self-sustained forward propagation of activity along the layers. Importantly, in addition to the propagation of the excitatory activity, we observe that the *inhibitory* neurons' activity progresses from one group to the next. This emphasizes the pivotal role played by inhibitory neurons for the observed dynamics, which is clearly beyond simply balancing the network.

To construct ccFFN networks we kept excitatory to excitatory and inhibitory to inhibitory connections uniform as outline above. We divide the network into layers consisting of both inhbitory and excitatory units and connected as outlined in Figure 4.1b: excitatory neurons in layer i are statistically biased to target the inhibitory neurons in their own layer with a weight ratio $W_{IE} = w^i_{IE}/w^{not[i]}_{IE}$, compared to the inhibitory neurons in the rest of the network. Similarly, the inhibitory neurons within a layer i target the excitatory neurons in the *next* layer $i+1$, more weakly according to the ratio $W_{EI} = (w^{i+1}_{EI}/w^{not[i+1]}_{EI})^{-1}$. In addition to modifying the weights, we control the analogous ratio of connections probabilities R_{IE} and R_{EI}. Note that W_{EI}, R_{EI} are defined with an inverse ratio, *i.e.*, a higher ratio means a *weaker* targeting corresponding to a stronger feedforward structure. To modulate the embedded feedforward level in the networks, we can thus vary the ratios $W_{IE}, W_{EI}, R_{IE}, R_{EI}$, while keeping the average weights and number of connections constant. For simplicity, in the following we kept all four ratios equal to a fixed value Q modulating the overall feedforward structure. Note, however, that feedforward activity can be observed by changing only the weights or the connectivity probabilities separately (for a related observation, see [Schaub et al., 2015]).

The network shown in Figure 4.1c consists of 5 layers of neurons with a feedforward ratio of $Q = 2.6$. To illustrate that increasing Q indeed results in an increased feedforward activity we calculated the cross-covariance (Fig. 4.1d) and the Pearson correlation coefficient (Fig. 4.1e), between the firing patterns of the neurons, averaged over the different layers.

Figure 4.1d shows the average cross-covariance functions within the same layer for the two conditions of $Q = 1$ and $Q = 2.6$, averaged over 10 realization of the network. To get a smooth estimated, we convolved the spike-train of every neuron with a Gaussian signal of

standard deviation $5ms$. For every neuron pair, the convolved signal ($f_i(t)$) was then used to calculate the pairwise cross-covariance $\phi_{ij} = \mathrm{cov}[f_i(t+\tau), f_j(t)]$:

$$\phi_{ij}(\tau) \approx \int [f_i(t+\tau) - \mu(f_i)][f_j(t) - \mu(f_j)]dt, \tag{4.3}$$

which we averaged over all neurons inside the same layer: Here $\mu(\cdot)$ denotes the mean of the signal. While there is no apparent temporal structure in the networks with $Q = 1$, there is a clearly visible increased synchrony in the networks with high feedforward ratio $Q = 2.6$, as indicated by the large peak at zero lag. Moreover peaks appearing at a lag of $\pm\delta$ corresponding to the repetition period of the firing, resulting from the the circular topology.

To investigate the tendency for each layer to fire in unison further we computed the Pearson correlation coefficient of the convolved spike-trains of all neuron pairs for varying levels of Q. We plot the average correlation coefficient within each group and layer in Figure 4.1e. For each value of Q, 10 network realizations were simulated. As can be seen, the larger the value of Q, the more synchrony there is within groups. This synchronous firing of groups is not decoupled but propagates along layers, as can be seen in Figure 4.1f. There we plot the average cross-covariance of a layer relative to all other layers (c.f 4.1d). As the regular shifts in the cross-covariance indicate, there is indeed a clear consecutive progression of activity from one layer to the next. We can thus conclude that the ratio Q directly influences the feedforward activity propagation in ccFFNs, demonstrating that directed information transmission is possible without an imposed excitatory-to-excitatory pathway in the network.

Interestingly, there exist some further experimental indications that a similar mechanism to direct activity might be implemented in canonical cortical microcircuits [Pluta et al., 2015]. The traditional view of these ubiquitous circuits is that inputs from layer 4 (L4) drive layers 2/3 (L2/3) that then excites layer 5 (L5). However, as Pluta and coworkers [Pluta et al., 2015] have shown in their work, the picture is likely to be more intricate: in particular it appears that L4 first suppresses L5 while driving L2/3, and only afterwards L5 shows elevated activity Pluta et al. [2015] — a finding that shows parallels to our proposed mechanism.

4.4.2 Disinhibitory feedforward networks

The aforementioned cross-coupling of excitatory and inhibitory neurons is not the only arrangement possible to create targeted feedforward activity driven by inhibitory units. To illustrate that our finding is general and should not be reduced to a single type of circuitry, we

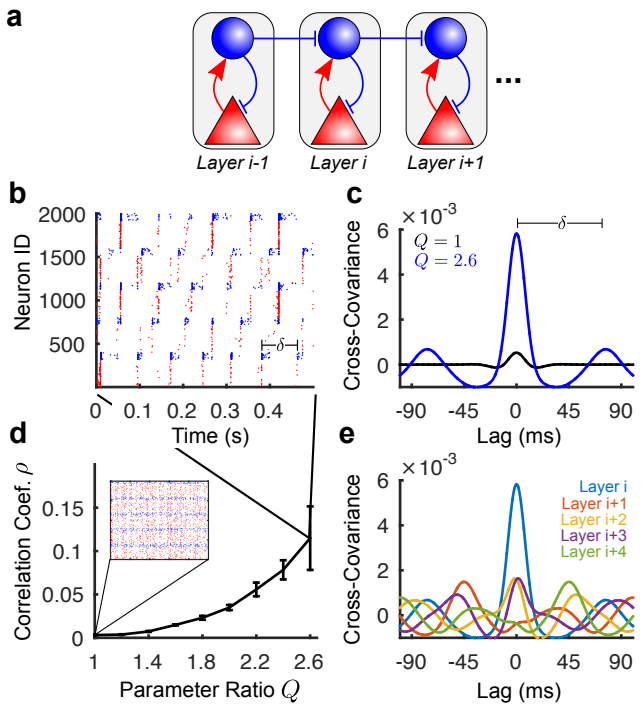

Figure 4.2: **Disinhibitory feedforward network. (a)** Schematic of dFFN architecture, in which feedforward activity propagates by the cascading of several disinhibitory structural motifs. For clarity, not all connections are shown in the schematic; importantly, excitatory to excitatory connections are randomly connected (see text). **(b)** Example raster plot of a network with 5 layers showing forward propagation ($Q = 2.6$). The inhibitory neurons again display forward propagating activity. Note that the final group connects back to the first (circular arrangement) and hence the activity propagates indefinitely. **(c)** Cross-covariance functions for the cases of $Q = 1$ and $Q = 2.6$ (see Figure 4.1). **(d)** The average Pearson correlation coefficient within layers as a function of Q. The larger the feedforward ratio, Q, the greater the correlation of firing within layers. Error bars are standard deviation. **(e)** Average cross-covariance function between different layers (see also Fig. 4.1). Note that due concurrently propagating multiple cascades, the crosscovaraince between different groups display multiple peaks (see text).

present here a second network architecture which may be implemented in the cortex, which takes inspiration from the specificity of inhibitory neurons' connectivity patterns [Harris and Shepherd, 2015; Roux and Buzsaki, 2015]. In disinhibition motifs, certain subtypes of interneurons inhibit other interneurons which normally suppress connected excitatory neurons. Such a disinhibition cascade can thus lead to an increase in firing rates in excitatory neurons which are normally suppressed. Inspired by this, we constructed a network model in which disinhibitory motifs enable the feedforward propagation of spiking activity. We hereafter call this architecture a disinhibitory feedforward network (dFFN). A diagram of the wiring scheme for a dFFN is shown in Figure 4.2a.

The functionality of this circuit can be explained schematically as follows: (i) each layer comprises a functional group of excitatory and inhibitory neurons more strongly connected

to each other than to the rest of the network; (ii) inhibitory neurons in one layer target preferentially the inhibitory neurons in the subsequent layer (disinhibition); (iii) thus, when the activity in the preceding layer increases, the inhibitory neurons' activity in the next layer will decrease. This in turn allows the excitatory neurons in the next layer to increase their firing; (iv) after a short delay, the inhibitory neurons increase in activity again, as they receive input from the excitatory neurons in their layer, which have elevated activity as a result of their disinhibition. This eventually reduces the total firing back to baseline in the group – however, not without the activity moving to the next group via disinhibition again; (v) in turn, every upstream layer is activated and the information propagates.

We implemented dFFN networks with varying number of layers confirming that the above circuitry results in a directed information propagation (Figure 4.2). Once again we observe feedforward signal propagation for both the excitatory and inhibitory neurons (Figure 4.2B). As before, to avoid boundary effects, we used a circular network layout in which the last group connects back to the first and hence the activity keeps propagating.

In dFFN networks, excitatory-to-excitatory connections remain again uniform. Every layer in this network is composed of groups of exctiatory and inhibitory units as shown in Figure 4.2a. Similar to above, we can control the imposed feedforward level with the ratio parameters $W_{IE} = w^i_{IE}/w^{not[i]}_{IE}$, $W_{EI} = w^i_{EI}/w^{not[i]}_{EI}$, and $W_{II} = w^{i+1}_{II}/w^{not[i+1]}_{II}$ or the analogous ratios of connections probabilities R_{IE}, R_{EI}, R_{II}. Again, for simplicity we set all six parameters equal to Q and vary them concurrently.

The displayed network in Figure 4.2b corresponds to a 5-layer network with $Q = 2.6$. When comparing the average cross-covariance function of such a network with uniformly connected networks $Q = 1$, we can again see an increased synchrony and a propagation of activity resulting from the dFFN architecture (Figure 4.2c). This is further demonstrated by the increasing Pearson correlation coefficient of neurons within layers as a function of Q (Figure 4.2d). Finally, we plot again the average cross-covariance between different layers in Figure 4.2e. Although the same behavior is qualitatively seen of the signal propagating along the cascaded layers of the dFFN topology, this appears to be far less pronounced than in the ccFFN architecture. The reason for this effect can be explained by inspecting the raster plot further (Figure 4.2b). Note that indeed multiple signals appear to be propagating simultaneously: in the ordered rastergram, the lowest group can fire again, even though the cascade that emanated from it previously has not reached the top group yet. For instance, the firing of a new cascade at layer i and a previous cascade at layer $i + 3$ can temporally overlap and this co-alignment results in the multiple peaks in the cross-covariance. Hence, the network propagates multiple signals concurrently, or, stated differently, the network's architecture is able to process multiple signals simultaneously effecting the cross-covariance.

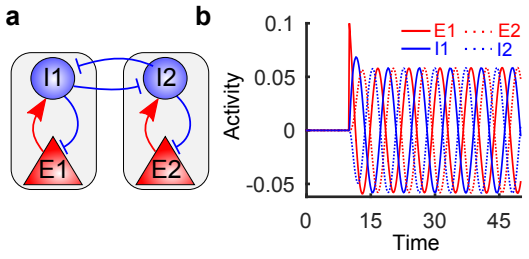

Figure 4.3: **Inhibitory connection profiles drive dynamics. (a)** Network architecture of stylized linear model. See text for dynamical equations. **(b)** Example simulation for $\beta = 0.5$ and a pulse of activity given to E_1 at $t = 10$ causing a permanent oscillation in all nodes of the network.

4.4.3 Feedforward activity in linear rate models

To gain further intuition about the behavior observed in the above circuits, it is insightful to consider stylized linear rate model with 2 layers implementing such mechanisms (Figure 4.3a). Let us denote the group of excitatory (inhibitory) neurons in layer 1 by E_1 (I_1), and the corresponding firing rates by r_{E_1} (r_{I_1}). Note that without loss of generality we may assume that the firing rates are measured relative to some baseline activity, and can thus be positive or negative. We can then write down the following coupled rate equations for layer 1 (equivalent equations for layer 2 not shown):

$$\frac{dr_{E_1}}{dt} = -\beta_{E_1} r_{E_1} - r_{I_1}, \tag{4.4}$$

$$\frac{dr_{I_1}}{dt} = -\beta_{I_1} r_{I_1} + r_{E_1} - r_{I_2}, \tag{4.5}$$

where β denotes the effective self-coupling of the individual groups, which is a compound of a constant decay (leak) term and the coupling of the neurons within the same group. For simplicity we assume here that there is no self-excitation term for the excitatory neurons, and hence the network has an excess of inhibition (not balanced). The magnitude of the parameter β will control how much dissipation there is for each neuron. If we set $\beta = 0.5$, which is the minimal requirement for a stable system, the dominant modes are associated with purely imaginary eigenvalues and correspond to (phase-shifted) wave-functions. If the network is now excited by an instantaneous input, such as an input to E_1 at $t = 10$ as shown in Figure 4.3b, indeed a propagation of activity can occur along the network as outlined above. Note that the general features of this finding are not an artifact of the chosen parameters or model size, but can be generalized to arbitrarily sized networks. The forward propagating activity is indeed a consequence of the described architecture: the particular inhibitory connection profiles drive the dynamics.

4.5 Discussion

In summary, we have demonstrated via LIF network simulations that one can construct networks in which feedforward activity is propagated even if connections between excitatory units are kept completely random. This is achieved by endowing inhibitory units with an active role in the propagation mechanism. Influenced by recent results in connectomics we proposed two different circuit layouts, in which both inhibitory and excitatory neurons display coherent feedforward dynamics. We believe that the recently discovered diversity of interneurons calls for a reassessment of the role of inhibitory neurons in neuronal circuits, beyond merely balancing the network, and would encourage scholars to assign them more active functional roles. Here we have demonstrated how such a functional role could be shaped in feedforward networks, but possibilities for such roles could go far beyond this.

Chapter 5

Sleep restriction in adolescent mice results in long term connectivity changes

In the overall theme of the thesis, my contribution in this chapter was to develop techniques to compare mesoscale connectomes from two groups of animals. Thus, this chapter considers brain networks at a coarser level than previous chapters and is concerned with developing tools for allowing identification of differences between brain networks. In this work, we demonstrate how experimental variabilities can have a profound effect on the data. Once, to the best of our knowledge, all variabilities were removed, we identify that indeed heterogeneous differences do exist between the two mice populations we considered. The actual comparison was done between control and sleep deprived animals, though the tools that I developed can be used on any behavioral or environmental manipulation.

5.1 Abstract

Neurodevelopment is a highly intricate process and there is increasing interest in understanding how sleep affects normal development. In particular, cortical circuits mature in stages, from early rapid synaptogenesis and synaptic pruning to late synaptic refinement, resulting in the adult anatomical connection matrix. Because the mature matrix is largely fixed, genetic or environmental factors that interfere with its development can have irreversible effects. Sleep disruption is rarely considered among those factors, and previous studies have mainly targeted very young animals and focused on the acute effects of sleep

deprivation on neuronal morphology and cortical plasticity. Adolescence is a time of intense brain remodeling and a sensitive period for the pathophysiology of mental disorders. Yet whether chronic sleep restriction (CSR) during adolescence has long-term effects on brain connectivity remains unclear. To start addressing this question we used viral-mediated axonal labeling, serial two-photon tomography of the entire brain, and the Allen Mouse Common Coordinate Framework. We focused on a high-order area with diffuse projections, the mouse secondary motor area (MOs), and for each of its targets we calculated the projection density, a combined measure of passing fibers and axonal terminals. We found no homogeneous difference in MOs projection density values between mice subjected to 5 days of CSR during early adolescence (P25-P30, $\geq 50\%$ decrease in daily sleep, n=14) and siblings that slept undisturbed (n=14). Machine learning algorithms, however, classified animals at significantly above chance levels, indicating that differences between the two groups exist, but are subtle and heterogeneous. Thus, sleep disruption in early adolescence may affect adult brain connectivity. However, because our method relies on a global measure of projection density and was not previously used to measure connectivity changes due to behavioral manipulations, definitive conclusions on the long-term structural effects of early CSR require additional experiments.

5.2 Introduction

From early development to the end of adolescence cortical circuits mature in stages, from early massive synaptogenesis and synaptic pruning, which result in large changes in the absolute number of synapses, to late synaptic refinement, when the initially homogeneous connectivity is reorganized without major changes in synaptic density [Innocenti and Price, 2005; Sanes and Yamagata, 2009; Tau and Peterson, 2010; Uddin et al., 2010]. The end result of these processes is the adult anatomical connection matrix. Because this matrix is largely fixed, genetic or environmental factors that interfere with its establishment during development can have irreversible effects. Sleep disruption is rarely considered among these factors, perhaps because one can always sleep longer or deeper at a later time. Thus, few studies have tested the hypothesis that sleep disruption during development may impair the maturation and maintenance of brain circuits [Roffwarg et al., 1966]. For example, early experiments used drugs to disturb neonatal sleep, and found long-term neurochemical and behavioral effects, for instance on anxious behavior (reviewed in [Frank, 2011]). However, these changes were likely caused not only by sleep loss, but also by other effects of the drugs used to enforce wake, many of which affect monoaminergic transmission [Frank, 2011]. More recent experiments in kittens combined monocular deprivation with one week

of rapid eye movement (REM) sleep deprivation before the end of the critical period and found a decrease in the size of neurons in the lateral geniculate nucleus of the thalamus [Shaffery et al., 1998], and similar results were obtained after total sleep deprivation [Pompeiano et al., 1995]. Chronic REM sleep deprivation alone also leads to the persistence of an immature form of synaptic potentiation in primary visual cortex, suggesting that sleep loss slows down the maturation of cortical circuits [Roffwarg et al., 1966; Shaffery et al., 2002, 2012]. Other studies in kittens found that a few hours of total sleep deprivation can immediately impair ocular dominance plasticity when sleep is prevented at the height of the critical period [Frank, 2011; Frank et al., 2001], and acute sleep deprivation in adolescent mice impairs the growth and maintenance of a subset of cortical spines formed after learning [Yang et al., 2014]. Most of these experiments focused on pre-adolescent animals, and morphological and electrophysiological effects were assessed immediately or soon after the end of sleep deprivation. Thus, whether sleep loss during development leaves permanent structural changes in the adult brain was unknown, and even less known were the consequences of sustained sleep disruption during the sensitive period of adolescence [Paus et al., 2008].

Here we tested in mice whether the occurrence of chronic sleep restriction (CSR) during early adolescence has long-term effects on the adult anatomical connection matrix. In rodents, adolescence can be broadly defined as the period from weaning at postnatal day 21 (P21) to sexual maturity (∼P50-P60) [Spear, 2000]. Massive synaptogenesis and synaptic pruning occur mainly during the second postnatal week [Aghajanian and Bloom, 1967; Ashby and Isaac, 2011; De Felipe et al., 1997; Koester and O'Leary, 1992; Maravall et al., 2004; Micheva and Beaulieu, 1996; Romand et al., 2011]. Synaptic refinement follows in the third and forth postnatal week, when the initially homogeneous connectivity is reorganized without major changes in synaptic density, and the functional optimization of cortical circuits continues throughout adolescence [Cancedda et al., 2004; Ko et al., 2013; Seelke et al., 2012; Zhang et al., 2002]. Thus, during early adolescence (∼P21-P34) the anatomical connection matrix is still being refined. During the same time electroencephalographic (EEG) patterns across the sleep/wake cycle are similar to those seen in adults, and so are total daily sleep amounts [Frank and Heller, 1997; Gramsbergen, 1976; Nelson et al., 2013].

Herein we show that, compared to the control group, there was no absolute affect that can be observed due to CSR. In others words there was no overall single difference throughout the whole brain between the two groups. Using a novel machine learning algorithm, however, we were able to classify the animals with significant accuracy indicating that the developed method can capture the fine effects of CSR on long-term structural connectivity. Future experiments and analysis will allow the teasing out of these differences further.

5.3 Materials and Methods

5.3.1 Animals

Five litters of C57BL/6J mice of the same age (n=32) were used in one single experiment that included 5 days of chronic sleep restriction (or sleep *ad libitum*) between postnatal day 25 and 30 (P25-P30), surgery for cortical injection of viral tracer at P44-P47, and sacrifice for brain collection at P65-P68 (see Figure 5.2A). All animal procedures followed the National Institutes of Health Guide for the Care and Use of Laboratory Animals and facilities were reviewed and approved by the IACUC of the University of Wisconsin-Madison, and were inspected and accredited by AAALAC.

5.3.2 Experimental procedure

At P21 mice were weaned, weighed, and housed in groups (4 per cage) in environmentally controlled conditions (12h:12h light-dark cycle; lights on at 8:00 am, room temperature 23 ± 1 degrees Celsius). At P24 body weight was re-checked and two groups of 16 animals, weight-balanced and sex-balanced, were created from the total pool of 32 mice. Each group was moved into a large cage (60 cm x 60 cm x 40 cm) where mice were free to interact. Food and water were provided *ab libitum* and replaced daily at 8 am. At P25 the control group was left undisturbed and video-monitored for 5 days, while the second group was subjected to 5 days of chronic sleep restriction (CSR) starting at 8 am. At that time adolescent mice show EEG patterns across the sleep/wake cycle similar to those of adult mice, with low voltage fast activity during wake and REM sleep and large slow waves during NREM sleep [Frank and Heller, 1997; Gramsbergen, 1976]. Total daily sleep amounts in young adolescent mice are also at adult levels [Nelson et al., 2013]. On the other hand, REM sleep in mice continues to decline during early adolescence, and sleep deprivation is followed by an increase in sleep duration but not in sleep intensity, suggesting that the mechanisms of homeostatic sleep regulation are not fully mature [Nelson et al., 2013].

CSR was enforced using multiple strategies to disrupt sleep, as described previously [de Vivo et al., 2016], with the only difference that caffeinated water was not used in the current study. Stimuli were selected to be as much as possible ecologically relevant, and included continuous exposure to novel objects, changes of cage and bedding, free access to multiple running wheels during the day, as well as mild forced locomotion on a slowly rotating platform during some parts of the night. The platform was located above a tray filled with 2-3 cm of water, and the rotation speed was low enough that mice could easily avoid falling into the water as long as they moved continuously. Video cameras and/or direct visual observation were used to monitor the mice at all times. A previous CSR study that

lasted 4 days (P25-P29) and used mice implanted with EEG electrodes found that total sleep time throughout the experiment was decreased by $\sim 60\%$ [de Vivo et al., 2016]. After CSR (or sleep *ad libitum*) mice were returned in their home cages (4 per cage), and continued to have access to novel objects (new sets of objects every morning) and running wheels until the end of the experiment at P65-68. All mice gained weight between P21 (weaning day) and P30 (end of CSR), but controls did so more than CSR mice (C: $+44.2 \pm 12.2\%$; CSR: $+20.7 \pm 9.3\%$, t-test, $p < 0.001$), and in each group males grew more than females (C/F $+34.8 \pm 12.7\%$, C/M $+51.1 \pm 12.2\%$, t-test, $p = 0.007$; CSR/F $+15.8 \pm 9.8\%$, CSR/M $+27.9 \pm 8.3\%$, t-test, $p = 0.008$).

5.3.3 Stereotaxic injection of AAV for anterograde axonal tracing

Surgery occurred over the course of 4 days (8 mice/day) between P44 and P47. Anterograde axonal tracing from secondary motor cortex (MOs) was performed by injecting AAV1.hSyn.eGFP.WPRE.bGH (1.79×10^{13} GC/mL) at two different depths using iontophoresis, which allows for small, focal injections [Harris et al., 2012]. The day before surgery, glass capillary tubing was heat-pulled to create pipet tips that were then cut and verified under a microscope to obtain tip widths of 10-30 μm. Just prior to surgery, these pipets were filled with virus using capillary action to prevent formation of bubbles. Mice were anesthetized under 2% isoflurane and maintained at 1-2% isoflurane for the duration of surgery. Using sterile technique, mice were fitted into a stereotaxic frame and an incision was made to expose the skull. The skull was cleaned with saline and hydrogen peroxide, and a small burr hole was made in the skull using a dental drill. Any exposed brain was kept moist by saline at all times. The filled pipet was then prepared by lowering a silver wire into the pipet until it contacted virus. An electrical lead was attached to the silver wire, and an electrical ground was connected to a metal clip placed on the skin near the skull. The pipet was then lowered to the surface of the brain 1.7 mm anterior and 0.75 mm lateral (right) from bregma. From the cortical surface, the pipet tip was lowered through the dura and into the brain 0.4 mm. A pause of 2 minutes was given to allow for a weak seal to form between the brain and the glass of the pipet. Current was then delivered through the pipet tip at 3 μA, alternating 7 seconds on and 7 seconds off, and repeating for 5 minutes to inject viral particles. The pipet tip was then lowered another 0.4 mm (to a total depth of 0.8 mm from the surface of the cortex) and the 5 minute current delivery was repeated. After the current was stopped, the pipet tip was kept in place 5 additional minutes to allow for any pressure to dissipate before removal of the tip. After removal of the tip, the incision was sealed using Vetbond, antibiotic gel was applied to the surgical site, and mice were removed from isoflurane. Mice were monitored daily for 7 days following surgery to ensure normal

recovery. Two control mice experienced health issues in this period and were sacrificed.

5.3.4 Perfusion

Three weeks after surgery (P65-P68), mice (16 CSR and 14 controls) were deeply anaesthetized with isoflurane (3% volume) and perfused transcardially with a flush (∼30 sec) of saline followed by 4% paraformaldehyde (PFA) in phosphate buffer (PB). Mice were then decapitated, and heads were kept in 4% PFA until shipping to the Allen Institute for serial two-photon tomography.

5.3.5 Serial two-photon (STP) tomography

Briefly, carefully dissected brains were prepared for STP tomography, which integrates optical imaging and vibratome sectioning, by first rinsing with PBS before embedding in an agarose block as previously described in detail [Oh et al., 2014]. Images were acquired on TissueCyte 1000 2-photon systems (TissueVision, Cambridge, MA) coupled with Mai Tai HP DeepSee lasers (Spectra Physics, Santa Clara, CA) using 925 nm wavelength light through a Zeiss 20x water immersion objective (NA = 1.0). One optical plane was imaged 75 m below the cutting surface. After an entire section was imaged at an XY resolution of ∼0.35 μm/pixel, a 100μm section was cut by the vibratome and then the specimen was returned to the objective for imaging of the next plane. Images from 140 coronal sections were collected to cover the full mouse brain. Data from one CSR mouse could not be used due to a problem in image alignment. Another CSR mouse was excluded due to problems during imaging, and thus the final analysis included 14 controls and 14 CSR mice.

5.3.6 Image data processing

The Allen Institute informatics data pipeline (IDP) managed processing and organization of the images and quantified data for analyses. Algorithms developed for the Allen Mouse Connectivity Atlas for signal detection and image registration were used on this dataset [Oh et al., 2014]. Detailed descriptions of the neuroinformatics developed for segmentation and registration for this Atlas were published recently [Kuan et al., 2015]. Briefly, the signal detection algorithm was applied to each image to segment positive fluorescent signals from background. Steps include low pass filtering to remove noise, followed by adaptive edge/line detection and classification, then integration of the detected results and rejection of artifacts or outliers. For registration, as STP tomography results in inherently aligned section images, we can simply stack the section images together to form a coherent reconstructed 3D volume. Each image stack is first registered to an intermediate “template” brain, cre-

ated by iteratively averaging across ∼1700 brains from the Allen Mouse Connectivity Atlas. Registration to this template occurs in two broad steps: global alignment followed by local alignment effected through a coarse-to-fine deformation registration. The final step is then to align to the 3D Allen Mouse Common Coordinate Framework model.

5.3.7 Projection density estimation

After segmentation and registration, signal was quantified for each voxel (10 μm × 10 μm × 10 μm) in the reference space and for each structure in the ontology by combining voxels from the same structure in the 3D reference model.

5.3.8 Thresholding

The regions that had mean projection fractions less than 0.1% were removed from the comparison analysis because the signal was too weak to be reliable across mice. For the ipsilateral side of the injection, 56 regions were removed (237 remained) and 92 regions (201 remained) for the contralateral side.

5.3.9 Injection volume normalization

Due to the experimental difficulties in controlling for the injection volume in every animal, we sought to account for the differences by normalizing by a factor proportional to the injection volume. Figure 5.1A illustrates the need for the normalization since there is a correlation between injection volume and the total summed projection fraction. Although such a dependence was expected, it is vital to account for as the larger the injection volume, the more neurons that become infected, resulting in more axonal signals and hence inaccurately larger projection fractions values. We observed that a direct division of the projection fractions by injection volumes was not suitable and resulted in a negative correlation (Figure 5.1B). Thus, we proceeded to normalize the data by fitting a power law:

$$\sum PF = A(InjVol)^n + B, \tag{5.1}$$

where $\sum PF$ is the sum of all projection fractions for an animal (after thresholding), $InjVol$ is the injection volume, and (A, B, n) are constants. It can be seen that $B = 0$ as there will be no fluorescence signal in the absence of any injection.

Taking the logarithm:

$$\log \sum PF = n \log(InjVol) + \log(A), \tag{5.2}$$

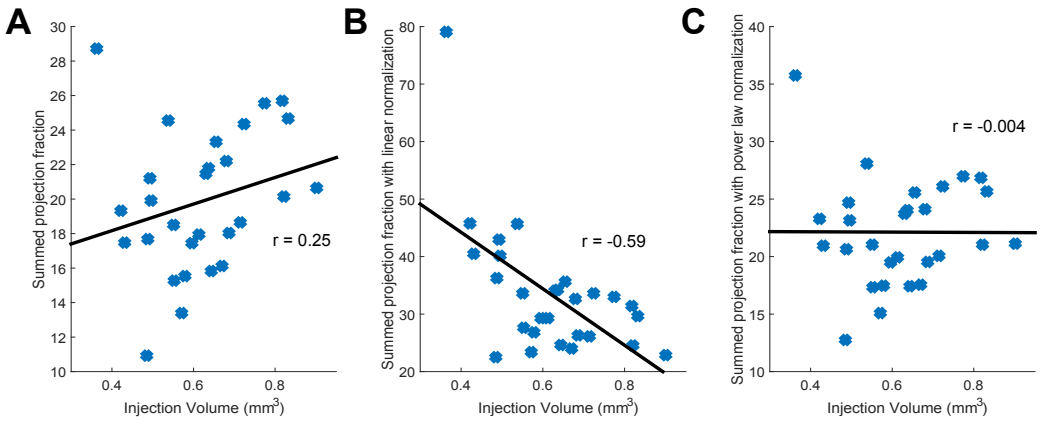

Figure 5.1: Effect of injection volume on total summed projection fraction. **A** Correlation of injection volume with the total summed projection fraction for every animal. Both the control and CSR groups are plotted together. The y-axis is summed projection volume for all regions considered and is thus technically bounded between 0 and 237. **B** Normalizing the total summed projection fraction by simple linear division with the injection volume results in over correction and a negative correlation with injection volume. **C** Normalizing the total summed projection fraction with a power law (see text) eliminates the dependence on injection volume. All values are given for Pearson's rho.

a linear regression fit allowed us to determine that $n = 0.2160$ was optimal and hence all the projection fractions were divided by $(InjVol)^n$ which is suitable as seen in Figure 5.1C. Note that once this normalization is done, however, the projection fraction values are no longer guaranteed to be less than 1.

5.3.10 General Linear Model

A general linear model (GLM) was used to control for differences in the centroids of the injections. This was done on the data after having normalized by the injection volume as described above. For every region, a GLM was fit to see the effect of the group type (CSR or control) on the normalized projection fraction. To allow comparisons between regions, every region was normalized to have unit mean. This was followed by fitting the following GLM:

$$PF_{i,j} = \beta_j + k_{xj}x_i + k_{yj}y_i + k_{zj}z_i + k_{cj}c_i \tag{5.3}$$

Where $PF_{i,j}$ is the normalized projection fraction (by injection volume as described previously) for animal i at region j. x_i is the medial-lateral distance of the injection site from the midline for animal i after adjusting for all animals such that $\bar{x}_i = 0$. Similarly, y_i corresponds to the anterior-posterior distance (from the anterior commissure) and z_i to depth measurements (from the pia). c_i corresponds to the condition of the animal where

control = 1 and CSR = -1. The terms k_{xj}, k_{yj}, k_{zj}, and k_{cj} are factors that capture the influence of their corresponding variable and are determined by maximum likelihood estimation with the MATLAB 2015b Statistics and Machine Learning Toolbox. β_j is the intercept term for the GLM for every region j. It should be noted that if k_{cj} is positive, then that indicates that control animals (c_i=1) will have an increased projection fraction due to their condition. The opposite is true if k_{cj} is negative. If k_{cj}=0 then there is no effect due to the condition. Although our analysis shows that individually k_{cj} values are not significant, we observe that all of the k_{cj} values are positively biased with a positive mean (Figure 5.3C). The mean of all k_{cj} values is called μ_{k_c}. However, the $\mu_{k_c} > 0$ result does not pass a bootstrap significance test as described below.

5.3.11 Bootstrap

To test the confidence of the positive μ_{k_c} result, we performed a bootstrap test on the data. Each group was separately resampled (with replacement) to create a total of 50 new data sets while maintaining the same size for each group. Thus any single resampling case may have some animals selected multiple times and others not selected at all. This allowed for 2500 (50 CSR $\times$ 50 control) comparisons where we determined the fraction of times that μ_{k_c} was negative as the p-value. Note that every comparison involved determining a new GLM for every region. With these comparisons, the best p-value that may be claimed is $1/50^2$ though the attained p-values were significantly larger. For the test where the female mice were investigated separately, μ_{k_c} changed sign ($\mu_{k_c} < 0$) and hence the opposite one-tailed bootstrap was performed (ratio of positive μ_{k_c} to total number of comparisons).

5.3.12 Classification techniques

Machine learning techniques were implemented to classify the animals between CSR and control groups. This was done after fitting a modified GLM as described above that did not include a condition parameter. The dependence of the injection centroid for every animal was then subtracted such that, to the best of our knowledge, only condition was a factor in influencing projection fractions. The algorithms were trained on $n-1$ animals and classification was tested on the excluded animals. This process was iterated and the classification performance was quantified as the ability to classify all animals in this manner. Using a standard logit-boost decision tree to minimize binomial deviance gave the best classification accuracy at 71% (8 errors). This was due to the large overlap between animal projection values, such that no specific regions were adequate to instantly determine if there were differences between the two groups. Hence, to improve the performance, a pre-processing step

as inspired by Role-Base Similarity (RBS) was applied [Beguerisse-Díaz et al., 2014]. Here, correlations between all regions were calculated to generate a 237×237 correlation matrix (ipsilateral hemisphere only considered). The matrix was thresholded at zero to create a positive-only correlation matrix in addition to zeroing the diagonal. Multiscale clustering was achieved using the dynamics-based clustering framework of Markov Stability [Billeh et al., 2014; Delvenne et al., 2010; Schaub et al., 2012]. The clustering algorithm used a Markov process to find clusters at different scales [Schaub et al., 2012]. Each cluster was then merged by summing each region within it. For instance, if a group of regions were grouped into a cluster, then for each animal the projection fractions were summed up to attain merged projection values for all regions in that cluster. At the different scales found by the Markov stability algorithm, a binary decision tree that minimizes the binomial deviance between a root node and targets was again used to classify the compacted data (MATLAB 2015b Statistics and Machine Learning Toolbox). The classification improved to an accuracy of 82% (5 errors). To further verify the significance of the result, 1000 random labels were made and the classification technique was repeated on these iterations. The classification was equal or better than 82% less than 1% of the time ($p < 0.01$).

5.4 Results

5.4.1 Animal handling and data collection

Mice of the same age from five litters were split into two groups (Figure 5.2A). One group (n=14) was subjected to 5 days of CSR in the middle of early adolescence, from P25 to P30, using ecologically relevant stimuli (see Methods), while during the same period control siblings (n=14) were allowed to sleep *ad libitum*. Mice were not equipped with EEG electrodes to avoid potential damage to the cortex. However, based on continuous visual monitoring and a previous study with EEG recordings [de Vivo et al., 2016], we estimate that overall sleep loss was between 50 and 60% (see Methods). Roughly two weeks later (P44-P47) CSR mice and controls were injected with recombinant adeno-associated virus (AAV) expressing enhanced green fluorescent protein (EGFP) in the right secondary motor area (MOs), to map its projections. We focused on MOs because it has diffuse projections [Zingg et al., 2014] and is highly plastic [Cao et al., 2015]. Exactly three weeks after each animal's injection the brains were perfused. Fluorescent signals were then imaged using serial two-photon tomography and informatically reconstructed within the Allen Mouse Common Coordinate Framework, a high-resolution coordinate system that allows the systematic analysis of the entire brain (see Methods, [Oh et al., 2014]).

As expected, robust anterograde tracing from right MOs was observed throughout the

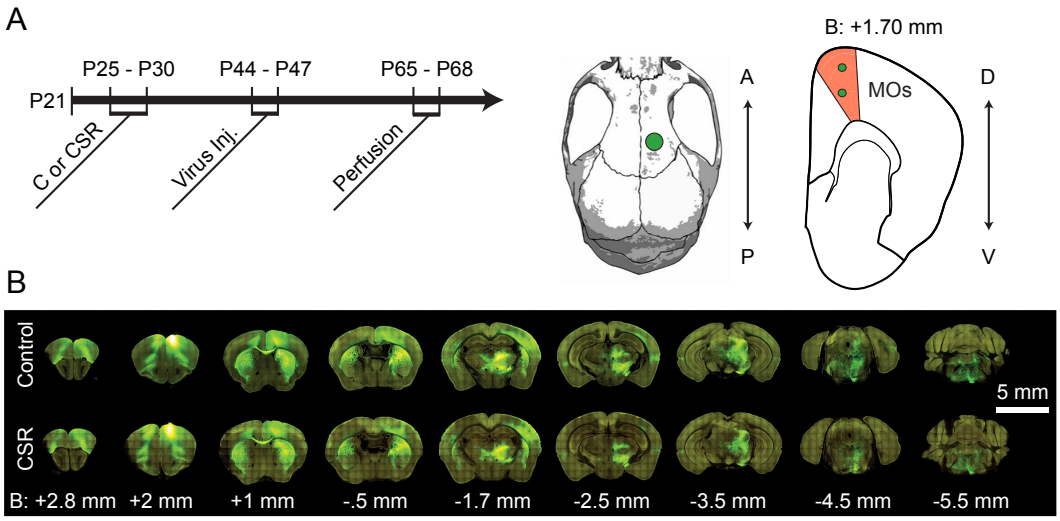

Figure 5.2: Experimental Timeline and MOs Projections. **A** Experimental timeline. Between P25 and P30, mice were allowed to sleep normally (C, n=14) or subjected to chronic sleep restriction (CSR, n=14). All mice were injected with AAV-GFP between P44 and P47, and each mouse was perfused exactly three weeks later. Middle and right panels show the location of the viral injections on the skull and in a coronal brain section. A=anterior, P=posterior, D=dorsal, V=ventral, B=bregma. **B** Example of projections from MOs in two representative mice (C and CSR) three weeks after injection of AAV-GFP. Measurements are given in mm from bregma.

brain (Fig. 5.2B). The overall pattern of projections was consistent across mice and similar to the one previously described for rat supplementary motor cortex, also known as medial agranular cortex [Reep et al., 2008; Stuesse and Newman, 1990]. Projections were always stronger, or only present, on the side of the injection, again consistent with previous studies showing that most MOs projections are ipsilateral [Oh et al., 2014; Reep et al., 2008; Stuesse and Newman, 1990]. Thus, subsequent analyses primarily focused on the right side.

5.4.2 Informatics connectivity analysis

To identify potential differences in connectivity between CSR mice and controls we examined the *projection fraction* (also referred to as projection density) values for each brain structure that receives axonal projections from MOs. The projection fraction is defined as the total number of voxels that fluoresce in a target brain structure divided by the total number of voxels in that structure (see Materials and Methods). Hence, projection fraction is positive-only, has a maximal value of 1, and its use allows to control for differences in volume across anatomically defined regions. Note that the projection fraction includes both passing fibers and axon terminals, because they could not be differentiated informatically from the images. Before direct comparisons between the two groups were made, projection fractions were normalized to control for small differences in injection volume across mice (see Materials

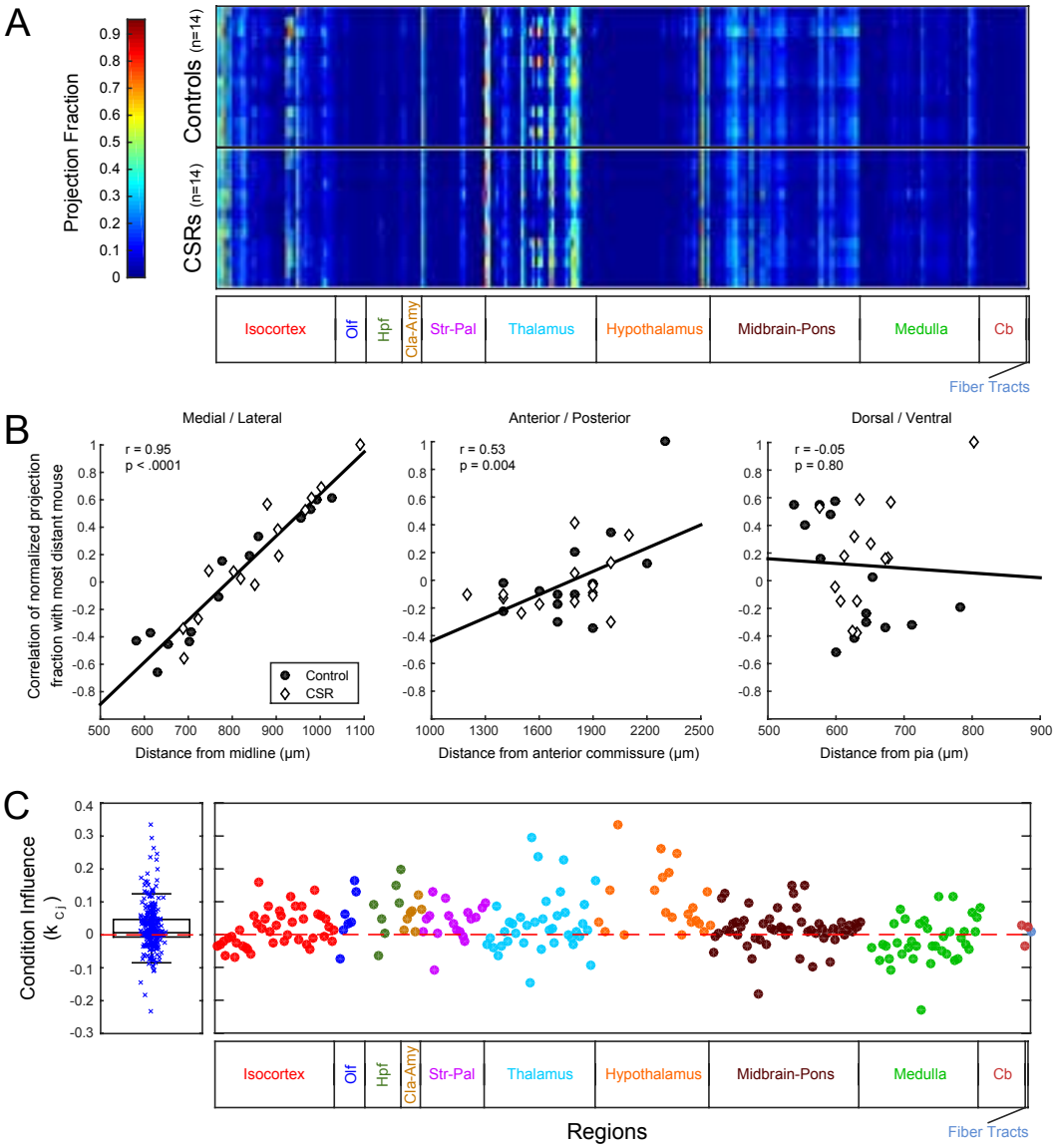

Figure 5.3: Projection Fractions in Controls and CSR Mice. **A** A plot of projection fractions (normalized by injection volume) across all mice (y axis) and all regions (x axis). **B** Correlations of projection fraction patterns in mice with injection site. The mouse with the most extreme distance in injection location is taken as the basis for correlation (and is by definition equal to one). All regions were first normalized to have unit means to account for differences between injections. Without this normalization, the mean pairwise correlation of all animals is 0.903 ± 0.07. Distance in microns is measured to the center of the injection site from the midline ($826 \pm 140 \mu m$, left panel), the midline merging of the anterior commissure ($1771 \pm 262 \mu m$, middle panel) or the pial surface ($636 \pm 60 \mu m$, right panel). Values are given for Pearson's rho and reported as mean $\pm$ standard deviation. **C** Left panel: box plot of the condition influence, k_{cj}, of the GLM for all regions. A positive k_{cj} indicates that the control group has a higher connectivity than the CSR group (see text for details). Right panel: an expanded version that is aligned and color-coded to match the subdivision in major anatomical regions. Regions that did not pass threshold have no corresponding k_{cj} point plotted.

and Methods, also Figure 5.1). Moreover, regions with very weak signal were removed (projection fractions $< 0.1\%$; see Materials and Methods). For the ipsilateral hemisphere, this thresholding resulted in dropping 56 regions out of a total of 293, leaving 237 regions for analysis. To ensure that signal within the weakest of the 237 regions was not simply due to false signal detection, detected signal overlays were compared to raw fluorescent images in two mice. Manual inspection in these mice confirmed that there were very few false positives and negatives, meaning that weak detected signal corresponded closely to real fluorescence. Figure 5.3A visualizes the normalized and thresholded data for the two groups; every row corresponds to a different mouse injection into MOs and every column is the projection fraction from MOs to that brain structure.

A challenge in the outlined experiments is the difficulty in precisely replicating the injection site position. To determine the effect of such variations, a plot of the mouse projection correlations relative to the most laterally injected mouse is shown in Figure 5.3B. Note the most laterally injected mouse has a perfect correlation of 1 as it is compared with itself. As can be seen there is a strong dependence relative to the injected medial-lateral position (r = 0.954). Similar patterns could be observed for the other dimensions (Figure 5.3B). We emphasize that in determining the correlation for Figure 5.3B, we normalize by the mean projection fraction of every region (to have unit mean) to account for differences between injections which is why we observe negative correlation values. Without this normalization and using solely the heat map in Figure 5.3A gives a high mean pairwise correlation between all injections (0.903 ± 0.07; mean $\pm$ standard deviation).

The effect also showed site specificity where, for instance, the anterior-posterior axis had a strong relationship in the isocortex (r = 0.602) and a significant influence by the depth axis was seen in the olfactory areas (r = 0.340). To account for this experimental variance, a general linear model (GLM) was fit to the normalized unit-mean projection fraction for every region (see Materials and Methods for more details). In the GLM, the effect of the condition (control or CSR) on a specific region is captured by a parameter k_{cj} where by construction, if k_{cj} is positive then the control group has a higher projection fraction to region j relative to the CSR group and vice versa. We found that $k_{cj}$was not significant on a region level, though the distribution of k_{cj} values did appear more positively biased with a positive mean (Figure 5.3C, left panel). However, running a bootstrap to test this effect yielded insignificant results ($\mu_{k_c} = 0.026$, $p = 0.252$, see Materials and Methods). Observing the distribution of k_{cj} across the different macro-regions (Figure 5.3C, right panel) indicates that certain brain regions may be more affected than others by sleep restriction. Once again, however, none of the divisions considered passed the bootstrap significance test. Performing the same analysis on the contralateral side yielded similar results.

To determine if there were sex-specific differences, we performed the same analysis on the male and female animals separately. This was possible as we had a similar number of males and females (14 males, 8 controls; 15 females, 6 controls). Our results show that the sleep deprivation paradigm did not influence the males' ($\mu_{k_c} = 0.013$, $p = 0.491$) nor the females' ($\mu_{k_c} = -0.034$, $p = 0.258$) mesoscale connectivity. Due to the small number of animals however, we cannot rule out that subtle effects do exist that we are unable to detect.

5.4.3 Animal classification

We also investigated if we could use the MOs normalized projection fractions (adjusted for injection positions, see Materials and Methods) to classify animals using machine-learning techniques. As is standard practice with very small data sets, we trained classification algorithms on all but one animal, used the algorithm to predict the group of the excluded mouse, and repeated the procedure for all mice. We then evaluated the performance of our algorithm by its ability to classify all 28 mice. The best performance we could attain on the normalized data was 71% (8 errors, see Materials and Methods). To see if we could improve classification accuracy, we applied a pre-processing step as inspired by a newly developed graph-theory technique termed Role-Base Similarity [Beguerisse-Díaz et al., 2014]. Briefly, we found the positive correlations between all regions to create a positive-only correlation matrix that was then clustered at different levels of granularity using a Markov stability algorithm [Billeh et al., 2014; Delvenne et al., 2010; Schaub et al., 2012]. By considering different levels of granularity and using classification tree algorithms on the compacted data, we reached a classification accuracy of 82% (5 errors, see Materials and Methods). These results again illustrate the difficulties of finding specific or even diffuse differences across the brain that are strong enough to differentiate the two groups. Since the classification problem is binomial in nature, the classification accuracy corresponds to a p-value of 0.0005 (determined from a binomial cumulative distribution function with 5 errors, 28 attempts, at a probability of 0.5). This indicates that although CSR does not affect the brain in a single homogenous direction, it does have an intricate heterogeneous effect that can be captured by machine learning classification. This was further verified by 1000 random labeling iterations and repeating the same classification (see Materials and Methods). The results indicate that to achieve 5 errors or better was again statistically significant ($p < 0.01$). Overall, the variability observed in the decision trees from dropping animals did not show a clear hypothesis for *post hoc* testing. Nonetheless, we conclude that long-term changes in brain connectivity at the mesoscopic level do occur and further investigations are required to fully uncover the differences.

5.5 Discussion

To our knowledge, this is the first study that tested whether there are structural changes in the adult mammalian brain after sleep was restricted during early adolescence. Brains were collected soon after mice reached adulthood, but younger mice were not tested. Thus, there may have been acute effects of CSR that we missed. Our goal, however, was to search for late, possibly irreversible effects of sleep loss on the adult connectivity.

Sleep in our mice was reduced by 50-60% for 5 days. In humans, this decrease would translate in sleeping ∼4.5 hours of sleep per day, assuming that teenagers (14-17 years) sleep the recommended 8-10 hours per day [Hirshkowitz et al., 2015]. Moreover, 5 days represent 12.5% of the adolescent period in mice (P21- P60), which would translate to 183 days in humans if we assume human adolescence lasts 4 years, although onset and offset vary widely depending on the parameter that is being measured [Roenneberg et al., 2004]. In short, the sleep loss that we enforced in mice was long and severe, and may rarely occur in humans. On the other end, epidemiological studies consistently find that adolescents build a chronic sleep debt during the school days, which they are assumed to compensate during the weekends by sleeping 1-2 hours longer [Leger et al., 2012; Roenneberg et al., 2007; Wolfson and Carskadon, 1998]. Whether this milder but repeated pattern of sleep restriction impairs the maturation of brain circuits is unknown. The inter-individual variability of the structural effects of chronic sleep loss in adolescents is also unknown, but adults vary in their susceptibility to the cognitive impairment caused by sleep deprivation [Rupp et al., 2012; Van Dongen et al., 2004], and differences in the microstructure of white and grey matter can predict inter-individual differences in the resistance to sleep loss [Bernardi et al., 2016; Cui et al., 2015; Rocklage et al., 2009].

We investigated the effect of gender in our study because sex differences in sleep exist in both humans and rodents [Mong et al., 2011]: relative to males, adult female C57BL/6J mice (the strain used in the current study) are awake ∼1.5h more per day, recover relatively more sleep after acute sleep deprivation [Paul et al., 2006], and respond to restraint stress with a smaller rebound in REM sleep [Paul et al., 2009]. Moreover, CSR mice were kept awake using mild forced locomotion, exposure to novel objects and social enrichment. None of these methods is routinely used in chronic variable stress paradigms. Yet sleep is tightly homeostatically regulated and sleep pressure becomes irresistible even after just a few hours of extended wake [Borbely et al., 2016]. Thus, extending wake beyond its physiological duration is inherently stressful, and chronic sleep loss in male adult rats leads to increased levels of catecholamines and, to a lesser extent, ACTH and glucocorticoids [Rechtschaffen and Bergmann, 2002]. The behavioral effects of stress are sexually dimorphic. For instance,

C57BL/6J mice that were kept awake by gentle handling for 3 hours daily from P5 to P42 show changes in sociability and repetitive behavior (but not in anxiety measures) when tested as adults, and these effects differ between males and females [Sare et al., 2016]. The structural effects of chronic stress are also sexually dimorphic in rodent prefrontal cortex and hippocampus, with loss of dendritic spines only seen in males but not in females [Leuner and Shors, 2013], although the underlying mechanisms are unclear and may include different sensitivity to glucocorticoids [Gillies and McArthur, 2010; Leuner and Shors, 2013], and/or differences in the response to other hormones and neurotransmitters involved in stress and arousal, including glutamate or noradrenaline [Valentino et al., 2012]. Since our two groups of mice included a similar number of males and females (14 males, 8 controls; 15 females, 6 controls) we specifically tested for any sex-related difference in our findings, but could not find any. However, we cannot rule out that the number of animals may have been too small to detect subtle differences.

Loss of sleep is associated with cellular stress, impaired protein synthesis, and increased energy demand [Borbely et al., 2016; de Vivo et al., 2016; Mackiewicz et al., 2008; Vecsey et al., 2012]. Furthermore, growth and maintenance of neural circuits is energetically expensive and requires continuous protein synthesis [Kleim et al., 2003; Li et al., 2004]. Of note, a recent study subjected flies to total sleep deprivation for 36 hours starting soon after eclosion and tested them as "young adults" (5-day old) [Kayser et al., 2014]. In these male flies courtship behavior was impaired, and the volume of one specific olfactory glomerulus was reduced [Kayser et al., 2014]. Intriguingly, this glomerulus was the one showing the largest growth after eclosion, suggesting that the most rapidly maturing brain regions are uniquely sensitive to sleep deprivation [Kayser et al., 2014]. It is possible, therefore, that had we tested younger mice, we would have found stronger and more prominent permanent structural effects caused by early chronic sleep loss. Moreover, another limitation of this study is that is relies on sterotaxically guided injections of the tracer to track the projections of a specific area. Not surprisingly, and as shown in this study, the specific pattern of these projections is exquisitely sensitive to the exact location of the injection site, resulting in significant inter-animal variability. Moreover, the projection fraction metric used combines both fibers of passage and axonal terminals, and thus specific effects on the terminals may have been missed and may be better assessed using array tomography combined with excitatory and inhibitory pre- and post-synaptic markers [Wang et al., 2014]. Finally, we targeted a high order area that is presumably still undergoing synaptic refinement during early adolescence, but only a systematic analysis of many brain regions can assess the full extent of the effects of chronic sleep loss.

Chapter 6

Conclusions

The field of neuroscience is increasing its multidisciplinary collaborations every day. Experiments and theories complement each other and will rely on one another as we uncover the mysteries of the brain. A common question posed is 'how does the brain work?' Unfortunately, for now, the question itself is ambiguous and too generic. There are numerous divisions to this question that are more concrete yet still elude scientists. For instance, researchers are trying to understand how memories are formed and retrieved, why human brains can learn complex languages, what are emotions, how the brain gives rise to consciousness, why do we sleep, to name but a few. The implications from finding these answers will be vast, from treating neurological diseases to applications only believed possible in science fiction.

The work in this thesis shines a small light into the complexity of the brain from a network perspective. With the 86 billion neurons in the human brain [Azevedo et al., 2009], each having incredible properties and diversity, the need for interdisciplinary collaboration is quickly evident. The chapters presented also show how brain networks can be studied at multiple levels of abstraction. Chapter 2 showed how the spiking patterns of neuron pairs can be associated to determine a functional connectivity network which can then be clustered to find communities of neurons with similar firing patterns. The methodology introduced can find groups of neurons at multiple scales to identify hierarchical partitions, in addition to potentially identifying embedded feedforward firing patterns, as well as the option of treating inhibitory neurons differently since they have opposite post-synaptic effects to allow for the identification of alternative partitions that are missed when the neuronal subtypes are not considered.

Networks at the microscopic scale where all the connections are known were studied in Chapter 3. Multiple clustered topologies were considered and shown to have almost identical spatiotemporal activity patterns. This illustrated that an extensive neuro-architectural basis is capable of exhibiting structured spatiotemporal dynamics. The work also demon-

strated that the spectral properties of the synaptic weight matrix can predict this behavior. One interesting topology studied had an inhibitory feedback loop between excitatory and inhibitory neurons showing that inhibitory neurons can play a role in the generation of spatiotemporal activity patterns and are not only responsible for balancing neuronal networks. This observation was extended in Chapter 4, which illustrated how inhibitory neurons can be the primary agent in allowing the propagation of feedforward activity. This was done via two different connectivity architectures and further illustrates the vast potential role of inhibitory neurons and the need to link network topology with observed activity.

Chapter 5 presented the examination of axonal projection data collected by the Allen Institute's Mesoscopic Connectivity Atlas [Oh et al., 2014] in sleep-deprived and control mice. The sleep deprivation was done during adolescence and the analysis showed that there is no unidirectional holistic difference between the groups. Based on the ability of machine learning classifiers to differentiate animals at above chance level, the chapter illustrates that subtle long term effects do arise due to the chronic sleep deprivation in mice. Additional analysis and experiments will be required to tease apart the long-term structural changes.

On a global scale, this thesis contributes to answering the larger-scale question of how network architecture and network computations are connected. As connectomic and neuronal activity data increase, we can start solving some vital problems. This can include why the cerebellum, which has more neurons than cortex [Herculano-Houzel, 2012], does not contribute to human consciousness (reported by patients with damaged cerebellums [Lemon and Edgley, 2010])? In parallel, what are the particular features of cortical architecture that allow for conscious experiences but not found elsewhere such as the enteric nervous systems of the digestive tract that has 500 million neurons [Furness, 2012].

In trying to understand brain circuits and architectures, neuronal stimulation techniques have allowed neuroscientists to uncover primitive behaviors controlled by subcortical structures such as feeding [Cai et al., 2014; Sweeney and Yang, 2015], drinking [Oka et al., 2015], mating [Lee et al., 2014], aggression [Lin et al., 2011], fear [Haubensak et al., 2010] and even loneliness [Matthews et al., 2016]. So this again poses the query of the role of cortex. Did it start evolving for planning purposes, allowing animals to prepare, for example, for upcoming winters where the better the planning the higher the chances of survival? That could explain why more cortical matter kept evolving and why animals can understand that postponing a reward could result in bigger future rewards (humans do this on the timescale of years and decades). Another essential question is why the brain evolved its current architecture? How and why is cortex divided into distinct roles such as the different sensory and motor cortices? What is the role of the prefrontal cortex? And what is special about cortex to allow for long term planning? On a smaller scale, what are the neuronal gene regulatory

mechanisms to allow for this? And what are the roles of plasticity?

The presented chapters show how some of these questions can be answered by studying the brain as a network. As discussed, networks can be viewed at multiple scales and all levels will be necessary and complementary to understand the computational principles performed by the brain. Of course this work goes hand in hand with other non-network centric research which is just as essential in our neuroscientific quest to answer the multitude of puzzling questions facing us.

Bibliography

Abeles, M., Gat, I., 2001. Detecting precise firing sequences in experimental data. Journal of Neuroscience Methods 107 (1-2), 141–154.

Adrian, E. D., Zotterman, Y., 1926. The impulses produced by sensory nerve-endings: Part ii. the response of a single end-organ. The Journal of Physiology 61 (2), 151–171.

Aertsen, A., Diesmann, M., Gewaltig, M., 1996. Propagation of synchronous spiking activity in feedforward neural networks. Journal of Physiology-Paris 90 (3-4), 243–247.

Aggarwal, C., Reddy, C., 2014. Data Clustering: Algorithms and Applications. Taylor & Francis Group.

Aghajanian, G., Bloom, F., 1967. The formation of synaptic junctions in developing rat brain: A quantitative electron microscopic study. Brain Research 6 (4), 716–727.

Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M., Keller, P. J., 2013. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nature Methods 10 (5), 413+.

Alexander, A. L., Lee, J. E., Lazar, M., Field, A. S., 2007. Diffusion tensor imaging of the brain. Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics 4 (3), 316–329.

Alivisatos, A., Chun, M., Church, G., Greenspan, R., Roukes, M., Yuste, R., 2012. The brain activity map project and the challenge of functional connectomics. Neuron 74 (6), 970–974.

Ambrosingerson, J., Granger, R., Lynch, G., 1990. Simulation of paleocortex performs hierarchical clustering. Science 247 (4948), 1344–1348.

Ashby, M., Isaac, J., 2011. Maturation of a recurrent excitatory neocortical circuit by experience-dependent unsilencing of newly formed dendritic spines. Neuron 70 (3), 510–521.

Aviel, Y., Mehring, C., Abeles, M., Horn, D., 2003. On embedding synfire chains in a balanced network. Neural Comput 15 (6), 1321–1340.

Awatramani, G., Slaughter, M., 2000. Origin of transient and sustained responses in ganglion cells of the retina. Journal of Neuroscience 20 (18), 7087–7095.

Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E., Leite, R. E., Filho, W. J., Lent, R., Herculano-Houzel, S., 2009. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. Journal of Comparative Neurology 513 (5), 532–541.

Barabsi, A.-L., Albert, R., 1999. Emergence of Scaling in Random Networks. Science 286 (5439), 509–512.

Barahona, M., Pecora, L. M., 2002. Synchronization in Small-World Systems. Physical Review Letters 89, 054101.

Barthó, P., Hirase, H., Monconduit, L., Zugaro, M., Harris, K. D., Buzsáki, G., 2004. Characterization of Neocortical Principal Cells and Interneurons by Network Interactions and Extracellular Features. Journal of Neurophysiology 92 (1), 600–608.

Basser, P. J., Mattiello, J., LeBihan, D., 1994. Mr diffusion tensor spectroscopy and imaging. Biophysical Journal 66 (1), 259–267.

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. In vivo fiber tractography using dt-mri data. Magnetic Resonance in Medicine 44 (4), 625–632.

Bassett, D. S., Brown, J. A., Deshpande, V., Carlson, J. M., Grafton, S. T., 2011. Conserved and variable architecture of human white matter connectivity. NeuroImage 54 (2), 1262–1279.

Batagelj, V., Brandes, U., 2005. Efficient generation of large random networks. Physical Review E 71, 036113.

Beggs, J., Plenz, D., 2003. Neuronal avalanches in neocortical circuits. Journal of Neuroscience 23 (35), 11167–11177.

Beguerisse-Díaz, M., Garduño Hernández, G., Vangelov, B., Yaliraki, S. N., Barahona, M., 2014. Interest communities and flow roles in directed networks: the Twitter network of the UK riots. Journal of The Royal Society Interface 11 (101).

Behrens, T. E., Sporns, O., 2012. Human connectomics. Current Opinion in Neurobiology 22 (1), 144–153.

Ben-Yishai, R., Bar-Or, R., Sompolinsky, H., 1995. Theory of orientation tuning in visual-cortex. Proceedings of the National Academy of Sciences of the United States of America 92 (9), 3844–3848.

Bernardi, G., Cecchetti, L., Siclari, F., Buchmann, A., Yu, X., Handjaras, G., Bellesi, M., Ricciardi, E., Kecskemeti, S. R., Riedner, B. A., Alexander, A. L., Benca, R. M., Ghilardi, M. F., Pietrini, P., Cirelli, C., Tononi, G., 2016. Sleep reverts changes in human gray and white matter caused by wake-dependent training. NeuroImage 129, 367–377.

Bertil, H., 2001. Ion Channels of Excitable Membranes, 3rd Edition. Sinauer Associates.

Betzel, R. F., Erickson, M. A., Abell, M., O'Donnell, B. F., Hetrick, W. P., Sporns, O., 2012. Synchronization dynamics and evidence for a repertoire of network states in resting eeg. Frontiers in Computational Neuroscience 6, 74–.

Billeh, Y. N., Rodriguez, A. V., Bellesi, M., Bernard, A., de Vivo, L., Funk, C. M., Harris, J., Honjoh, S., Mihalas, S., Ng, L., Koch, C., Cirelli, C., Tononi, G., 2016. Effects of chronic sleep restriction during early adolescence on the adult pattern of connectivity of mouse secondary motor cortex. eNeuro, –.

Billeh, Y. N., Schaub, M. T., Anastassiou, C. A., Barahona, M., Koch, C., 2014. Revealing cell assemblies at multiple levels of granularity. Journal of Neuroscience Methods 236 (0), 92 – 106.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P10008.

Borbely, A. A., Daan, S., Wirz-Justice, A., Deboer, T., 2016. The two-process model of sleep regulation: a reappraisal. Journal of Sleep Research.

Bortone, D., Olsen, S., Scanziani, M., 2014. Translaminar inhibitory cells recruited by layer 6 corticothalamic neurons suppress visual cortex. Neuron 82 (2), 474–485.

Briggman, K. L., Helmstaedter, M., Denk, W., 2011. Wiring specificity in the direction-selectivity circuit of the retina. Nature 471 (7337), 183–188.

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual Web search engine. Computer Networks and Systems 30 (1-7), 107 – 117.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience 10 (3), 186–198.

Bullmore, E., Sporns, O., 2012. The economy of brain network organization. Nature Reviews Neuroscience 13 (5), 336–349.

Buzsaki, G., 2004. Large-scale recording of neuronal ensembles. Nature Neuroscience 7 (5), 446–451.

Buzsaki, G., 2010. Neural Syntax: Cell Assemblies, Synapsembles, and Readers. Neuron 68 (3), 362–385.

Buzsaki, G., Mizuseki, K., 2014. The log-dynamic brain: how skewed distributions affect network operations. Nature Reviews Neuroscience 15 (4), 264–278.

Buzsaki, G., Moser, E. I., 2013. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. Nature Neuroscience 16 (2), 130–138.

Cai, H., Haubensak, W., Anthony, T. E., Anderson, D. J., 2014. Central amygdala pkc-[delta]+ neurons mediate the influence of multiple anorexigenic signals. Nature Neuroscience 17 (9), 1240–1248.

Cancedda, L., Putignano, E., Sale, A., Viegi, A., Berardi, N., Maffei, L., 2004. Acceleration of visual system development by environmental enrichment. The Journal of Neuroscience 24 (20), 4840–4848.

Cao, V., Ye, Y., Mastwal, S., Ren, M., Coon, M., Liu, Q., Costa, R., Wang, K., 2015. Motor learning consolidates arc-expressing neuronal ensembles in secondary motor cortex. Neuron 86 (6), 1385–1392.

Cateau, H., Fukai, T., 2001. Fokker-Planck approach to the pulse packet propagation in synfire chain. Neural Networks 14 (6-7, SI), 675–685.

Chamberlin, N. L., Du, B., de Lacalle, S., Saper, C. B., 1998. Recombinant adeno-associated virus vector: use for transgene expression and anterograde tract tracing in the cns. Brain Research 793 (12), 169–175.

Chen, S. X., Kim, A. N., Peters, A. J., Komiyama, T., 2015. Subtype-specific plasticity of inhibitory circuits in motor cortex during motor learning. Nature Neuroscience 18 (8), 1109–1115.

Chu, C. J., Kramer, M. A., Pathmanathan, J., Bianchi, M. T., Westover, M. B., Wizon, L., Cash, S. S., 2012. Emergence of stable functional networks in long-term human electroencephalography. The Journal of Neuroscience 32 (8), 2703–2713.

Clopath, C., Buesing, L., Vasilaki, E., Gerstner, W., 2010. Connectivity reflects coding: a model of voltage-based STDP with homeostasis. Nature Neuroscience 13 (3), 344–352.

Conturo, T. E., Lori, N. F., Cull, T. S., Akbudak, E., Snyder, A. Z., Shimony, J. S., McKinstry, R. C., Burton, H., Raichle, M. E., 1999. Tracking neuronal fiber pathways in the living human brain. Proceedings of the National Academy of Sciences of the United States of America 96 (18), 10422–10427.

Cooper-Knock, J., Kirby, J., Ferraiuolo, L., Heath, P. R., Rattray, M., Shaw, P. J., 2012. Gene expression profiling in human neurodegenerative disease. Nature Reviews Neurology 8 (9), 518–530.

Cui, J., Tkachenko, O., Gogel, H., Kipman, M., Preer, L. A., Weber, M., Divatia, S. C., Demers, L. A., Olson, E. A., Buchholz, J. L., Bark, J. S., Rosso, I. M., Rauch, S. L., Killgore, W. D., 2015. Microstructure of frontoparietal connections predicts individual resistance to sleep deprivation. NeuroImage 106, 123–133.

Davie, C. A., 2008. A review of parkinson's disease. British Medical Bulletin 86 (1), 109–127.

De Felipe, J., Marco, P., Fairn, A., Jones, E. G., 1997. Inhibitory synaptogenesis in mouse somatosensory cortex. Cerebral Cortex 7 (7), 619–634.

de Vivo, L., Nelson, A. B., Bellesi, M., Noguti, J., Tononi, G., Cirelli, C., 2016. Loss of sleep affects the ultrastructure of pyramidal neurons in the adolescent mouse frontal cortex. Sleep in press.

Delmotte, A., Tate, E. W., Yaliraki, S. N., Barahona, M., 2011. Protein multi-scale organization through graph partitioning and robustness analysis: application to the myosin-myosin light chain interaction. Physical Biology 8 (5), 055010.

Delvenne, J.-C., Schaub, M. T., Yaliraki, S. N., Barahona, M., 2013. The Stability of a Graph Partition: A Dynamics-Based Framework for Community Detection. In: Mukherjee, A., Choudhury, M., Peruani, F., Ganguly, N., Mitra, B. (Eds.), Dynamics On and Of Complex Networks, Volume 2. Springer New York, pp. 221–242.

Delvenne, J.-C., Yaliraki, S. N., Barahona, M., 2010. Stability of graph communities across time scales. Proceedings of the National Academy of Sciences 107 (29), 12755–12760.

Denk, W., Briggman, K. L., Helmstaedter, M., 2012. Structural neurobiology: missing link to a mechanistic understanding of neural computation. Nature Reviews Neuroscience 13 (5), 351–358.

Diba, K., Buzsaki, G., 2007. Forward and reverse hippocampal place-cell sequences during ripples. Nature Neuroscience 10 (10), 1241–1242.

Diesmann, M., Gewaltig, M., Aertsen, A., 1999. Stable propagation of synchronous spiking in cortical neural networks. Nature 402 (6761), 529–533.

Du, J., Blanche, T. J., Harrison, R. R., Lester, H. A., Masmanidis, S. C., 2011. Multiplexed, High Density Electrophysiology with Nanofabricated Neural Probes. PLoS ONE 6 (10).

Ernst, U., Pawelzik, K., Sahar-Pikielny, C., Tsodyks, M., 2001. Intracortical origin of visual maps. Nature Neuroscience 4 (4), 431–436.

Feldt, S., Waddell, J., Hetrick, V. L., Berke, J. D., Żochowski, M., 2009. Functional clustering algorithm for the analysis of dynamic network data. Physical Review E 79, 056104.

Felleman, D. J., Van Essen, D. C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex 1 (1), 1–47.

Fellous, J., Tiesinga, P., Thomas, P., Sejnowski, T., 2004. Discovering spike patterns in neuronal responses. Journal of Neuroscience 24 (12), 2989–3001.

Fortunato, S., 2010. Community detection in graphs. Physics Reports 486 (3-5), 75 – 174.

Fortunato, S., Barthélemy, M., 2007. Resolution limit in community detection. Proceedings of the National Academy of Sciences 104 (1), 36–41.

Foster, D., Wilson, M., 2006. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature 440 (7084), 680–683.

Frank, M., Heller, H., 1997. Development of diurnal organization of eeg slow-wave activity and slow- wave sleep in the rat. American Journal of Physiology 273, R472–478.

Frank, M. G., 2011. Chapter 14 - sleep and developmental plasticity: not just for kids. In: Eus J.W. Van Someren, Ysbrand D. Van Der Werf, P. R. R. H. D. M., Silva, F. H. L. D. (Eds.), Slow Brain Oscillations of Sleep, Resting State and Vigilance. Vol. Volume 193. Elsevier, pp. 221–232.

Frank, M. G., Issa, N. P., Stryker, M. P., 2001. Sleep enhances plasticity in the developing visual cortex. Neuron 30 (1), 275–287.

Furness, J. B., 2012. The enteric nervous system and neurogastroenterology. Nature Reviews Gastroenterology & Hepatology 9 (5), 286–294.

Galán, R. F., 2008. On How Network Architecture Determines the Dominant Patterns of Spontaneous Neural Activity. PLoS ONE 3 (5), e2148.

Gansel, K. S., Singer, W., 2012. Detecting multineuronal temporal patterns in parallel spike trains. Frontiers in Neuroinformatics 6 (18).

Gerstner, W., Kistler, W. M., Naud, R., L., P., 2014. Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition. Cambridge University Press.

Gewaltig, M., Diesmann, M., Aertsen, A., 2001. Propagation of cortical synfire activity: survival probability in single trials and stability in the mean. Neural Networks 14 (6-7, SI), 657–673.

Gillies, G. E., McArthur, S., 2010. Estrogen actions in the brain and the basis for differential action in men and women: A case for sex-specific medicines. Pharmacological Reviews 62 (2), 155–198.

Goldman, M. S., 2009. Memory without Feedback in a Neural Network. Neuron 61 (4), 621–634.

Gollisch, T., Meister, M., 2008. Rapid neural coding in the retina with relative spike latencies. Science 319 (5866), 1108–1111.

Golub, H., Van Loan, F. C., 1996. Matrix Computations, 3rd Edition. John Hopkins University Press, Baltimore and London.

Gramsbergen, A., 1976. The development of the eeg in the rat. Developmental Psychobiology 9 (6), 501–515.

Grubb, M. S., Rossi, F. M., Changeux, J.-P., Thompson, I. D., 2003. Abnormal functional organization in the dorsal lateral geniculate nucleus of mice lacking the 2 subunit of the nicotinic acetylcholine receptor. Neuron 40 (6), 1161–1172.

Hahnloser, R., Kozhevnikov, A., Fee, M., 2002. An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419 (6902), 65–70.

Harris, J. A., Wook Oh, S., Zeng, H., 2012. Adeno-associated viral vectors for anterograde axonal tracing with fluorescent proteins in nontransgenic and cre driver mice. In: Current Protocols in Neuroscience. John Wiley & Sons, Inc., pp. –.

Harris, K., 2005. Neural signatures of cell assembly organization. Nature Reviews Neuroscience 6 (5), 399–407.

Harris, K., Csicsvari, J., Hirase, H., Dragoi, G., Buzsaki, G., 2003. Organization of cell assemblies in the hippocampus. Nature 424 (6948), 552–556.

Harris, K. D., Mrsic-Flogel, T. D., 2013. Cortical connectivity and sensory coding. Nature 503 (7474), 51–58.

Harris, K. D., Shepherd, G. M. G., 2015. The neocortical circuit: themes and variations. Nature Neuroscience 18 (2), 170–181.

Haubensak, W., Kunwar, P. S., Cai, H., Ciocchi, S., Wall, N. R., Ponnusamy, R., Biag, J., Dong, H.-W., Deisseroth, K., Callaway, E. M., Fanselow, M. S., Luthi, A., Anderson, D. J., 2010. Genetic dissection of an amygdala microcircuit that gates conditioned fear. Nature 468 (7321), 270–276.

Hebb, D. O., 1949. The Organization of Behavior. JohnWiley & Sons.

Hendry, S. H. C., Reid, R. C., 2000. The koniocellular pathway in primate vision. Annual Reviews Neuroscience 23 (1), 127–153.

Herculano-Houzel, S., 2012. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. Proceedings of the National Academy of Sciences of the United States of America 109 (Suppl 1), 10661–10668.

Herculano-Houzel, S., Mota, B., Lent, R., 2006. Cellular scaling rules for rodent brains. Proceedings of the National Academy of Sciences 103 (32), 12138–12143.

Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., Hazen, N., Herman, J., Adams Hillard, P. J., Katz, E. S., Kheirandish-Gozal, L., Neubauer, D. N., ODonnell, A. E., Ohayon, M., Peever, J., Rawding, R., Sachdeva, R. C., Setters, B., Vitiello, M. V., Ware, J. C., 2015. National sleep foundations updated sleep duration recommendations: final report. Sleep Health 1 (4), 233–243.

Hodgkin, A. L., Huxley, A. F., 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. The Journal of Physiology 117 (4), 500–544.

Hromadka, T., DeWeese, M. R., Zador, A. M., 2008. Sparse representation of sounds in the unanesthetized auditory cortex. PLoS Biology 6 (1), 124–137.

Huang, Z. J., 2014. Toward a Genetic Dissection of Cortical Circuits in the Mouse. Neuron 83 (6), 1284–1302.

Humphries, M. D., 2011. Spike-Train Communities: Finding Groups of Similar Spike Trains. Journal of Neuroscience 31 (6), 2321–2336.

Hyman, J., Wyble, B., Goyal, V., Rossi, C., Hasselmo, M., 2003. Stimulation in hippocampal region CA1 in behaving rats yields long-term potentiation when delivered to the peak of theta and long-term depression when delivered to the trough. Journal of Neuroscience 23 (37), 11725–11731.

Innocenti, G. M., Price, D. J., 2005. Exuberance in the development of cortical networks. Nature Reviews Neuroscience 6 (12), 955–965.

Isaacson, J. S., Scanziani, M., 2011. How Inhibition Shapes Cortical Activity. Neuron 72 (2), 231–243.

Ito, M., Masuda, N., Shinomiya, K., Endo, K., Ito, K., 2013. Systematic Analysis of Neural Projections Reveals Clonal Composition of the Drosophila Brain. Current Biology 23 (8), 644–655.

J Pavlides, C., Greenstein, Y., Grudman, M., J, W., 1988. Long-term potentiation in the dentate gyrus is induced preferentially on the positive phase of theta-rhythm. Brain Research 439 (1-2), 383–387.

Jbabdi, S., Johansen-Berg, H., 2011. Tractography - where do we go from here? Brain connectivity 1 (3), 169–183.

Johansen-Berg, H., 2013. Human connectomics what will the future demand? NeuroImage 80, 541–544.

Johnston, D., Wu, S. M.-S., 1995. Foundations of Cellular Neurophysiology. The MIT Press.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., Hudspeth, A. J., 2013. Principles of Neural Science, 5th ed., 5th Edition. McGraw-Hill.

Kayser, M. S., Yue, Z., Sehgal, A., 2014. A critical period of sleep for development of courtship circuitry and behavior in drosophila. Science 344 (6181), 269–274.

Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., Arieli, A., 2003. Spontaneously emerging cortical representations of visual attributes. Nature 425 (6961), 954–956.

Kennedy, M. B., 2000. Signal-processing machines at the postsynaptic density. Science 290 (5492), 750–754.

Kepecs, A., Fishell, G., 2014. Interneuron cell types are fit to function. Nature 505, 318326.

Kistler, W., Gerstner, W., 2002. Stable propagation of activity pulses in populations of spiking neurons. Neural Computation 14 (5), 987–997.

Klausberger, T., Somogyi, P., 2008. Neuronal diversity and temporal dynamics: The unity of hippocampal circuit operations. Science 321 (5885), 53–57.

Kleim, J. A., Bruneau, R., Calder, K., Pocock, D., VandenBerg, P. M., MacDonald, E., Monfils, M. H., Sutherland, R. J., Nader, K., 2003. Functional organization of adult motor cortex is dependent upon continued protein synthesis. Neuron 40 (1), 167–176.

Ko, H., Cossell, L., Baragli, C., Antolik, J., Clopath, C., Hofer, S. B., Mrsic-Flogel, T. D., 2013. The emergence of functional microcircuits in visual cortex. Nature 496 (7443), 96–100.

Ko, H., Hofer, S. B., Pichler, B., Buchanan, K. A., Sjoestroem, P. J., Mrsic-Flogel, T. D., 2011. Functional specificity of local synaptic connections in neocortical networks. Nature 473 (7345), 87–U100.

Koch, C., 1999. Biophysics of computation: information processing in single neurons. Oxford University Press.

Koester, S., O'Leary, D., 1992. Functional classes of cortical projection neurons develop dendritic distinctions by class-specific sculpting of an early common pattern. The Journal of Neuroscience 12 (4), 1382–1393.

Kohara, K., Pignatelli, M., Rivest, A. J., Jung, H.-Y., Kitamura, T., Suh, J., Frank, D., Kajikawa, K., Mise, N., Obata, Y., Wickersham, I. R., Tonegawa, S., 2014. Cell type-specific genetic and optogenetic tools reveal hippocampal ca2 circuits. Nature Neuroscience 17 (2), 269–279.

Kremkow, J., Perrinet, L. U., Masson, G. S., Aertsen, A., 2010. Functional consequences of correlated excitatory and inhibitory conductances in cortical networks. Journal of Computational Neuroscience 28 (3), 579–594.

Kreuz, T., Chicharro, D., Houghton, C., Andrzejak, R. G., Mormann, F., 2013. Monitoring spike train synchrony. Journal of Neurophysiology 109 (5), 1457–1472.

Kuan, L., Li, Y., Lau, C., Feng, D., Bernard, A., Sunkin, S. M., Zeng, H., Dang, C., Hawrylycz, M., Ng, L., 2015. Neuroinformatics of the allen mouse brain connectivity atlas. Methods 73, 4–17.

Kumar, A., Rotter, S., Aertsen, A., 2008. Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. Journal of Neuroscience 28 (20), 5268–5280.

Kumar, A., Rotter, S., Aertsen, A., 2010. Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. Nature Reviews Neuroscience 11 (9), 615–627.

Lambiotte, R., Delvenne, J.-C., Barahona, M., 2009. Laplacian Dynamics and Multiscale Modular Structure in Networks. arxiv:0812.1770, 1–29.

Lambiotte, R., Rosvall, M., 2012. Ranking and clustering of nodes in networks with smart teleportation. Physical Review E 85, 056107.

Larremore, D. B., Shew, W. L., Restrepo, J. G., 2011. Predicting Criticality and Dynamic Range in Complex Networks: Effects of Topology. Physical Review Letters 106 (5).

Laubach, M., Shuler, M., Nicolelis, M., 1999. Independent component analyses for quantifying neuronal ensemble interactions. Journal of Neuroscience Methods 94 (1), 141–154.

Laughlin, S., Sejnowski, T., 2003. Communication in neuronal networks. Science 301 (5641), 1870–1874.

Lee, H., Kim, D.-W., Remedios, R., Anthony, T. E., Chang, A., Madisen, L., Zeng, H., Anderson, D. J., 2014. Scalable control of mounting and attack by esr1+ neurons in the ventromedial hypothalamus. Nature 509 (7502), 627–632.

Lee, S., Kruglikov, I., Huang, Z. J., Fishell, G., Rudy, B., 2013. A disinhibitory circuit mediates motor integration in the somatosensory cortex. Nature Neuroscience 16 (11), 1662–1670.

Lefort, S., Tomm, C., Sarria, J. C. F., Petersen, C. C. H., 2009. The Excitatory Neuronal Network of the C2 Barrel Column in Mouse Primary Somatosensory Cortex. Neuron 61 (2), 301–316.

Leger, D., Beck, F., Richard, J.-B., Godeau, E., 2012. Total sleep time severely drops during adolescence. PLoS ONE 7 (10), e45204–.

Lemon, R. N., Edgley, S. A., 2010. Life without a cerebellum. Brain 133 (3), 652–654.

Leuner, B., Shors, T., 2013. Stress, anxiety, and dendritic spines: What are the connections? Neuroscience 251, 108–119.

Li, Z., Okamoto, K.-I., Hayashi, Y., Sheng, M., 2004. The importance of dendritic mitochondria in the morphogenesis and plasticity of spines and synapses. Cell 119 (6), 873–887.

Lin, D., Boyle, M. P., Dollar, P., Lee, H., Lein, E. S., Perona, P., Anderson, D. J., 2011. Functional identification of an aggression locus in the mouse hypothalamus. Nature 470 (7333), 221–226.

Litvak, V., Sompolinsky, H., Segev, I., Abeles, M., 2003. On the transmission of rate code in long feedforward networks with excitatory-inhibitory balance. Journal of Neuroscience 23 (7), 3006–3015.

Litwin-Kumar, A., Doiron, B., 2012. Slow dynamics and high variability in balanced cortical networks with clustered connections. Nature Neuroscince 15 (11), 1498–1505.

Litwin-Kumar, A., Doiron, B., 2014. Formation and maintenance of neuronal assemblies through synaptic plasticity. Nature Communications 5.

Llorens-Martin, M., Blazquez-Llorca, L., Benavides-Piccione, R., Rabano, A., Hernandez, F., Avila, J., DeFelipe, J., 2014. Selective alterations of neurons and circuits related to early memory loss in alzheimer's disease. Frontiers in Neuroanatomy 8, –.

Lopes-Dos-Santos, V., Conde-Ocazionez, S., Nicolelis, M. A. L., Ribeiro, S. T., Tort, A. B. L., 2011. Neuronal Assembly Detection and Cell Membership Specification by Principal Component Analysis. PLoS ONE 6 (6), e20996.

Lopes-Dos-Santos, V., Ribeiro, S., Tort, A. B. L., 2013. Detecting cell assemblies in large neuronal populations. Journal of Neuroscience Methods 220 (2), 149–66.

Lyttle, D., Fellous, J.-M., 2011. A new similarity measure for spike trains: Sensitivity to bursts and periods of inhibition. Journal of Neuroscience Methods 199 (2), 296–309.

Mackiewicz, M., Naidoo, N., Zimmerman, J. E., Pack, A. I., 2008. Molecular mechanisms of sleep and wakefulness. Annals of the New York Academy of Sciences 1129 (1), 335–349.

Maravall, M., Koh, I. Y., Lindquist, W. B., Svoboda, K., 2004. Experience-dependent changes in basal dendritic branching of layer 2/3 pyramidal neurons during a critical period for developmental plasticity in rat barrel cortex. Cerebral Cortex 14 (6), 655–664.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., Wu, C., 2004. Interneurons of the neocortical inhibitory system. Nature Reviews Neuroscience 5 (10), 793–807.

Matthews, G., Nieh, E., VanderWeele, C., Halbert, S., Pradhan, R., Yosafat, A., Glober, G., Izadmehr, E., Thomas, R., Lacy, G., Wildes, C., Ungless, M., Tye, K., 2016. Dorsal raphe dopamine neurons represent the experience of social isolation. Cell 164 (4), 617–631.

McGinley, M. J., Westbrook, G. L., 2013. Hierarchical excitatory synaptic connectivity in mouse olfactory cortex. Proceedings of the National Academy of Sciences of the United States of America 110 (40), 16193–16198.

Mehring, C., Hehl, U., Kubo, M., Diesmann, M., Aertsen, A., 2003. Activity dynamics and propagation of synchronous spiking in locally connected random networks. Biological Cybernetics 88 (5), 395–408.

Meila, M., 2003. Comparing clusterings by the variation of information. In: Scholkopf, B., Warmuth, M. (Eds.), Learning Theory and Kernel Machines. Vol. 2777 of Lecture Notes in Artificial Intelligence. Springer-Verlag Berlin, pp. 173–187.

Meila, M., 2007. Comparing clusterings–an information based distance. Journal of Multivariate Analysis 98 (5), 873 – 895.

Meister, M., Pine, J., Baylor, D., 1994. Multi-neuronal Signals from the Retina - Acquisition and Analysis. Journal of Neuroscience Methods 51 (1), 95–106.

Merboldt, K.-D., Hanicke, W., Frahm, J., 1985. Self-diffusion nmr imaging using stimulated echoes. Journal of Magnetic Resonance (1969) 64 (3), 479–486.

Meunier, D., Lambiotte, R., Bullmore, E. T., 2010. Modular and hierarchically modular organization of brain networks. Frontiers in Neuroscience 4 (200).

Micheva, K. D., Beaulieu, C., 1996. Quantitative aspects of synaptogenesis in the rat barrel field cortex with special reference to gaba circuitry. Journal of Comparative Neurology 373 (3), 340–354.

Mong, J. A., Baker, F. C., Mahoney, M. M., Paul, K. N., Schwartz, M. D., Semba, K., Silver, R., 2011. Sleep, rhythms, and the endocrine brain: Influence of sex and gonadal hormones. Journal of Neuroscience 31 (45), 16107–16116.

Murphy, B. K., Miller, K. D., 2009. Balanced Amplification: A New Mechanism of Selective Amplification of Neural Activity Patterns. Neuron 61 (4), 635 – 648.

Nadakuditi, R. R., Newman, M. E. J., 2013. Spectra of random graphs with arbitrary expected degrees. Physical Review E 87, 012803.

Nasrallah, K., Piskorowski, R. A., Chevaleyre, V., 2015. Inhibitory plasticity permits the recruitment of ca2 pyramidal neurons by ca3. eNeuro 2(4), 1–12.

Nelson, B. A., Faraguna, U., Zoltan, T. J., Tononi, G., Cirelli, C., 2013. Sleep patterns and homeostatic mechanisms in adolescent mice.

Newman, M., 2004. Detecting community structure in networks. European Physical Journal B 38 (2), 321–330.

Newman, M. E. J., Girvan, M., 2004. Finding and evaluating community structure in networks. Physical Review E 69 (2), 026113.

Niven, J. E., Laughlin, S. B., 2008. Energy limitation as a selective pressure on the evolution of sensory systems. Journal of Experimental Biology 211 (11), 1792–1804.

Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A. M., Mortrud, M. T., Ouellette, B., Nguyen, T. N., Sorensen, S. A., Slaughterbeck, C. R., Wakeman, W., Li, Y., Feng, D., Ho, A., Nicholas, E., Hirokawa, K. E., Bohn, P., Joines, K. M., Peng, H., Hawrylycz, M. J., Phillips, J. W., Hohmann, J. G., Wohnoutka, P., Koch, C., Bernard, A., Dang, C., Jones, A. R., Zeng, H., Gerfen, C. R., 2014. A mesoscale connectome of the mouse brain. Nature 508 (7495), 207–214.

Oka, Y., Ye, M., Zuker, C. S., 2015. Thirst driving and suppressing signals encoded by distinct neural populations in the brain. Nature 520 (7547), 349–352.

Okatan, M., Wilson, M. A., Brown, E. N., 2005. Analyzing Functional Connectivity Using a Network Likelihood Model of Ensemble Neural Spiking Activity. Neural Computation 17 (9), 1927–1961.

Olsen, S. R., Bortone, D. S., Adesnik, H., Scanziani, M., 2012. Gain control by layer six in cortical circuits of vision. Nature 483 (7387), 47–52.

Otsuka, T., Kawaguchi, Y., 2011. Cell Diversity and Connection Specificity between Callosal Projection Neurons in the Frontal Cortex. Journal of Neuroscience 31 (10), 3862–3870.

Paul, K., Dugovic, C., Turek, F., Laposky, A., 2006. Diurnal sex differences in the sleep-wake cycle of mice are dependent on gonadal function. Sleep 29, 1211–1223.

Paul, K. N., Losee-Olson, S., Pinckney, L., Turek, F. W., 2009. The ability of stress to alter sleep in mice is sensitive to reproductive hormones. Brain Research 1305, 74–85.

Paus, T., Keshavan, M., Giedd, J. N., 2008. Why do many psychiatric disorders emerge during adolescence? Nature Reviews Neuroscience 9 (12), 947–957.

Perin, R., Berger, T. K., Markram, H., 2011. A synaptic organizing principle for cortical neuronal groups. Proceedings of the National Academy of Sciences of the United States of America 108 (13), 5419–5424.

Peyrache, A., Benchenane, K., Khamassi, M., Wiener, S. I., Battaglia, F. P., 2010. Principal component analysis of ensemble recordings reveals cell assemblies at high temporal resolution. Journal of Computational Neuroscience 29 (1-2, SI), 309–325.

Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., Scanziani, M., 2013. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nature Neuroscience 16 (8), 1068–1076.

Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J. I., Huang, Z. J., Kepecs, A., 2013. Cortical interneurons that specialize in disinhibitory control. Nature 503 (7477), 521+.

Pluta, S., Naka, A., Veit, J., Telian, G., Yao, L., Hakim, R., Taylor, D., Adesnik, H., 2015. A direct translaminar inhibitory circuit tunes cortical output. Nature Neuroscience 18 (11), 1631–1640.

Pompeiano, O., Pompeiano, M., Corvaja, N., 1995. Effects of sleep deprivation on the postnatal development of visual-deprived cells in the cat's lateral geniculate nucleus. Archives Italiennes de Biologie 134, 121–140.

Pouzat, C., Mazor, O., Laurent, G., 2002. Using noise signature to optimize spike-sorting and to assess neuronal classification quality. Journal of Neuroscience Methods 122 (1), 43–57.

Prescott, A. R., Bales, A., James, J., Trinkle-Mulcahy, L., Sleeman, J. E., 2014. Time-resolved quantitative proteomics implicates the core snrnp protein smb together with smn in neural trafficking. Journal of Cell Science 127 (4), 812–827.

Quiroga, R. Q., Panzeri, S., 2009. Extracting information from neuronal populations: information theory and decoding approaches. Nature Reviews Neuroscience 10 (3), 173–185.

Rajan, K., Abbott, L. F., 2006. Eigenvalue Spectra of Random Matrices for Neural Networks. Physical Review Letters 97, 188104.

Ramón y Cajal, S., 1888. Estructura de los centros nerviosos de las aves.

Ramón y Cajal, S., 1933. Histology, 10th Edition. Wood.

Rechtschaffen, A., Bergmann, B., 2002. Sleep deprivation in the rat: an update of the 1989 paper. Sleep 25, 18–24.

Reep, R. L., Wu, J. H., Cheatwood, J. L., Corwin, J. V., Kartje, G. L., Mir, A., 2008. Quantification of synaptic density in corticostriatal projections from rat medial agranular cortex. Brain Research 1233, 27–34.

Reimann, M. W., Anastassiou, C. A., Perin, R., Hill, S. L., Markram, H., Koch, C., 2013. A Biophysically Detailed Model of Neocortical Local Field Potentials Predicts the Critical Role of Active Membrane Currents. Neuron 79 (2), 375–390.

Reyes, A., 2003. Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. Nature Neuroscience 6 (6), 593–599.

Riehle, A., Grun, S., Diesmann, M., Aertsen, A., 1997. Spike synchronization and rate modulation differentially involved in motor cortical function. Science 278 (5345), 1950–1953.

Rocklage, M., Williams, V., Pacheco, J., Schnyer, D., 2009. White matter differences predict cognitive vulnerability to sleep deprivation. Sleep 32, 1100–1103.

Roenneberg, T., Kuehnle, T., Juda, M., Kantermann, T., Allebrandt, K., Gordijn, M., Merrow, M., 2007. Epidemiology of the human circadian clock. Sleep Medicine Reviews 11 (6), 429–438.

Roenneberg, T., Kuehnle, T., Pramstaller, P. P., Ricken, J., Havel, M., Guth, A., Merrow, M., 2004. A marker for the end of adolescence. Current Biology 14 (24), R1038–R1039.

Roffwarg, H. P., Muzio, J. N., Dement, W. C., 1966. Ontogenetic development of the human sleep-dream cycle. Science 152 (3722), 604–619.

Rokem, A., Watzl, S., Gollisch, T., Stemmler, M., Herz, A., Samengo, I., 2006. Spike-timing precision underlies the coding efficiency of auditory receptor neurons. Journal of Neurophysiology 95 (4), 2541–2552.

Romand, S., Wang, Y., Toledo-Rodriguez, M., Markram, H., 2011. Morphological development of thick-tufted layer v pyramidal cells in the rat somatosensory cortex. Frontiers in Neuroanatomy 5, –.

Ronan, J. L., Wu, W., Crabtree, G. R., 2013. From neural development to cognition: unexpected roles for chromatin. Nature Reviews Genetics 14 (5), 347–359.

Rosvall, M., Bergstrom, C. T., 2008. Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences 105 (4), 1118–1123.

Roux, L., Buzsaki, G., 2015. Tasks for inhibitory interneurons in intact brain circuits. Neuropharmacology 88 (SI), 10–23.

Rupp, T., Wesensten, N., Balkin, T., 2012. Trait-like vulnerability to total and partial sleep loss. Sleep 35, 1163–1172.

Rutishauser, U., Douglas, R. J., Slotine, J.-J., 2011. Collective stability of networks of winner-take-all circuits. Neural Computation 23 (3), 735–773.

Sanes, J. R., Yamagata, M., 2009. Many paths to synaptic specificity. Annual Review of Cell and Developmental Biology 25 (1), 161–195.

Sare, R. M., Levine, M., Hildreth, C., Picchioni, D., Smith, C. B., 2016. Chronic sleep restriction during development can lead to long-lasting behavioral effects. Physiology & Behavior 155, 208–217.

Sasaki, T., Matsuki, N., Ikegaya, Y., 2007. Metastability of active CA3 networks. Journal of Neuroscience 27 (3), 517–528.

Savic, I., Gulyas, B., Larsson, M., Roland, P., 2000. Olfactory functions are mediated by parallel and hierarchical processing. Neuron 26 (3), 735–745.

Schaub, M. T., Billeh, Y. N., Anastassiou, C. A., Koch, C., Barahona, M., 2015. Emergence of slow-switching assemblies in structured neuronal networks. PLoS Computational Biology 11 (7), e1004196.

Schaub, M. T., Delvenne, J.-C., Yaliraki, S. N., Barahona, M., 2012. Markov Dynamics as a Zooming Lens for Multiscale Community Detection: Non Clique-Like Communities and the Field-of-View Limit. PLoS ONE 7 (2), e32210.

Schreiber, S., Fellous, J., Whitmer, D., Tiesinga, P., Sejnowski, T., 2003. A new correlation-based measure of spike timing reliability. Neurocomputing 52-4, 925–931.

Schwartz, O., Pillow, J. W., Rust, N. C., Simoncelli, E. P., 2006. Spike-triggered neural characterization. Journal of Vision 6 (4), 484–507.

Seelke, A. M. H., Dooley, J. C., Krubitzer, L. A., 2012. The emergence of somatotopic maps of the body in s1 in rats: The correspondence between functional and anatomical organization. PLoS ONE 7 (2), e32322–.

Selhorst, J. B., Chen, Y., 2009. The optic nerve. Seminars in neurology 29 (1), 29–35.

Seung, S., 2012. Connectome - How the Brain's Wiring Makes Us Who We Are. Houghton Mifflin Harcourt Publishing Company.

Shaffery, J., Oksenberg, A., Marks, G., Speciale, S., Mihailoff, G., HP, R., 1998. Rem sleep deprivation in monocularly occluded kittens reduces the size of cells in lgn monocular segment. Sleep 21, 837–845.

Shaffery, J., Sinton, C., Bissette, G., Roffwarg, H., Marks, G., 2002. Rapid eye movement sleep deprivation modifies expression of long-term potentiation in visual cortex of immature rats. Neuroscience 110 (3), 431–443.

Shaffery, J. P., Lopez, J., Roffwarg, H. P., 2012. Brain-derived neurotrophic factor (bdnf) reverses the effects of rapid eye movement sleep deprivation (remsd) on developmentally regulated, long-term potentiation (ltp) in visual cortex slices. Neuroscience Letters 513 (1), 84–88.

Shimono, M., Beggs, J. M., 2014. Functional Clusters, Hubs, and Communities in the Cortical Microconnectome. Cerebral Cortex.

Silberberg, G., Markram, H., 2007. Disynaptic inhibition between neocortical pyramidal cells mediated by martinotti cells. Neuron 53 (5), 735–746.

Simon, H. A., Ando, A., 1961. Aggregation of Variables in Dynamic Systems. Econometrica 29 (2), 111–138.

Snell, R. S., 2006. Clinical neuroanatomy, 6th Edition. Lippincott Williams & Wilkins.

Somers, D., Nelson, S., Sur, M., 1995. An emergent model of orientation selectivity in cat visual cortical simple cells. Journal of Neuroscience 15 (8), 5448–5465.

Sommers, H. J., Crisanti, A., Sompolinsky, H., Stein, Y., 1988. Spectrum of Large Random Asymmetric Matrices. Physical Review Letters 60, 1895–1898.

Song, S., Sjostrom, P., Reigl, M., Nelson, S., Chklovskii, D., 2005. Highly nonrandom features of synaptic connectivity in local cortical circuits. Plos Biology 3 (3), 507–519.

Sotiropoulos, S. N., Jbabdi, S., Xu, J., Andersson, J. L., Moeller, S., Auerbach, E. J., Glasser, M. F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D. A., Yacoub, E., Lenglet, C., Van Essen, D. C., Ugurbil, K., Behrens, T. E., 2013. Advances in diffusion mri acquisition and processing in the human connectome project. NeuroImage 80, 125–143.

Spear, L., 2000. The adolescent brain and age-related behavioral manifestations. Neuroscience & Biobehavioral Reviews 24 (4), 417–463.

Sporns, O., 2013. Network attributes for segregation and integration in the human brain. Current Opinion in Neurobiology 23 (2), 162–171.

Sporns, O., Tononi, G., Ktter, R., 2005. The human connectome: A structural description of the human brain. PLoS Computational Biology 1 (4), e42.

Stewart, G. W., 2001. Matrix Algorithms Volume 2: Eigensystems. Vol. 2. Siam.

Stuesse, S. L., Newman, D. B., 1990. Projections from the medial agranular cortex to brain stem visuomotor centers in rats. Experimental Brain Research 80, 532–544.

Sweeney, P., Yang, Y., 2015. An excitatory ventral hippocampus to lateral septum circuit that suppresses feeding. Nature Communications 6, –.

Taniguchi, H., 2014. Genetic dissection of GABAergic neural circuits in mouse neocortex. Frontiers in Cellular Neuroscience 8.

Tasic, B., Menon, V., Nguyen, T. N., Kim, T. K., Jarsky, T., Yao, Z., Levi, B., Gray, L. T., Sorensen, S. A., Dolbeare, T., Bertagnolli, D., Goldy, J., Shapovalova, N., Parry, S., Lee, C., Smith, K., Bernard, A., Madisen, L., Sunkin, S. M., Hawrylycz, M., Koch, C., Zeng, H., 2016. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nature Neuroscience 19 (2), 335–346.

Tau, G. Z., Peterson, B. S., 2010. Normal development of brain circuits. Neuropsychopharmacology 35 (1), 147–168.

Taxidis, J., Anastassiou, C., Diba, K., Koch, C., 2015. Local field potentials encode place cell ensemble activation during hippocampal sharp wave ripples. Neuron 87 (3), 590–604.

Taylor, D. G., Bushell, M. C., 1985. The spatial mapping of translational diffusion coefficients by the nmr imaging technique. Physics in Medicine and Biology 30 (4), 345–.

Tetzlaff, T., Buschermohle, M., Geisel, T., Diesmann, M., 2003. The spread of rate and correlation in stationary cortical networks. Neurocomputing 52-4, 949–954.

Trefethen, L. N., Embree, M., 2005. Spectra and pseudospectra: the behavior of nonnormal matrices and operators. Princeton University Press.

Uddin, L. Q., Supekar, K., Menon, V., 2010. Typical and atypical development of functional human brain networks: insights from resting-state fmri. Frontiers in Systems Neuroscience 4, –.

Valentino, R. J., Reyes, B., Van Bockstaele, E., Bangasser, D., 2012. Molecular and cellular sex differences at the intersection of stress and arousal. Neuropharmacology 62 (1), 13–20.

Van Dongen, H., Baynard, M., Maislin, G., Dinges, D., 2004. Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. Sleep 27, 423–433.

van Rossum, M., 2001. A novel spike distance. Neural Computation 13 (4), 751–763.

van Rossum, M., Turrigiano, G., Nelson, S., 2002. Fast propagation of firing rates through layered networks of noisy neurons. Journal of Neuroscience 22 (5), 1956–1966.

van Vreeswijk, C., Sompolinsky, H., 1998. Chaotic balanced state in a model of cortical circuits. Neural Computation 10 (6), 1321–1371.

Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., Chklovskii, D. B., 2011. Structural properties of the caenorhabditis elegans neuronal network. PLoS Computational Biology 7 (2), e1001066–.

Vecsey, C. G., Peixoto, L., Choi, J. H. K., Wimmer, M., Jaganath, D., Hernandez, P. J., Blackwell, J., Meda, K., Park, A. J., Hannenhalli, S., Abel, T., 2012. Genomic analysis of sleep deprivation reveals translational regulation in the hippocampus. Physiological Genomics 44 (20), 981–991.

Victor, J., Purpura, K., 1996. Nature and precision of temporal coding in visual cortex: A metric-space analysis. Journal of Neurophysiology 76 (2), 1310–1326.

Vincent, K., Tauskela, J. S., Thivierge, J.-P., 2012. Extracting functionally feedforward networks from a population of spiking neurons. Frontiers in Computational Neuroscience 6.

Vogels, T., Rajan, K., Abbott, L., 2005. Neural network dynamics. Annual Reviews Neuroscience 28, 357–376.

Vogels, T. P., Abbott, L. F., 2005. Signal Propagation and Logic Gating in Networks of Integrate-and-Fire Neurons. Journal of Neuroscience 25 (46), 10786–10795.

Vogels, T. P., Abbott, L. F., 2009. Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. Nature Neuroscience 12 (4), 483–491.

von Luxburg, U., 2007. A tutorial on spectral clustering. Statistics and Computing 17 (4), 395–416.

Wang, G. X., Smith, S. J., Mourrain, P., 2014. Fmr1 ko and fenobam treatment differentially impact distinct synapse populations of mouse neocortex. Neuron 84 (6), 1273–1286.

Watts, D. J., Strogatz, S. H., 1998. Collective dynamics of 'small-world' networks. Nature 393 (6684), 440–442.

White, J. G., Southgate, E., Thomson, J. N., Brenner, S., 1986. The structure of the nervous system of the nematode caenorhabditis elegans. Philosophical Transactions of the Royal Society of London B: Biological Sciences 314 (1165), 1–340.

Wolfson, A. R., Carskadon, M. A., 1998. Sleep schedules and daytime functioning in adolescents. Child Development 69 (4), 875–887.

Xu, H., Jeong, H.-Y., Tremblay, R., Rudy, B., 2013. Neocortical Somatostatin-Expressing GABAergic Interneurons Disinhibit the Thalamorecipient Layer 4. Neuron 77 (1), 155–167.

Yang, G., Lai, C. S. W., Cichon, J., Ma, L., Li, W., Gan, W.-B., 2014. Sleep promotes branch-specific formation of dendritic spines after learning. Science 344 (6188), 1173–1178.

Yao, Z., Zhang, Y., Lin, L., Zhou, Y., Xu, C., Jiang, T., the Alzheimer's Disease Neuroimaging Initiative, 2010. Abnormal cortical networks in mild cognitive impairment and alzheimer's disease. PLoS Computational Biology 6 (11), e1001006.

Yassin, L., Benedetti, B. L., Jouhanneau, J.-S., Wen, J. A., Poulet, J. F. A., Barth, A. L., 2010. An Embedded Subnetwork of Highly Active Neurons in the Neocortex. Neuron 68 (6), 1043–1050.

Yoshimura, Y., Dantzker, J., Callaway, E., 2005. Excitatory cortical neurons form fine-scale functional networks. Nature 433 (7028), 868–873.

Zhang, L. I., Bao, S., Merzenich, M. M., 2002. Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. Proceedings of the National Academy of Sciences 99 (4), 2309–2314.

Zhang, X., Rao Nadakuditi, R., Newman, M. E. J., 2013. Spectra of random graphs with community structure and arbitrary degrees, arXiv:1310.0046.

Zingg, B., Hintiryan, H., Gou, L., Song, M., Bay, M., Bienkowski, M., Foster, N., Yamashita, S., Bowman, I., Toga, A., Dong, H.-W., 2014. Neural networks of the mouse neocortex. Cell 156 (5), 1096–1111.