

EÖTVÖS LORÁND UNIVERSITY
FACULTY OF SCIENCE

**SIGNAL FLOW HIERARCHY
IN A MULTI-SCALE DYNAMIC
NETWORK MODEL OF THE
CEREBRAL CORTEX**

BÁLINT VARGA

Thesis, Biophysics MSc.

Consultants:

László Négyessy, Ph.D.

Zoltán Somogyvári, Ph.D.

Dept. of Computational Sciences, Wigner Research Centre for Physics

Internal advisor:

Imre Derényi, Ph.D.

Dept. of Biological Physics, Eötvös Loránd University



Budapest, Hungary

2020

Abstract

The cerebral cortex is a network with a nested hierarchy. Neural communication gives rise to complex dynamics characterizing the signal flow in the cortical network. However, the relationship between structure and function is highly non-trivial. While physical connections do impose constraints on signal flow, the observable statistical interdependencies between the elements of the network show time-varying patterns. The understanding of this dynamics is of special interest since it would grant us the ability to diagnose the workings of the brain in health and disease. The goal of my studies was to implement and analyze an anatomically and physiologically realistic dynamic model of a large-scale network of cortical areas. To this end the functioning of the cortical areas were modelled by a multi-compartment neural mass model in a directed and weighted hierarchical network. Structure-function relationship was studied by correlating three measures: 1) convergence degree (CD), a topological measure of signal flow, 2) an anatomical index of cortical hierarchy and 3) dynamical dependence by computing spectral Granger causalities between areas. To obtain a biologically relevant CD I introduced a modified version of the shortest path structure of the weighted graph, which favours robustness in the expense of the winner-take-all approach. CD exposed a densely connected component of higher-order areas resembling the rich club of the network that was not seen in the anatomical hierarchy. Due to computational limitations structure-function relationship was studied in a sub-network of eight areas. Remarkably, CD significantly correlated with the empirical index of anatomical hierarchy as well as with the causal relationship of areal activities. This is the first study showing the close correspondence of network topology and dynamics in a weighted and hierarchical model of the large-scale cortical network.

Acknowledgements

First and foremost I would like to thank my advisors László Négyessy and Zoltán Somogyvári. Laci would patiently answer all my multifarious and not always entirely relevant questions at any time (middle of the night and weekends included), with an average response time that no other human could beat. His meticulous revisions were crucial in achieving this polished final form. Soma gave vital help on the implementation of the dynamical model and the causality analysis. This work would not exist without their constant support and encouragement.

I would also like to thank the periodic input from other members of our research team at the Wigner Research Centre, namely László Zalányi, Fülöp Bazsó, Zsigmond Benkő, Marcell Stippinger, Bálint File, Tamás Kiss and András Telcs. I also thank Péter Erdős of the Rényi Institute for the meetings we had; as a matter of fact the reanimation of the convergence degree happened in part because of his encouraging words. I would also mention my college and ‘brother-in-arms’ Kristóf Furuglyás, with whom we shared this trial of academic life.

I thank the original authors of the model, Dr. Jorge F. Mejias and Prof. Xiao-Jing Wang for answering my plea for help at a time when I got desperately stuck with my reverse engineering of their work. Although I did not directly use the code they provided, it helped nonetheless, verifying my implementations and also revealing a few crucial technical details that I could not glean from their publications.

Last but not least I thank my family: my brother who provided basic life support for me while I was immersed beyond recognition in heaps of code, and my parents and foster father who made it possible for me to get to this point in the first place. My main goal in life is to make you proud.

Contents

1	Introduction	1
1.1	Neurobiological foundations	1
1.1.1	Hierarchical organization of the cortex	3
1.1.2	Anatomical hierarchy in the cortex	4
1.1.3	Functional hierarchy in the cortex	7
1.2	Dynamic models of brain activity	8
1.3	Brain as a complex network	11
1.3.1	Graph theoretical foundations	11
1.3.2	Signal flow in complex networks	12
1.3.3	Convergence in complex networks	15
2	Goals	18
3	Materials and methods	19
3.1	A multi-scale dynamic model of the primate cerebral cortex	19
3.1.1	Intralaminar local circuit	20
3.1.2	Interlaminar local circuit	22
3.1.3	Interareal model	22
3.1.4	Large-scale model	25
3.2	Network analysis of signal flow	29
3.2.1	Convergence degree	29
3.2.2	Node-reduced convergence degree	30
3.2.3	Beyond shortest paths	31
4	Results	34
4.1	A modified shortest path structure	34
4.2	Analysis of the CD flow	38
4.3	Correlations of topology, anatomy and dynamics	42
5	Discussion	46
	References	49
	Appendices	53
A	Retrograde tract-tracing	53
B	Electrocorticography	53
C	Causality analysis	53

D	Technical notes on the implementation	54
E	Parameter values for the dynamic model	55
F	Adjacency matrices	56

1 Introduction

It is easy to get lost in the details when studying living systems. The diversity and complexity presented by life, as we know it, is fascinating and also daunting. This may be especially true for the brain, with all its fine, enigmatic inner structures and workings. Accordingly, we have been persistently picking and prodding at it for the last 250 years, accumulating knowledge that is vast and diverse. From a certain perspective it is quite regrettable to give up this hard-earned, detailed knowledge to an abstract model described by a few obscure mathematical equations (1).

As a matter of fact, the definition of *modelling* as the extraction of significant details in order to make sense of the babel of reality, would be quite an accurate summation of perception itself (2). A model which does not generalize would be simply redundant. To quote Larry Abbott, a leader figure of the field of theoretical neuroscience (who in turn refers to a short story from Borges, written in 1946, titled *On Exactitude in Science* (3)):

The term ‘realistic’ model is a sociological rather than a scientific term. The truly realistic model is as impossible and useless a concept as Borges’ “map of the empire that was of the same scale as the empire and that coincided with it point for point.” (4)

That being said, it is not at all trivial what components of the system should be assumed relevant and which ones to ‘condemn to irrelevancy’ in the quest for understanding. In the following sections of this chapter I aim to outline some important neurobiological findings, deemed relevant at least for the present approach. I proceed to describe the general class of neural dynamic models used later in my thesis, and conclude with a summary of the graph theoretical results I ventured to extend by the work presented here.

1.1 Neurobiological foundations

The primate brain is a gigantic network of intricately interconnected neural and non-neural cells. It contains on the order of 10^{11} neurons that can be classified into thousands of different categories. Yet the behaviour of the animal (humans included), in all its complexity, presumably depends less on the diversity of cells, than on their local and global organization into circuits with precise, but adaptable functions (5). In the cerebral cortex, where most of the high-level computations of the brain are taking place, this organization of circuits leads to the highest scale of *cortical areas*: macroscopic regions of neural tissue with more-or-less clearly defined boundaries (Figure 1), connected via bundles of axonal processes. Therefore cortical areas are thought of as structurally and functionally differentiated, unique large-scale units of the *cortical network* (6, 7).

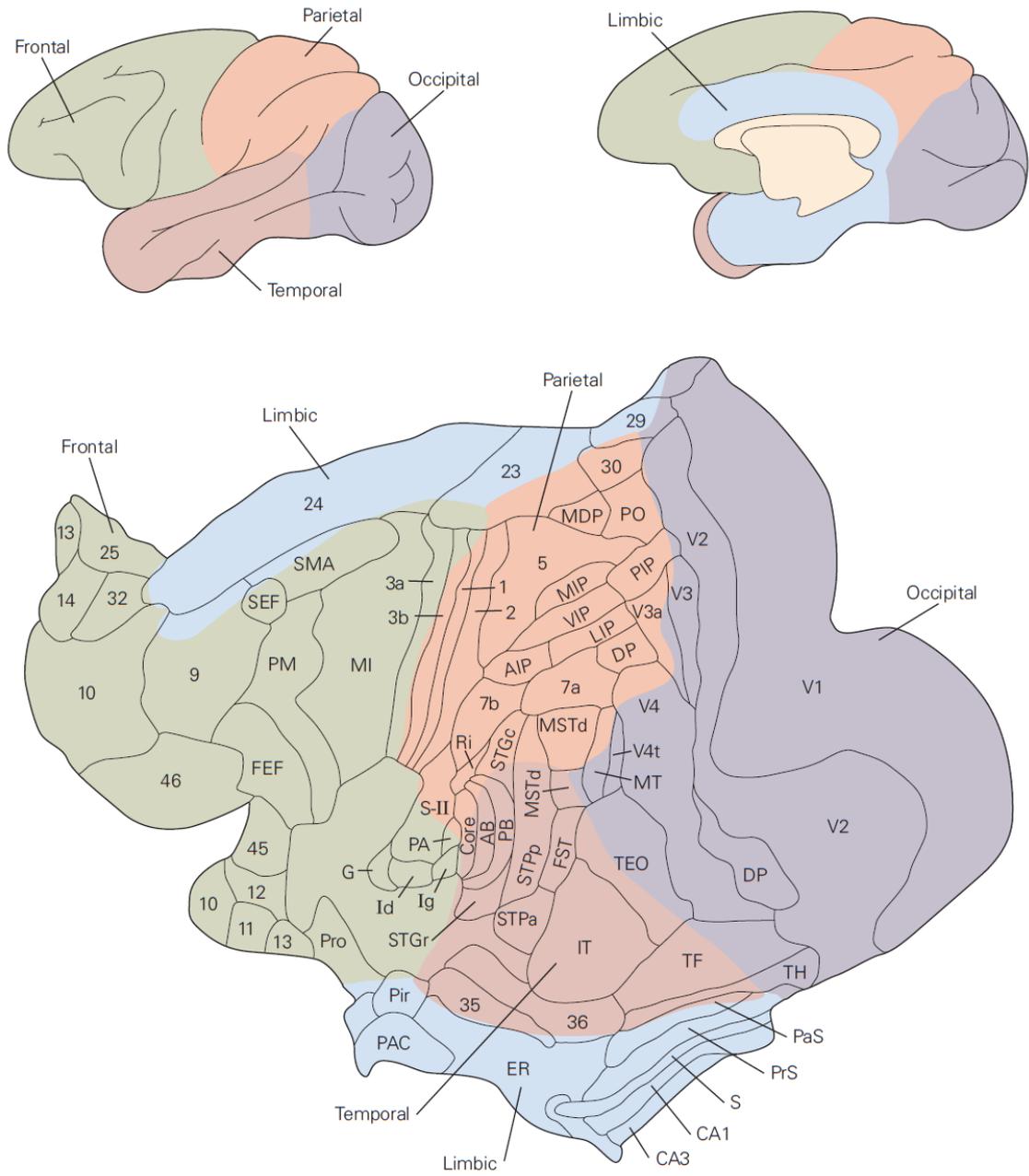


Figure 1: Schematic diagrams representing the cortical parcellation of the rhesus macaque, a primate that is a sufficiently close relative of ours (yet, unfortunately for them, also distant enough) to have received the unsought honour of becoming the number one test subject for primate brain research. Top: lateral (left), and medial (right) views of the brain, showing colour coding of the five cortical lobes which are named according to the corresponding cranial bones. Bottom: an unfolded, ‘spread-out’ view of the entire cortex from one hemisphere. The visual system, which is the main focus of this work, starting in the occipital cortex with the primary visual area V1, and has components spread across all lobes, for example the superior temporal polysensory (STP) areas in the temporal lobe, the frontal eye field (FEF) areas in the frontal lobe, subdivisions of the classical area 7 in the dorsal lobe and area 24 in the limbic lobe. From (8).

1.1.1 Hierarchical organization of the cortex

The relationships between cortical areas is said to be hierarchical, with *feedforward* (FF) communication directed from hierarchically lower-level areas to higher-level ones, and *feedback* (FB) communication going in the reverse direction (6). Here it should be noted, that as much as it is prevalent in the field of neuroscience and especially systems neuroscience, the term ‘hierarchy’ has its pitfalls (9). In particular, a vague definition leads to multiple assumed meanings which do not necessarily overlap and thus can be misleading. While I am going to use this term profusely in the following, I will aspire to state clearly the particular sense which it is meant to represent in the given context.

In the classical sense, cortical hierarchy assumes the successive processing of information in the brain, from input to output, also known as the *sensory-motor arc* (10, 11). Primary sensory areas receiving input from the sensory organs (e.g. retina), in most cases through the thalamus, stand on the bottom of this hierarchy (e.g. primary visual area, V1). Here, neurons are sensitive to simple features of the stimulus, for example so-called simple and complex cells to edges of different orientations in the V1 (12). Primary sensory areas are followed by higher-order, unimodal association areas, where neurons encode increasingly complex features of the sensory information. Regions that integrate information from multiple sensory sources (multimodal association areas) assume the ‘top’ positions in this sensory hierarchy.

In the multimodal areas the collated information of different sensory modalities gets coordinated with plans for action by interacting with motor cortices. In higher-order motor areas programs for movement are calculated and conveyed to primary motor areas for implementation, which in turn directly or indirectly (through subcortical formations) influence the efferent motor neurons of the spinal column, responsible for muscle movements (8, 13). The main principles of cortical organization in this sense are *integration* and *segregation*, ensuring the integrated functioning of distinct specialized structures (14, 15).

Between every level of the above described sensory hierarchy FF connections are thought to convey the processed information of the lower area, while FB communication reaches back from the higher area to modulate the FF stream (6, 7, 16, 17). In particular, feedback connections play a key role in higher cognitive functions including attention control (selecting features of interest based on behavioural context) and prediction (conveying our implicit expectations about the world, based on inherent and learned rules). The integration of said expectations and the processed sensory stimuli leads to inferences about the outside world (2, 16). One presumed interesting and arguably unsolicited corollary of this pattern is our ability for illusions (or inability thereof to perceive them correctly).

1.1.2 Anatomical hierarchy in the cortex

One case of a more specified meaning for hierarchy is generally denoted as *anatomical hierarchy* (9, 17). Thus far I implicitly wrote about the cortex as if it were a two dimensional sheet, ‘horizontally’ divided into separate, interconnected areas; of course, the cortex does have a third, ‘vertical’ dimension. Based on cytoarchitectonical (i.e. the characterization of cellular morphologies) and physiological observations, in its most developed form six cortical layers can be distinguished, each with a characteristic distribution of specific classes of excitatory and inhibitory neurons. Most prominent of these layers is L4, also called the granular layer, which receives most of the subcortical input, thus dividing the six layers into two compartments, a supragranular (L2/3) and an infragranular (L5/6) (13).

Classical and recent neuroanatomical studies show that several neural features change systematically across the cortex (18), along spatial axes highly similar to that of classical hierarchical succession. These features can be thought of as markers of differentiation, involving multiple levels of organization, such as genetic, cytoarchitectonical and connectivity based. A core observation is that all these features tend to change in a concerted manner along the *differentiation axis* (Figure 2). One of the main examples is the prominence of the granular layer, increasing as we move from less differentiated towards more differentiated regions. For example, V1 (which is a highly differentiated region) has perhaps the most pronounced granular layer, while it is less prominent in classical higher-order frontal and limbic areas (which show a less differentiated architectonic profile). Another feature is overall neuron density, which also increases as we move from less to more differentiated areas. Overall cellular soma size decreases along the axis for projection neurons (i.e. exclusively excitatory cells which project to other areas, in contrast to the ones only forming synapses locally, with other neurons of the same area). The proportion of supra- versus infragranular soma sizes tend to increase in the same direction. Connections exist predominantly between areas of similar standing, and neurons in less differentiated areas tend to form more connections, constituting a densely connected core of higher-order areas, also called a rich-club (19).

Moreover, laminar patterns of connectivity also show a similar tendency. Anatomical projections between areas close to each other (in the differentiation sense) do not discriminate between layers. However, areas of distinct differentional levels show a skewed unilaminar pattern based on their hierarchical relationship, in correlation with classical feedforward and feedback directions (9). Specifically, projections from more to less differentiated areas (corresponding to the FF direction) tend to target supragranular layers (relayed through L4), and they originate preferentially also in supragranular layers; the wider the differentional gap between the source and target areas, the stronger this preference is. Projections in the reverse (FB) direction on the other hand avoid L4, targeting

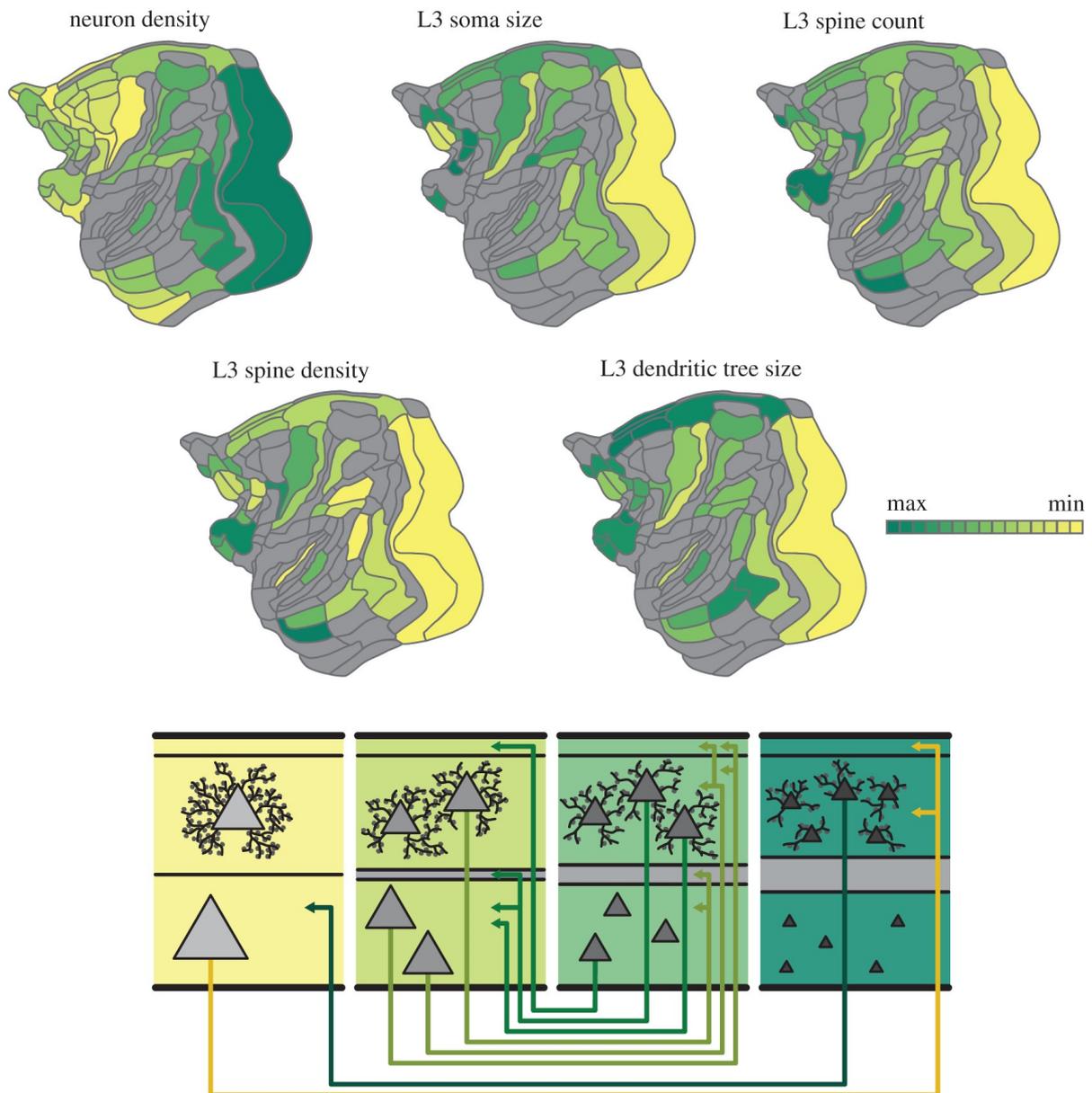


Figure 2: Diagrams showing the distribution of architectonic types across the cortex, especially the visual system. Top: colour maps for five architectonic features. The first panel shows the overall density of neurons in the cortical layers; area V1 has the highest cellular density. The other four panels focus on supragranular, L3 neurons; all four features have a minimal value in V1. Bottom: the concerted variation of the architectural features along the differentiation axis, from less to more differentiated, left to right (e.g. V1 would look like the far right block). L4 prominence (gray band) and neuron density increases, while soma and dendritic tree size decreases along the axis. Connections also show a distinct pattern. From (18).

both supra- and infragranular layers, while they originate preferentially in L5/6, where this preference shows the same correlation with differentiatonal distance as for the forward case (6, 7, 18).

The layer-specificity described here provides an opportunity for quantifying the anatomical hierarchy of the cortical network. Results harnessing this opportunity are mostly based on the seminal paper of Felleman and Van Essen (FVE), published in 1991 (6), which estimated the anatomical hierarchy by collating data from multiple sources. The most recent great upgrade to this database was brought by Markov et al. (7), who focused on the visual system employing modern techniques, and found correlating, but not identical results to that of FVE. In particular, based on retrograde tract-tracing (see Appendix A for a concise description) to label the input neurons in 29 cortical areas they defined two indices for the characterization of interareal projections: a relative connection strength (fraction of all labeled neurons, FLN) and a measure of hierarchical distance (proportion of supragranular labeled neurons within an area, SLN; Figure 3).

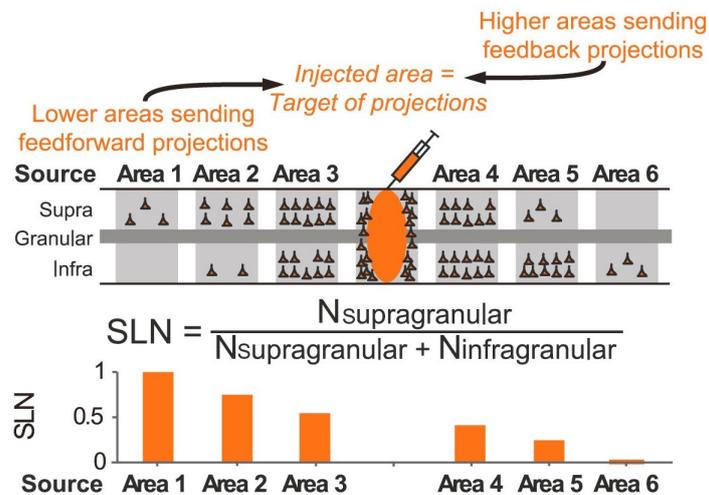


Figure 3: A schematic showing a tracer injection and the distribution of the identified projecting neurons. Supra- and infragranular neurons are both present in FF and FB pathways, but their relative contributions differ characteristically. More differentiated, ‘lower’ areas form FF projections originating in their L2/3, while less differentiated, ‘higher’ areas have FB projections originating in their L5/6. Therefore SLN estimates how FF a connection is. Areas reached from an injection can be then arranged by decreasing SLN, thus representing their anatomical hierarchical order. From (17), based on (7).

The SLN index of an interareal connection can take any value between the two extremes of zero (all the projecting neurons are from L5/6, the connection is totally feedback) and one (all projections originate from L2/3, the connection is absolutely feedforward). For values in between, the projecting neurons show FF and FB characteristics in the given proportions. The FLN also takes values from the closed interval $[0, 1]$, where a connection with an FLN value of one would mean that the target area receives all its input through that pathway (i.e. from that single source area), while a zero FLN would mean no connection at all (see Figure 30 and Figure 31 in Appendix F for matrix representations of FLN and SLN data, respectively, for the network of 29 areas). As a simplistic illustration

for a three-area network, if area A gives three projections to C, involving two supra- and one infragranular neurons, while area B has a total of two projections to C, one from the supra- and one from the infragranular layer, then $SLN(A,C)=2/3$ and $FLN(A,C)=3/5$, while $SLN(B,C)=1/2$ and $FLN(B,C)=2/5$; note that the FLN values for target C add up to one.

SLN is therefore assumed to be a measure of the anatomical hierarchical distance between two areas. To calculate the global anatomical hierarchical ranks from it, a generalized linear model was fitted to the SLN values (specifically, a beta-binomial model fitted by maximum likelihood, using the Nelder-Mead algorithm; see (7) for the exact procedural description). The result is a hierarchical ladder of the areas, from 1 to 10, owing to the traditional notion that there are up to 10 levels of hierarchy in the visual system (6).

1.1.3 Functional hierarchy in the cortex

Recent electrophysiological studies of the visual cortex, in particular electrocorticography (ECoG, see Appendix B for a brief summary) measurements documented also a dynamical layer-specificity of synchronization in the firing rate of neurons. L2/3 neurons show local and interareal synchronization in their firing rate mostly at higher frequencies, in particular in the gamma-band (60-80 Hz in (17) and 40-90 Hz according to (20)). L5/6 neurons on the other hand synchronize more at lower frequencies, specifically in the lower beta- (14-18 Hz in (17)) or alpha-band (5-15 Hz in (20)). A further synchronization was observed by (17) in the very low, theta-band frequencies (around 4 Hz), which was not documented by (20).

By applying spectral multivariate Granger causality analysis (see Appendix C for details on the computation) to ECoG time series data from eight visual areas, a consistent frequency-dependent directed influence was identified in the gamma- and theta-band for projections that are FF (based on their SLN value). Conversely, a directed influence in the alpha/low-beta band was found for anatomical connections that are anatomically FB. Based on these observations, Bastos et al. (17) defined an index for *functional hierarchy*, called directed influence asymmetry index (DAI), as the normalized difference between the Granger causality (GC) spectra in the frequency-domain, that is

$$DAI_{s \rightarrow t}(f) = \frac{GC_{s \rightarrow t}(f) - GC_{t \rightarrow s}(f)}{GC_{s \rightarrow t}(f) + GC_{t \rightarrow s}(f)} \quad (1)$$

where the indices s and t signify the source and target area, respectively. Note that $DAI_{s \rightarrow t}(f) = -DAI_{t \rightarrow s}(f)$. Defined this way, DAI values show positive correlation with the SLN in the gamma- and theta-bands, and negative correlation in the lower beta-band (Figure 4, left).

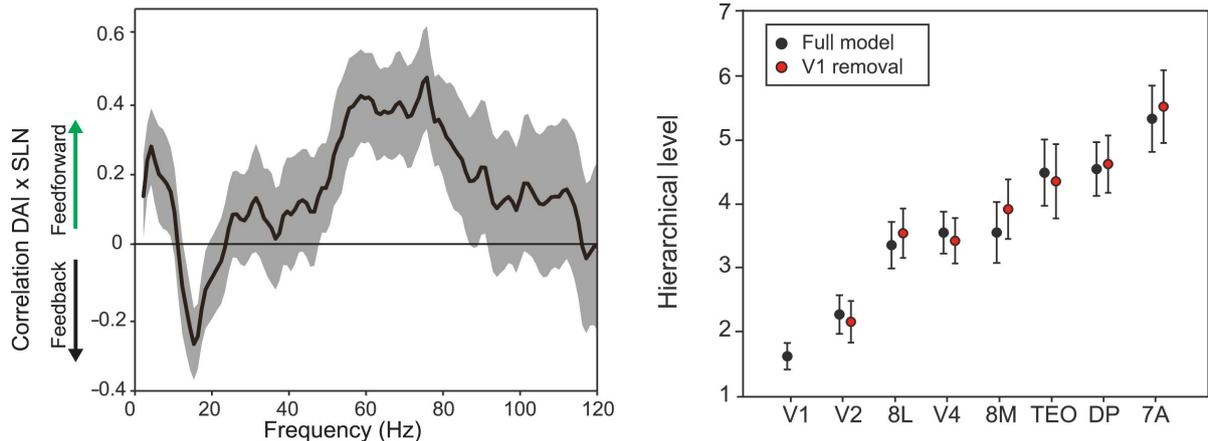


Figure 4: Left: rank correlation between DAI and SLN values across area pairs combined from eight areas, as a function of frequency. A negative correlation can be seen in the low-beta-band, and a positive correlation for higher frequencies. A further positive correlation can also be seen for very low, theta-band frequencies, which was not reported by (20). Right: functional hierarchy among the eight areas. To test for the robustness of this ranking, they computed it also without area V1 and found similar results. Adapted from (17).

They went further to define the multifrequency-band DAI (mDAI) by averaging the DAI of the theta, beta and gamma ranges (after inverting the beta values because of their negative correlation with the SLN); the mDAI strongly correlates with the SLN across all pairs of areas studied. The functional hierarchical ranks are then computed by a relatively simple series of transformations of the mDAI values (17). First, the mDAI was rescaled from the closed interval $[-1, 1]$ to $[-5, 5]$, owing to the 10 levels notion (6). Second, they considered each area in turn as a target, and shifted all source mDAI values, such that the lowest was one. This was done to achieve a positive scale, starting with one. Third and last, they averaged the resulting levels across all target areas. The result is a hierarchical ladder from 1 to 10 (Figure 4, right), similar to that of the anatomical hierarchy by Markov et al. (7).

One key difference between the anatomical and functional hierarchy is that the latter shows task-dependent variability (17). This instability is a general property of functional networks of the brain (in contrast with the relative stability of large-scale structural connectivity), and is observable on multiple temporal scales (21).

1.2 Dynamic models of brain activity

Yet another sense in which the brain can be described as hierarchical, is its hierarchically modular, embedded composition. This comes especially to light when we try to construct a model of the brain in a general and parsimonious way. Figuratively speaking, one could set the aperture to the resolution of individual neurons. Alternatively, one could go even

lower in spatial resolution and model the biophysical and biochemical interactions between molecules making up the cells. Choosing the elementary components of our model sets explicit limitations on its scope by determining the phenomena it might be able to reproduce. If one would venture to construct a global model of the entire primate brain on the cellular level (even only a static one, never mind dynamics), one would quickly start to appreciate both the sheer size of the brain and the relatively slim capacity of even our most powerful computers. Luckily, as I already implicated before, one might not have to go to such lengths to advance our understanding of the cortical network (1, 21–23).

Just as classical statistical mechanics can account for most of the large-scale phenomena involving matter without the need to evoke quantum mechanical principles, we can model collective neural phenomena with mean field models without explicit assumptions about the cellular workings of a single neuron (1, 22, 23). Moreover, perception, movement and overall behaviour are the products of collective neural activity, where the presence or absence of a single spike of a neuron does not lead to an observable difference. It is also the reality of the neuroscientific field, that a considerable amount of available functional experimental data represent the mean activity of large neuronal populations (e.g. fMRI, EEG, ECoG); fitting a model to this data requires considerations of the same level of function (1). Naturally, while evaluating their objective usefulness, we have to keep in mind at all times also the inherent limitations of such large-scale models (e.g. if we are interested in developing a pharmaceutical substance, we should probably consider quantum mechanical models of molecular interactions instead).

The activity of embedded neuronal populations inevitably occurs in the presence of noise. Stochastic fluctuations due to thermal energy, the inherent stochastic nature of ion channels and the cross-talk from nearby neurons all add up to form a noisy environment, which is also subject to external perturbations (experimentalists know the real struggle against ‘electrosmog’) (1). To address this circumstance, mathematical models of realistic large-scale neural dynamics employ stochastic differential equations, with two general components: a deterministic coupling term (i.e. drift, taken from classical stochastic calculus) and a noise term (also called diffusion) (22, 24).

There are several different realizations of this general scheme, perhaps the most basic being that of the neural ensembles. This approach is based on the assumption that if we consider large enough scales, the firing rates of individual neurons are uncorrelated, therefore according to the central limit theorem, the ensemble activity can be represented by a normal distribution, specifically with its mean and variance. This mean firing rate responds to the total synaptic input of the population by shifting towards higher or lower frequencies, while the variance represents the magnitude of all the noise effects (1). Alternatively, we can assume that the coherence between the firing rate of individual neurons is strong, in which case we can discard the variance and represent the ensemble

activity solely by its mean. With the degrees of freedom thus reduced to one, multiple different (e.g. inhibitory or excitatory; Figure 5, middle panels) neuronal populations can be modeled by a system of stochastic differential equations, each representing an ensemble mean; this is the main feature of the so called *neural mass models* (1, 22).

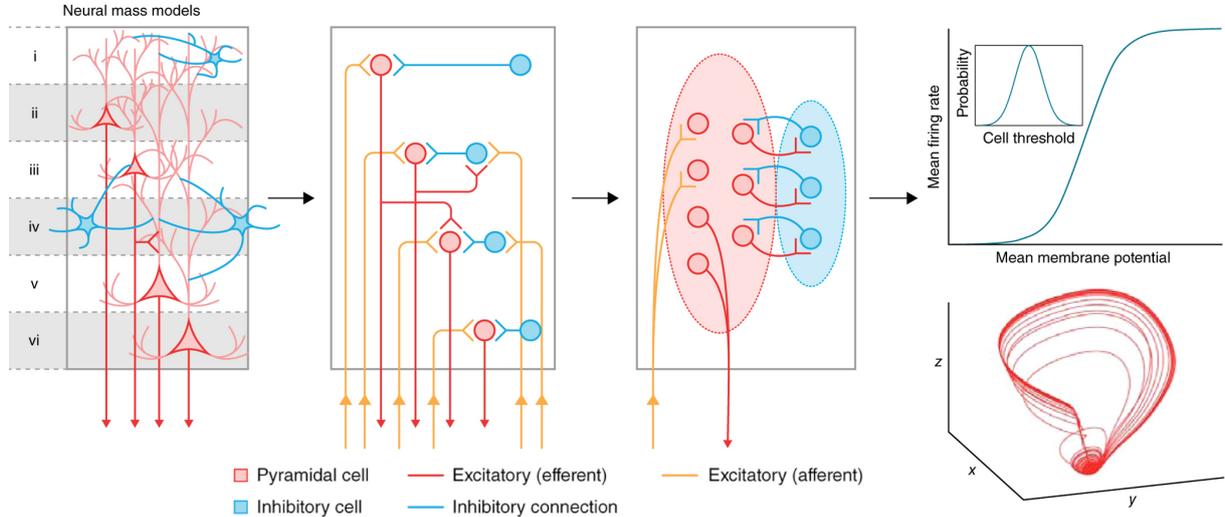


Figure 5: The general steps of abstractions taken when forming a neuronal mass model. All the micro-scale complexities of the neuronal tissue are reduced to a few populations of excitatory and inhibitory neurons, which can in turn be represented by a few coupled differential equations. The variability in the individual firing thresholds of the neurons can be represented as a sigmoidal relationship between the external input of a population and its mean activity. Integrating this system of equations results in attractor dynamics, for example in the form of a limit cycle, that is, stationary oscillations of activity. From (1).

One type of neural mass models can be defined by assuming that the coherence is so strong that the activity of the entire ensemble can be represented with the same biophysical equation that describes the activity of a single neuron. Alternatively a more empirical approach can be taken by informing the model by actual observations of biological neuronal populations. The model I used in my simulations is one such, in particular a variation on the classical Wilson–Cowan equations (23, 24). These equations employ a sigmoid (or a sigmoid-like) function for representing the probabilistic nature of the relationship between the inputs and the outputs of the population. The solutions of the system are time-series of the mean activity of the individual populations, which exhibit attractor dynamics. Based on the concrete parameter configuration different solutions can be achieved, usually the most desirable being a simple limit cycle oscillation.

Multiple neural mass models can be connected into an anatomy-based network, representing a patch of the cortex, or even a whole set of cortical areas, with dynamics unfolding on the underlying structure. I will describe in detail the equations and other components of the model I used in my work in the Materials and methods chapter.

1.3 Brain as a complex network

It is quite intuitive to think about the brain as a *complex network* (as far as *thinking about the brain* can be intuitive). In an implicit or explicit fashion I have adhered to this notion in the preceding sections too; here I lay out a more formal framework for capturing the brain in complex networks terms.

To reiterate, the elements of such a network can be chosen by different levels of abstraction. The appropriate domain (in spatial as well as in temporal terms) is selected based on the nature of the inquiry, thus identifying the relevant features for constructing the network (21, 25, 26). If we consider a potential spatial resolution ranging from molecular levels to the behaviour of the whole organism (and perhaps even further, to the societal networks of such organisms), and a corresponding temporal resolution from the milliseconds of molecular interactions to the neural development across the lifespan of the organism, then for this project I have set the aperture to ‘somewhere in the middle’ (see Figure 1 of (26)). In more concrete terms, I consider the stationary activity of large neuronal populations coupled by a constant anatomical connectivity in the time-span of a few minutes.

1.3.1 Graph theoretical foundations

After choosing the basic building blocks, their relations are then described and analyzed numerically as a mathematical *graph* (27). A graph $G(V, E)$ consists of a set of nodes or vertices (used interchangeably) V , and a set of connections or edges E . Generally, a graph is a topological object, which does not explicitly incorporate any spatial or temporal features. Adhering to this, it can be represented by a simple list of length n , consisting of the n vertex pairs for whom a connection exists (to be thorough, disconnected nodes might exist, or even entire disconnected components, but I will not consider this case in the following).

As a more visual alternative to the edge list representation, a binary adjacency matrix of size $n \times n$ can be constructed, where the cell ij is 1 if there is an edge between i and j , and zero otherwise. To somewhat enrich the topological parsimony, graphs can also be extended to a metric topology by associating real values with the edges; in this case the adjacency matrix holds the weights of the edges, with a zero weight meaning no connection whatsoever. If we do not consider ‘loops’, that is, edges that connect the vertex to itself, then the diagonal cells are always zero. Adjacency matrices are symmetric for undirected graphs, where the directionalities of individual edges are not taken into account, and are potentially asymmetric for directed graphs, where we differentiate between an edge going from i to j and an edge from j to i . A node-pair with edges in both directions between them is said to be reciprocally connected. It follows, that undirected graphs represent a

subtype of directed ones, with all the edges being reciprocal. Because direction is crucial in graphs that exhibit signal flow, in the following I will consider directed graphs (with and without explicit weights).

For any set of nodes several connectivity configurations are possible. To be precise, of the potential $n(n-1)/2$ undirected or $n(n-1)$ directed edges usually only a specific subset exists in a specific graph (with the all-to-all extreme case called a complete graph). A birds-eye view of the network configuration would be the proportion of the actual number of edges to the possible maximum number, referred to as the density of the graph. On the other hand, if we zoom in on a given node, the number of edges directly connected to the node represents its degree. For directed graphs the incoming and outgoing edges are differentiated, therefore an in-degree and an out-degree can be computed. One simple way to characterize a graph is to plot its degree distribution (27).

There is a multitude of measures for characterizing graphs, each describing different aspects and features (27). While some of them seek to describe features of vertices, others explicitly target edges. In terms of scope, some provide insight into the local structure, considering for example a single vertex and its direct neighbours, while others quantify general, global aspects of the entire graph. Of this menagerie of measures there are some select favourites for every subfield, based on the nature of the network under investigation, but none of the measures offer a single general answer for all questions. This might be exceptionally true for the intersection of graph theory and neuroscience, namely the young field of network neuroscience (26), with its short history of not more than twenty-few years.

1.3.2 Signal flow in complex networks

A structural network thus constructed can be furnished with dynamics, as I described in the previous section. On such an integrated system observable dynamical functional patterns can arise as a consequence of communication, or *signal flow* between parts of the system. These functional connections are measured and quantified in terms of statistical interdependence (Granger causality serving as an advanced example). Remarkably, the functional patterns unfolding on the underlying structure can chart out a whole different network of functional connections, which is determined by the structural skeleton, but not at all in a trivial manner. Signals spread in convoluted ways, resulting in effects that shape the functional connectivity (see Figure 6). Most importantly, the paths taken by the signal flow are not available for direct observation in most cases; this is true for many real-world networks, and especially true in case of the brain. The complex relationship thus outlined between structure and function is the main object of interest in my present work.

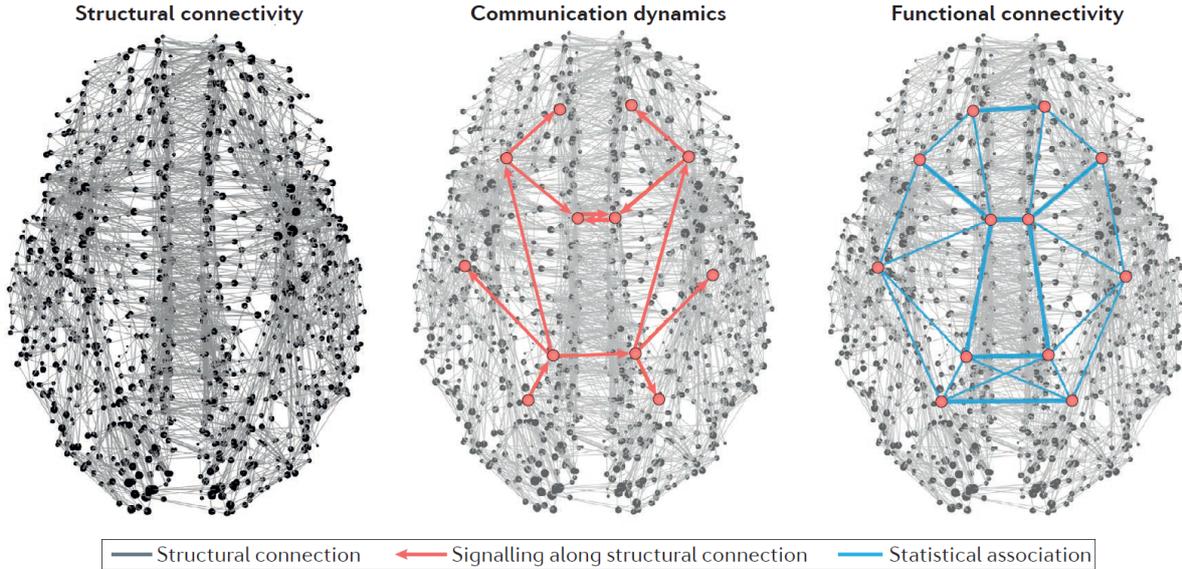


Figure 6: The relationship between structure and function, conceptualized as communication dynamics, or signal flow. A structural network model of the brain represents the physical connections between the nodes (left). These connections support and constrain the spreading of signals between nodes (middle), which is mostly hidden from direct measurement in the brain. The resulting effects in the activity of the nodes are observed as statistical dependencies, that is, functional connectivity (right). From (28).

One of the most basic graph theoretic concepts relating network topology and the communication that it supports is the idea of *shortest paths*, also called geodesic paths. As a matter of fact, it is the debuting concept of graph theory, described by Euler in the 18th century. In a binary graph, several paths can lead from one vertex to another, defined as the succession of vertices visited (or equivalently, edge crossed) on the way. The shortest of these paths (or rather, all the shortest paths, for more than one can be present with the same length) holds special significance: if we assume effective communication across the graph, it has to happen along these paths (at least in the absolute optimal case) (29).

Shortest paths can also be computed for weighted graphs by finding all the paths between vertex-pairs that minimize cost, which is usually defined as the multiplicative inverse of the weight (29). This definition should be employed by keeping the underlying meaning of the weights in mind: for example, the number of lanes on a road (perhaps serving as an analogy for the number of axons in a neural pathway) might be considered as a weight, while its inverse as a cost. On the other hand, the length of the road (or analogously the length of an axonal projection) should be thought of as cost, and ought to be inverted for the role of weight. It might also be noted that zero weights could pose a technical problem when trying to calculate the cost (i.e. zero division). One way to alleviate this problem is to use NaN instead of zero in the absence of a weighted

connection (which can be especially useful when a zero weight is assumed to have a whole different meaning; in this case the cost would be infinity). Then again, using NaNs in the adjacency matrix poses problems of its own; in short, one should always keep in mind non-connections when algorithmically calculating weighted graph properties.

Computing shortest paths by the classical cost-minimizing algorithm (called Dijkstra’s algorithm) can lead to some hard-to-justify results in certain cases. In particular, one problem is that it evaluates path length exclusively by minimizing the cost, without considering the potentially high number of jumps (i.e. nodes visited). In semi-metric networks, where the edges have associated distances, while nodes are also assumed to have some inner structure, or simply that the delay they impose on the transmission is non-negligible, the shortest paths we get with Dijkstra’s algorithm will not be accurate representations of the most efficient communication channels (28, 30, 31). Multiple (at least partial) solutions exist that address this problem, of which I will describe the ones that I employed in the Materials and methods chapter.

The distribution of shortest paths in the graph can be of special interest in multiple situations. Specifically, the sum of the fractions of shortest paths for all node-pairs that traverse a given edge is denoted as the *edge betweenness* centrality (EB). That is, the betweenness of an edge e is

$$EB(e) = \sum_{s,t \in V} \frac{\sigma(s,t|e)}{\sigma(s,t)} \quad (2)$$

where V is the set of nodes, $\sigma(s,t)$ is the number of shortest paths between nodes s and t , while $\sigma(s,t|e)$ is the number of those shortest paths between s and t that go through the edge e . In a simple case where only a single shortest path exists for every node-pair, the EB is simply the number of all shortest paths that pass through the given edge. Similarly to the EB, if we are considering nodes and their participation in the shortest path structure, we get to the node betweenness centrality (27, 30, 31). These indices quantify the importance of a given edge or node in the signal flow on the graph. If we consider a real-world signal-flow-supporting network with elements that are prone to failure, the elements with the highest centralities are the ‘weak-spots’ of the network. In particular, if we start eliminating nodes or edges by their centrality, starting from the least central, a lot of elimination will have to be done before considerable effects would show up in the signal flow. On the other hand, if we start the elimination with the most central elements, the signal flow would quickly collapse. How quickly depends on the shortest path structure of the graph: the more balanced their distribution among the elements, the more robust the signal flow on the graph (30).

Several topological indices are defined based on the shortest paths, and through them on the assumption that there is an optimization for communication efficiency in the network. This assumption is much debated for semi-metric real-world networks (28, 30), and

while efficiency of communication is thought to be a common feature in many physical, and especially in biological networks (with the brain as a fine example), it is not at all clear how much they are optimized for this end. While it is prevalent in the literature, assuming that brain network nodes communicate *explicitly* along shortest paths returns an oversimplified, and therefore unrealistic picture. On one hand, it is a fact that axonal projections not only serve the purpose of communication, they also hold an exact metabolic cost, leading to the observation that long projections are handed out sparingly in the brain (32). In particular, the metabolic wiring cost results in the distance rule that is described in a mathematical form by an exponential decay in the number of projections with increasing spatial length. In network scientific terms this imposes a constraint appearing in the form of edge costs, but as I mentioned above, the number of jumps should incur a cost too.

On the other hand, several observations have been made that suggest a more relaxed, diffuse manner of signal flow instead of communicating through strictly shortest paths. For example, if there is a bottleneck in the shortest path structure of a network (that is, multiple overlapping communication routes exist), that would lead to congestion and interference in the signal flow. This brings us to the debate over the question: is the brain routing its internal communication (i.e. signals flow through determined paths) or is it the case that signals spread entirely by diffusion along the underlying structure (28). In this work I assume a balanced position, relying heavily on shortest paths in my analysis, but acknowledging their limitations by some mitigating measures to lessen the rigour of the ‘absolute optimum’ assumption.

1.3.3 Convergence in complex networks

If one would try to summarize everything I wrote thus far in a single concise slogan, probably the best bet would be something like: *signal flow in the cerebral cortex is organized in a hierarchy of integration and segregation*. A reasonable objective would be then to quantify the integrative properties of the cortical network. A viable way to think about integration in network terms is to consider the convergence of connections and therefore signal flow; however, translating this idea to an exact definition is not a trivial task. A naive approach to the problem could be to try and define convergence based on the in- and out-degree of a node, but this would not lead to success for several reasons. The main obstacle is that the cortical network is too dense and too many edges are reciprocated for a local perspective like this one to give a sensible account of the global integrative properties of the graph (33).

Considering the previously outlined role of shortest paths, they might seem a worthy candidate for a starting point. In fact, introduced by Négyessy et al. (33) and developed into a rigorous mathematical framework by Bányai et al. (34), the *convergence degree*

(CD) seeks to characterize the integrative properties of a graph based on its shortest path structure. It is fundamentally an edge property, quantifying the convergent or divergent nature of individual directed edges. This is a desirable quality, in contrast to measures defined explicitly for nodes; a simple illustration for this being that since usually there are many more edges than nodes in a graph, greater robustness of the results can be achieved by characterizing the edges (33).

As far as shortest paths represent communication channels between node pairs, the convergence degree of an edge can be intuitively understood as the normalized difference between the number of nodes ‘talking’ through channels that comprise the given edge, and the number of nodes ‘listening’ to those channels. It is calculated as

$$CD(i, j) = \frac{|In(i, j)| - |Out(i, j)|}{|In(i, j) \cup Out(i, j)|} \quad (3)$$

where $|In(i, j)|$ and $|Out(i, j)|$ are the number of starting and ending nodes for shortest paths that pass through an edge (i, j) , and the denominator is a normalization factor, as the number of nodes in the union of these two sets (see the Materials and methods for a precise derivation). CD takes values from the open interval $(-1, 1)$. If $|In(i, j)|$ and $|Out(i, j)|$ is of the same size, that is, the same number of nodes are ‘talking’ as there are nodes ‘listening’, the CD value of the given edge (i, j) is zero. If more nodes are ‘talking’, the CD is positive, signifying a convergent connection, and if there are more of the ‘listening’ nodes, the resulting negative CD means a divergent connection.

Convergence and divergence are largely complementary, as indicated by the observation that the CD values of reciprocated edges have mostly inverse signs (33). Thus, based on the CD value of edges two complementary subnetworks can be identified, sharing the same nodes. Remarkably, the divergent and convergent subnetwork was found to correspond well to the feedback and feedforward subnetworks, respectively, identified in the cortical hierarchy (33). Note that this was done on the collated FVE database (6), where the SLN was not yet explicitly defined.

The edge-based formulation of the CD has undeniable advantages, but it is useful to have the ability to categorize individual nodes too. A simple condensed node measure based on CD is to take the difference between the sum of CD values for the outgoing and the sum of CD values for the incoming edges of the given node. This measure is named the *CD flow* and it was found to correspond to the hierarchical classification of nodes (34). Based on this one possible interpretation of the convergence degree for cortical networks could be ‘signal flow hierarchy’.

To assume a more detailed view, the *node-reduced CD* was introduced, which is actually the decomposition of the CD flow into four components. It is calculated by summing for the convergent and divergent, incoming and outgoing CD values, for all combinations

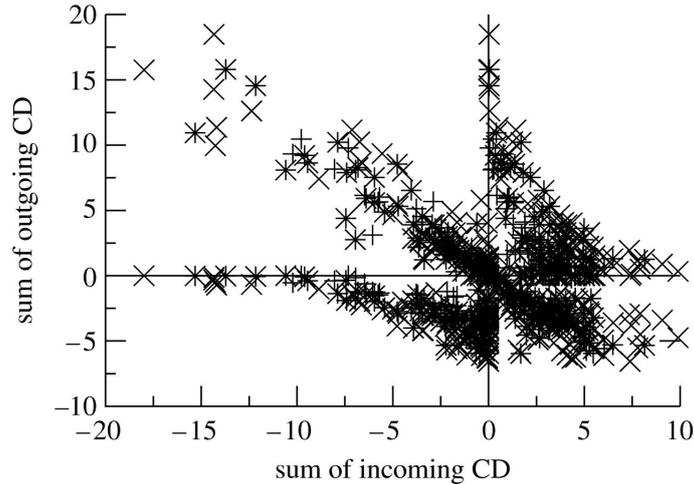


Figure 7: A node-reduced CD (nrCD) representation of the three different networks (macaque visuotactile cortex, whole macaque cortex and whole cat cortex) investigated in (33). Note that every area is present in all four quadrants (see the text for details). The distance of node from the origin in the respective quadrant signifies its sink (upper left), source (lower right) or relay (lower left and upper right) properties. The sink and source properties of nodes align clearly to show a negative correlation, while the relay properties form hyperbolic-like distributions, meaning that there are no explicit relay nodes in the cortical network. This can be understood as a topological proof that the cerebral cortex is hierarchical (34). Note that this version of the nrCD is not normalized. From (33).

separately (i.e. sum of positive in-CDs, sum of negative in-CDs, sum of positive out-CDs and sum of negative out-CDs). This results in four numbers for each node, two for the incoming and two for the outgoing direction, thus a plot with four separate quadrants can be drawn up, with a given node present in each of the quadrants (Figure 7).

Of these four quadrants the upper left $(-, +)$ and lower right $(+, -)$ holds special interest in brain networks, based on the fact that cortical areas mainly have afferents and efferents with CD values of opposing signs. Convergent input paired with divergent output (lower right quadrant) shows the source properties ('sourceness') of nodes, that is, the further a node is from the origin in this quadrant, the more it is estimated to behave as a source of signal flow. On the other hand, divergent input along with convergent output (upper left quadrant) indicates sink properties ('sinkness') of a the nodes. In these two quadrants a clear negative correlation can be seen, showing a gradual change between sink and source properties. The lower left and upper right quadrants represent signed relay characteristics, where both the in- and output is divergent (negative relay) or convergent (positive relay) (34).

2 Goals

In this work I aim to make a first step towards gainful contribution to the field of systems neuroscience. In particular, it is well documented that in terms of the cerebral cortex the relationship between structure and function is a highly non-trivial matter (28). On the timescale of behavioural phenomena (seconds to hours), the mostly stable large-scale anatomical network gives rise to fast-changing statistical interdependencies in the functional patterns (17, 21). The exact nature of this relationship is of special interest in applications ranging from pathological diagnosis to brain-computer interfaces.

The bridge between structure and function is the hierarchically organized communication among neuronal elements (28). One viable approach for characterizing the communication, i.e. signal flow properties of a network is the topological convergence degree, which is rooted in the concept of shortest paths (33, 34). Therefore, in a nutshell, my directive is to implement an empirically informed multi-scale dynamic network model of the cerebral cortex, and analyze its structure-function relations in a graph theoretic framework relying mostly on the convergence degree. My project has three main objectives:

1) The basis of my investigations is the multi-scale dynamic model described by Mejias et al. (23), which incorporates extensive anatomical and physiological data to reproduce many empirical observations of the cortical structure and function. I will implement and test this model framework to be able to analyse the signal flow in a controlled environment. By recreating the results of Mejias et al. (23) I also aim to contribute to the globally growing concern about the value of reproducibility in the sciences.

2) My main analytical tool is the convergence degree, described by Négyessy et al. (33) and Bányai et al. (34). While it was defined and thus far only applied to binary networks, to increase its scope I will also extend it to the domain of weighted graphs. To do this in an empirically informed manner I employ a combination of modifications to the classical shortest path-finding algorithm.

3) I will use the convergence degree to characterize the signal flow properties of the anatomical network that is the backbone of the dynamic model, both in a binary and in a weighted case. Finally, to better understand the structure-function relationship in the cerebral cortex, I will compare the obtained topological signal flow structure to the empirical anatomical structure and the simulated dynamical features of the model.

3 Materials and methods

In this chapter I will provide all the technical details of my work. In its raw form it is a several thousand lines of code (and an additional few thousand that was written and left unused or simply got jettisoned). I had to build up all my programs from scratch, writing many low-level routines for the simulation and the ensuing analyses, on several occasions reaching the limits in terms of computational power. See Appendix D for further details on my implementations.

3.1 A multi-scale dynamic model of the primate cerebral cortex

A significant amount of the work presented in my thesis encompassed the implementation of the empirically informed multi-scale dynamical model published by Mejias et al. (23) and the recreation of their results. In this section I describe this model in detail, with my recreations of most of the diagrams from the original article, and a few of my own addition. I also try to provide some further insight in specific cases, not explicitly stated in the original article or its supplementary materials, which can nonetheless hopefully clarify some aspects of the model. In the following, I mainly use the plural ‘we’, to denote the original authors, with my humble participation included, and refer to the singular ‘I’ only where I have to bear complete responsibility for the ensuing statements. It is also to be noted (as a disclaimer, perhaps), that this is not an exhaustive recollection of the results presented by Mejias et al.: some, which I deemed irrelevant for the present work, are with due respect omitted.

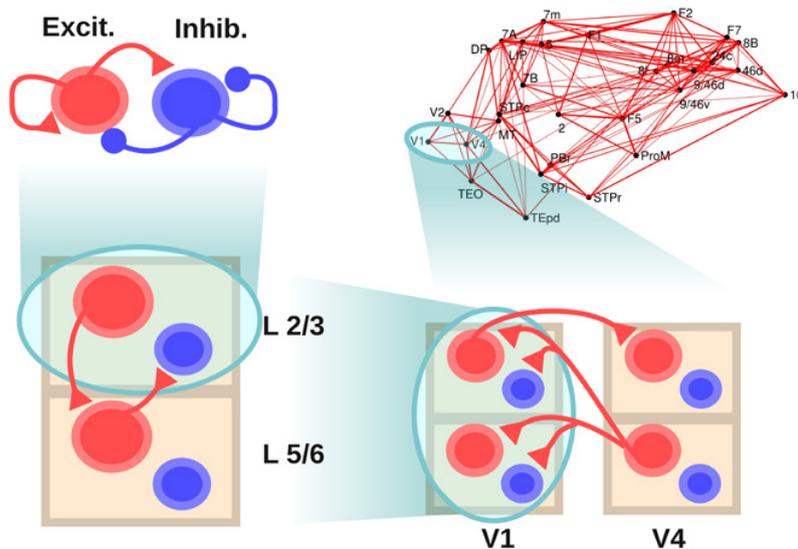


Figure 8: Layout of the multi-scale model (starting with the upper left panel, in a counterclockwise direction): intralaminar, interlaminar, interareal and large-scale level. Excitatory populations are in red, inhibitory in blue. The connections of the previous level are omitted for clarity. From (23).

The multi-scale model describes four embedded levels of neuronal structure: i) local excitatory and inhibitory populations in a given cortical layer, ii) four populations, two for the supragranular and two for the infragranular layer, representing a macroscopic cortical area, iii) two coupled areas, both represented by a laminar model, and iv) the large-scale cortical network of 29 areas, mainly in the visual pathways (Figure 8). In the following I describe in detail these four levels of modelling.

3.1.1 Intralaminar local circuit

The basic building block of the model is a pair of neuronal populations, excitatory and inhibitory, representing the respective pyramidal cells and interneurons in one (supra- or infragranular) layer of a given cortical region. Mathematically we employ a nonlinear Wilson–Cowan-type neuronal mass model, described by a pair of coupled stochastic differential equations, of the form

$$\tau_E \frac{dr_E}{dt} = -r_E + \Theta(I_E^{net} + I_E^{ext}) + \sqrt{\tau_E} \xi_E(t) \quad (4)$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \Theta(I_I^{net} + I_I^{ext}) + \sqrt{\tau_I} \xi_I(t) \quad (5)$$

where $r_{E,I}$ are dimensionless activities (i.e. mean firing rates) of the excitatory and inhibitory populations, respectively, $\tau_{E,I}$ are the corresponding time constants, $\xi_{E,I}$ are Gaussian noise terms with zero mean, and $\sigma_{E,I}$ strength and $\Theta(x) = x/(1 - \exp^{-x})$ is the transduction function (used instead of a sigmoid function), while $I_{E,I}^{net}$ and $I_{E,I}^{ext}$ represent the input from other parts of the network and from exterior sources (e.g from the thalamus, which is not included as an individual node), respectively. If we assume isolated intralaminar populations with only local contributions, then the network input is

$$I_E^{net} = J_{EE}r_E + J_{EI}r_I \quad (6)$$

$$I_I^{net} = J_{IE}r_E + J_{II}r_I \quad (7)$$

where $J_{\alpha\beta}$ is the coupling strength (mean synaptic strength), from population β to population α . The populations are connected both reciprocally and recurrently (Figure 8, upper left). Note that in the absence of all input (recurrent, network and external) to a population the equations still produce oscillations, but if we cut the noise term too, the mean activities die down exponentially. Two different parameter sets are employed in accordance with electrophysiological studies (17, 20), one for the supragranular and one for the infragranular populations: the former is chosen to achieve a noise-driven gamma (40 Hz) rhythm, while the latter shows alpha to lower beta (10-30 Hz) oscillations (Figure 9, left); see Appendix E for the exact parameter values used in this study. A case of elevated

activity can be simulated by a constant external input to the excitatory populations (e.g. the effect of contrast in V1; Figure 9, right).

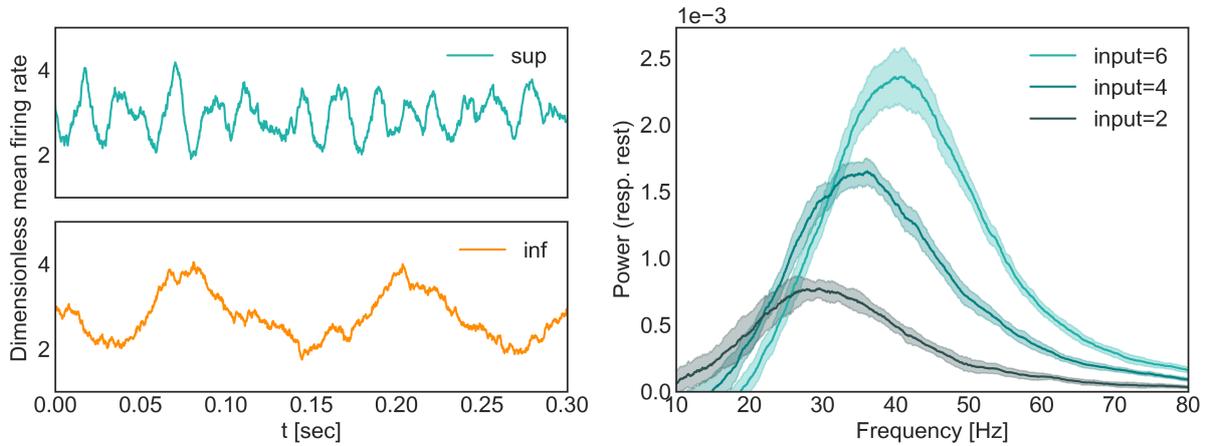


Figure 9: Noise-driven oscillatory activity of an excitatory-inhibitory circuit showing gamma-band oscillations in the supragranular (upper left) and alpha-band oscillation in the infragranular parameter sets (lower left). Right: power spectra displaying the effect of a constant external input to the excitatory population of the supragranular layer. The power spectrum of zero input has been subtracted from each curve. Recreated following (23).

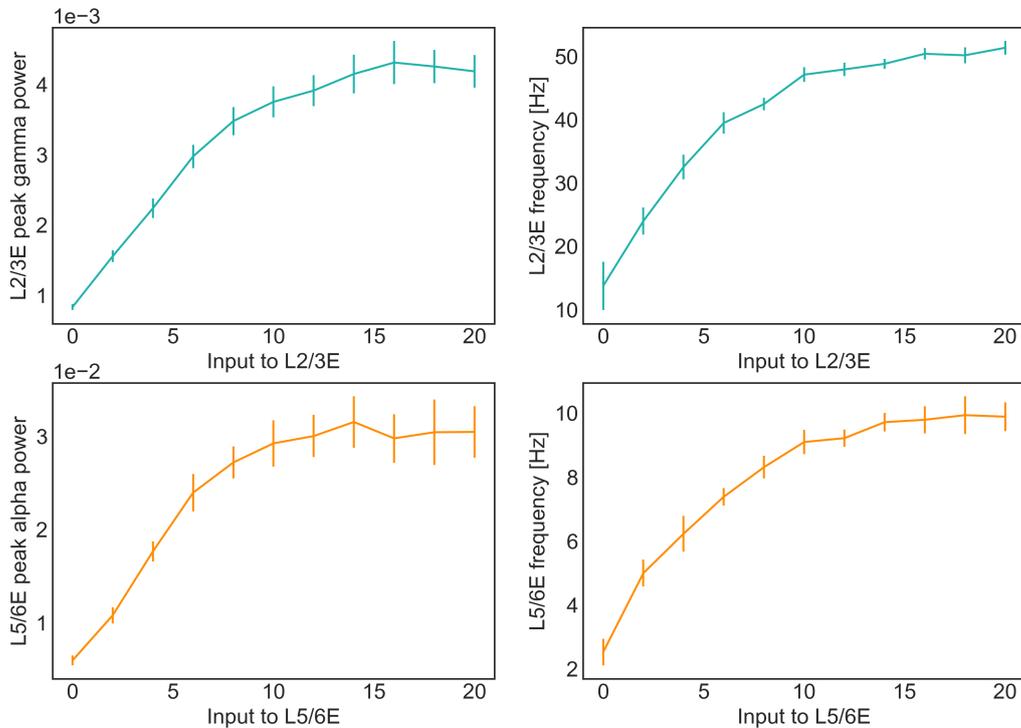


Figure 10: Effect of input to the excitatory population on the power spectrum peak (left) and corresponding frequency (right) for an isolated supra- (top) and an isolated infragranular layer (bottom). Recreated following (23).

The stimulation can be made in case of both isolated layers, and has two conspicuous effects: i) higher inputs lead to an overall increase in the magnitude of the characteristic power peak (Figure 10, left), and ii) a continuous shift of the peak towards higher frequencies (Figure 10, right). Both phenomena can also be observed in real-life physiological studies (see (23) for references).

3.1.2 Interlaminar local circuit

On the next level we couple the supra- and infragranular layers (L2 and L5 in the mathematical notation, for clarity) to form a local circuit, which we assume to represent a whole area. We have 2×2 populations, thus four different potential projections, of which we use only the two strongest, according to the anatomical studies (7): one from the supragranular excitatory (L2E) to the infragranular excitatory (L5E) population and one from the infragranular excitatory (L5E) to the supragranular inhibitory (L2I) population (Figure 8, lower left). We achieve this by extending equations 6 and 7. Using a convenient matrix notation, the input arriving to each of the four populations is

$$\begin{bmatrix} I_{L2E} \\ I_{L2I} \\ I_{L5E} \\ I_{L5I} \end{bmatrix} = \begin{bmatrix} J_{EE} & J_{EI} & 0 & 0 \\ J_{IE} & J_{II} & J_{2,5} & 0 \\ J_{5,2} & 0 & J_{EE} & J_{EI} \\ 0 & 0 & J_{IE} & J_{II} \end{bmatrix} \begin{bmatrix} r_{L2E} \\ r_{L2I} \\ r_{L5E} \\ r_{L5I} \end{bmatrix} \quad (8)$$

with $J_{2,5}$ and $J_{5,2}$ being the strength of the interlaminar projection from supra- to infragranular and infra- to supragranular layer, respectively. For a given area α we can rewrite this in a compact notation as $\mathbf{I}_\alpha = \mathbf{J}_\alpha \mathbf{r}_\alpha$. The result is what we would expect: via these interlaminar projections intralaminar rhythms can spread across layers and modulate each other's intrinsic oscillation (Figure 11).

3.1.3 Interareal model

We proceed to couple two of the previously described interlaminar circuits, simulating the interaction of two distinct cortical areas. Thus far, our model does not differentiate between areas: the connectivity pattern of the local circuits as well as their parametrization is assumed to be universal across the cortex. The layer-specific connection strengths between the individual areas however depend on their relative positions in the anatomical hierarchy (6, 7).

Here we choose the area pair of V1 and V4 which have a distinctive hierarchical relationship, so that the projection from V1 to V4 is exclusively feedforward, while from V4 to V1 it is solely feedback (note that in the large-scale case we are going to use a more general approach, where projections are mixtures of FF and FB communication, according

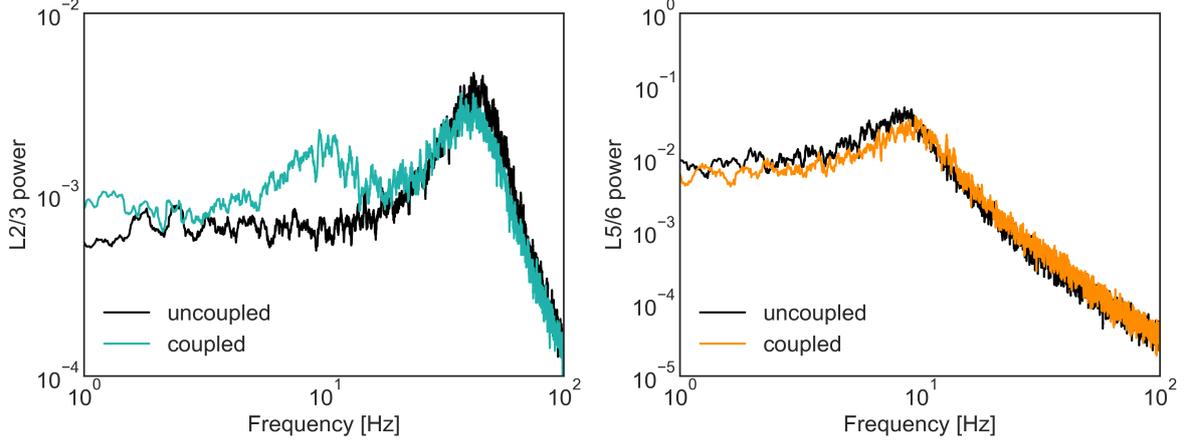


Figure 11: Power spectrum of the supra- (left) and infragranular (right) layer in the isolated and coupled case. A strong alpha component can be seen in L2/3, which is not present in the isolated circuit. In L5/6 a slight shift can be observed to higher frequencies. All axes are logarithmic. Recreated following (23).

to their SLN value). Anatomical studies (7) show that in the visual pathways feedforward projections originate mainly from L2/3 pyramidal neurons and target layer 4 excitatory neurons, which in turn establish synaptic connections with L2/3 pyramidal neurons; we approximate this with an interareal projection from L2/3E of the source area to L2/3E of the target area. Feedback communication in turn mainly originates in L5/6E but is not so focused, giving collaterals to all four populations of the supra- and infragranular layer (see Figure 8, lower right). We represent this with two distinct input equations for the two areas,

$$\mathbf{I}_{V4} = \begin{bmatrix} I_{V4L2E} \\ I_{V4L2I} \\ I_{V4L5E} \\ I_{V4L5I} \end{bmatrix} = \mathbf{J}_{V4}\mathbf{r}_{V4} + \begin{bmatrix} J_{FF1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_{V1L2E} \\ r_{V1L2I} \\ r_{V1L5E} \\ r_{V1L5I} \end{bmatrix} \quad (9)$$

for the feedforward direction, which can be written as $\mathbf{I}_{V4} = \mathbf{J}_{V4}\mathbf{r}_{V4} + \mathbf{J}_{FF}\mathbf{r}_{V1}$, and

$$\mathbf{I}_{V1} = \begin{bmatrix} I_{V1L2E} \\ I_{V1L2I} \\ I_{V1L5E} \\ I_{V1L5I} \end{bmatrix} = \mathbf{J}_{V1}\mathbf{r}_{V1} + \begin{bmatrix} 0 & 0 & J_{FB1} & 0 \\ 0 & 0 & J_{FB2} & 0 \\ 0 & 0 & J_{FB3} & 0 \\ 0 & 0 & J_{FB4} & 0 \end{bmatrix} \begin{bmatrix} r_{V4L2E} \\ r_{V4L2I} \\ r_{V4L5E} \\ r_{V4L5I} \end{bmatrix} \quad (10)$$

for the feedback stream, also compacted to $\mathbf{I}_{V1} = \mathbf{J}_{V1}\mathbf{r}_{V1} + \mathbf{J}_{FB}\mathbf{r}_{V4}$.

To test our model, we simulate a scenario, where both areas receive a constant input to their excitatory populations, and we analyze their frequency specific interactions. For

this purpose we calculate a weighted combination of the simulated time series data for both areas as

$$S_\alpha(t) = (1 - \eta)r_{\alpha L2E}(t) + \eta r_{\alpha L5E}(t) \quad (11)$$

representing the signal measured by an electrode recording in area α , with the parameter $\eta = 0.8$ reflecting the depth in the cortex. Analyzing the two signals we observe a frequency-specific relationship in each direction; to resolve the directionality of influences in the joint spectrum we employ spectral Granger causality (GC) analysis (see Appendix C for a more detailed description of the computation). This reveals the two distinct peaks for the FF and FB communication, in the gamma and alpha band, respectively (Figure 12), supporting electrophysiological observations (17, 20). To further express this causal relationship, we calculate the DAI (directed influence asymmetry index) as in Equation 1), adapted from Bastos et al. (17). With indices explicitly for the V1-V4 case, the equation looks like

$$DAI_{V1 \rightarrow V4}(f) = \frac{GC_{V1 \rightarrow V4}(f) - GC_{V4 \rightarrow V1}(f)}{GC_{V1 \rightarrow V4}(f) + GC_{V4 \rightarrow V1}(f)} \quad (12)$$

As expected, the DAI is negative for lower frequencies and positive for higher frequencies, signifying a clear functional hierarchy (V1 lower, V4 higher). Note, that if we were to find the inverted version of this curve (i.e. positive in lower, negative in higher frequencies), that would mean the inverse hierarchical relationship.

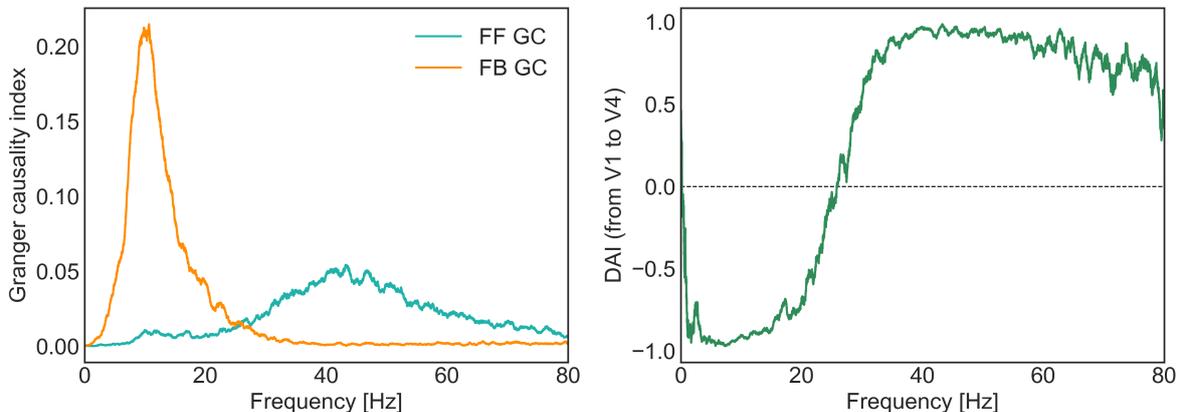


Figure 12: Frequency-specific interactions for the interareal circuit of V1 and V4. Left: to establish the causal relationship between the two signals we compute the Granger causality index in the frequency domain, and observe that each of the peaks correspond to a particular direction of influence. Right: the DAI is negative in the feedback frequency range and positive in the feedforward. Recreated following (23).

3.1.4 Large-scale model

To construct the large-scale network model, we use the anatomical connectivity data and indices described by Markov et al. (7). The graph has 29 vertices, with a density of 66%. We set a multilaminar circuit in each node, and define the edge weights as follows. (i) The FLN value of the corresponding anatomical projection serves as a basic strength for the edge, after a log-linear transformation, addressing the fact that FLN values span a range of about five orders of magnitude. Using $w_{ij} = c_1 FLN_{ij}^{c_2}$ with optimal parameter values for c_1 and c_2 , the connectivity information is preserved, while the actual strength values are compressed to a range that the rate model can handle effectively. (ii) The SLN value of the projection characterises the hierarchical relationship between the areas, and therefore the laminar specificity of the origin and target of the projections. As described before, an SLN value of unity would mean that the communication on the edge is totally FF, while a zero SLN signifies a pure FB direction; an SLN in between represents a specific proportion of the two, with a factor of SLN_{ij} for the FF and $(1 - SLN_{ij})$ for the FB flow. (iii) Lastly, while at the local levels the closeness of the populations did not warrant a delay for the transmissions, on the large-scale level we incorporate a mean empiric delay term calculated as the product of reported anatomical distances of the projections and an average action potential propagation speed of 1.5 m/s.

Putting together all these components, the net input received by neuronal populations in a given area i at time t is

$$\mathbf{I}_i(t) = \mathbf{J}_i \mathbf{r}_i(t) + \sum_{j \neq i}^{areas} [W_{FF}^{ij} \mathbf{J}_{FF} + W_{FB}^{ij} \mathbf{J}_{FB}] \mathbf{r}_j(t - \Delta_{ij}) \quad (13)$$

where $W_{FF}^{ij} = w_{ij} SLN_{ij}$ and $W_{FB}^{ij} = w_{ij} (1 - SLN_{ij})$ are the total weights for the FF and FB component, respectively (with w_{ij} as defined above), and Δ_{ij} is the interareal delay term.

From here on, while I generated the time series data for the whole 29 node network, for computational reasons, and also following the example of Mejias et al. (23), I mostly only analyzed a subnetwork of 8 regions of interest, selected by Bastos et al. (17). These 8 regions are selected roughly evenly from the visual hierarchy and are namely: V1, V2, V4, DP, 8m, 8l, TEO and 7A (see Figure 13 for matrix representations of the corresponding anatomical data).

In their article Mejias et al. investigated a parameter configuration with a uniform constant external ('thalamic') input to every area, plus a constant external input on top of this to area V1, representing a strong stimulation arriving from the retina. In addition to this setting, initially I simulated a 'resting-state' case too, without the extra retinal input. As expected based upon experimental findings of sensory-driven gamma

enhancements (see (23) for references), the stimulated case shows a drastically elevated peak for the gamma power of early visual areas V1 and V2, compared to the resting-state results (Figure 14).

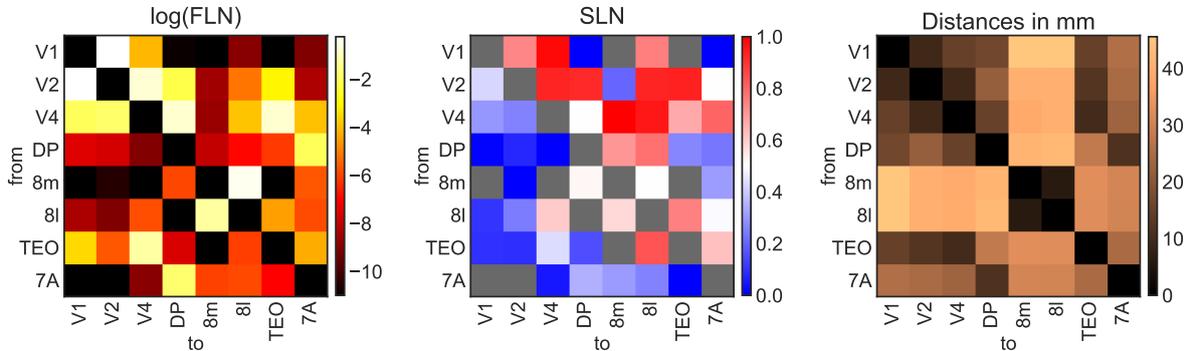


Figure 13: Adjacency matrices for the 8 regions of interest. The FLN and SLN matrices are highly asymmetric, while the distance matrix is symmetric. The areas on the axes are in anatomical hierarchical order. Note that because an SLN of zero still holds information, we have to differentiate it from nonexistent connections: gray cells in the SLN matrix represent NaN values, corresponding to zeros in the FLN matrix. See Figures 30, 31, 32 in Appendix F for adjacency matrices of the entire graph. Recreated following (23).

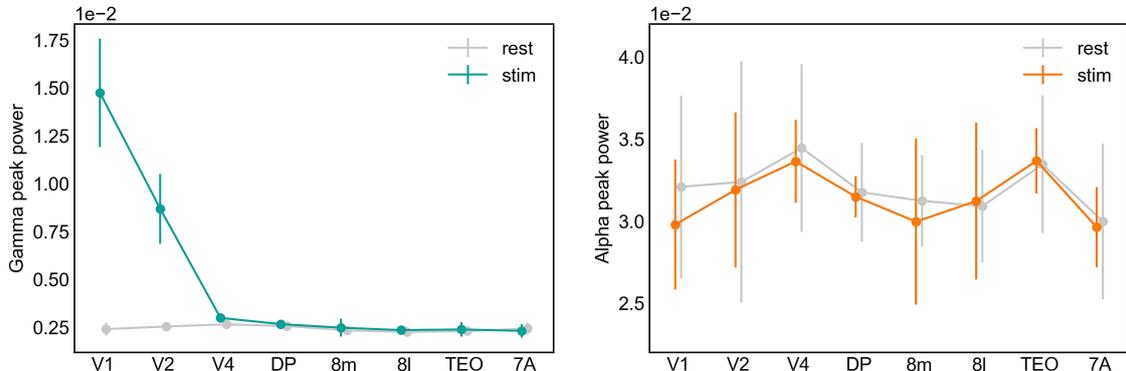


Figure 14: Power peaks for the 8 regions of interest in the gamma- and alpha-band. Retinal stimulus induces a gamma peak increase in V1, with V2 also indirectly effected, showing moderate increase to its strong coupling to V1. In the alpha band retinal input does not induce such a drastic effect, although a slight increase in V1 can be perhaps ascribed to weaker FB from V2. For both cases a slightly higher alpha peak in ventral areas V4 and TEO can be seen, but is yet unexplained. In the following I will not consider the resting-state, since later results show no significant deviations relative to the stimulated case. Partially recreated following (23).

We proceed to repeat the same procedure as for the interareal level (Figure 12), but for all area pairs from the 8×8 subnetwork. A multivariate spectral GC analysis (in contrast

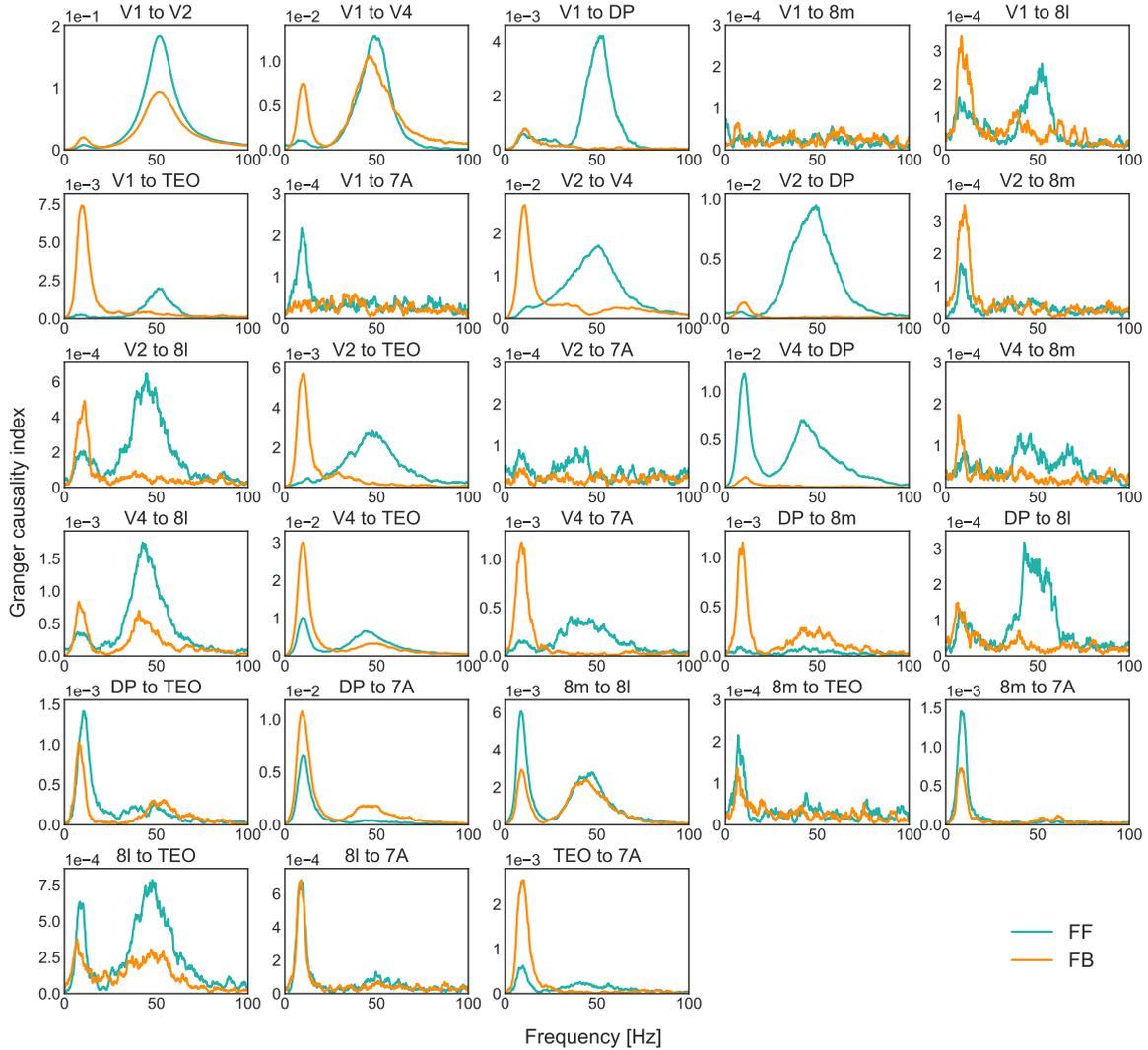


Figure 15: Conditional spectral GC profiles for the large-scale model, in particular for the 28 combinations present of the 8×8 subnetwork (to be precise, 2×28 directions, of the $64 - 8$ possible 2-tuples, omitting the diagonal). Note that these are conditional causality spectra, meaning that in theory only direct influences are considered (conditioned on all the indirect ones). Nonetheless, this is not directly constrained by the existing anatomical projections: for example from V1 to TEO no direct projection can be found (see the FLN matrix in Figure 13), yet there is a (weak but existent) functional influence between them. A constant ‘thalamic’ input of value 6 is added to every area, plus an extra 6 to area V1, simulating a state where the eyes are open. Note the different scaling of the vertical axes. Recreated following (23).

with the pairwise version used before this has the ability to filter the indirect influences; see Appendix C for more details) shows a complex picture of frequency dependent directed influences. The relationship between FF/FB and gamma/alpha rhythms is much more varied than for the pure FF/FB directionality of the interareal model. The overall influence is stronger for lower areas (due to the retinal input, as can be seen from comparisons with

the resting-state case), weaker but still discernible for areas further away from V1 (Figure 15).

To make the information present in the complexities of Figures 15 more accessible, we obtain the DAI spectrum by Equation 1 for all the area pairs. To check for the relationship with the anatomical structure, as did Bastos et al. (17), we compute the frequency dependent correlation between the DAI and the SLN, composing our version of their recordings, present here as Figure 4. As expected, we observe a negative correlation for the alpha-/low-beta-band, and a positive one for the gamma-band (Figure 16, left). In my results a distinguishable positive correlation can also be seen in the narrow theta-band (as reported by (17)), which Mejias et al. (23) do not report upon, therefore I also chose to omit it from my analyses; due to its proportionally small contribution, it would not significantly alter our results on this scale.

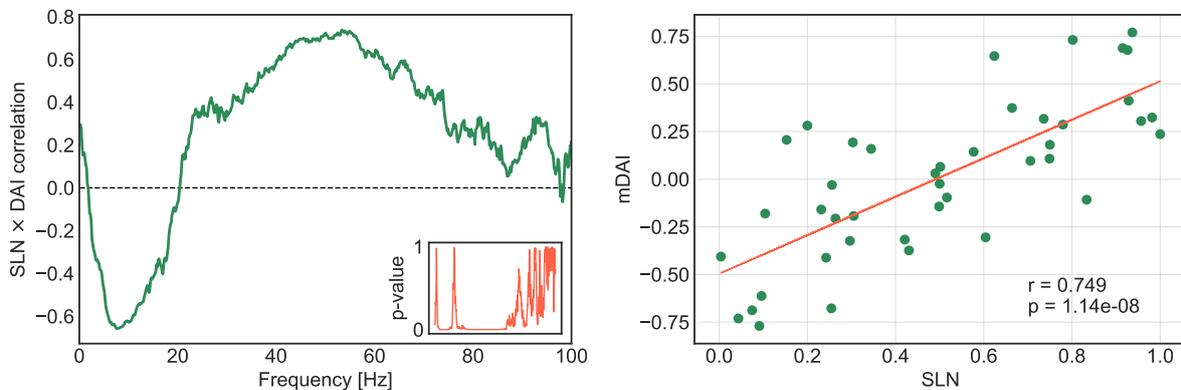


Figure 16: Relationship between anatomical projections and functional interactions. Left: correlation between the corresponding SLN and DAI values as a function of frequency (with the inset showing the p-values). Negative correlation can be seen in the alpha range, and positive in the gamma (and theta, see the text) range. Right: SLN shows a significant positive correlation with the mDAI. Note that following Mejias et al. the zero and one valued SLNs are not included in the correlation plot due to their erratic behaviour (which is not explained in (23), but is probably due to measurement errors). Partially recreated following (23).

On our way to functional hierarchy, we continue to follow Bastos et al., and combine the alpha and gamma ranges of the DAI spectrum (excluding theta, as discussed) for each area to form the mDAI (multifrequency DAI), which serves as a point estimate for the regions of interest, as

$$mDAI_{i \rightarrow j} = \frac{DAI_{i \rightarrow j}(\gamma) - DAI_{j \rightarrow i}(\alpha)}{2} \quad (14)$$

where the DAI for a given frequency range ω is

$$DAI(\omega) = \int_{\omega_{min}}^{\omega_{max}} DAI(f)df \quad (15)$$

where the alpha/low-beta range is taken to be 6-18 Hz, while the gamma range is taken as 30-70 Hz. We correlate the thus obtained mDAI values with the SLN and see a significant positive correlation (Figure 16, right).

3.2 Network analysis of signal flow

To analyse the model thus implemented, I used the graph theoretic framework outlined in the Introduction. There I presented only a rough sketch of the mathematical basis for computing the CD, focusing more on its properties and the results it lead to. Here I will recount the exact set theoretical formulation, following Bányai et al. (34). Note that they actually described two versions, a local and a global one, for the CD and all its derivations; here I will present and use only the global CD. Furthermore, just as in the previous section of this chapter, the plural we should be understood as the original authors, including myself as a hitchhiker, while I use the singular case only where it is appropriate. After thus setting the foundation, I will introduce the measures I have taken to extend the convergence degree into the domain of weighted graphs. More on this front can be found in the Results chapter.

3.2.1 Convergence degree

For any arbitrary graph G a set $SP(G)$ can be defined as the set containing all the shortest paths of the graph. Each path is present in the form of an ordered list of the nodes it visits, beginning with the node where the path originates and ending with the one where it terminates, signifying the optimal communication route between these two points. For a given edge (i, j) (going from node i to node j), a subset of $SP(G)$ can be identified that pass through (i, j) ; we can denote this set as $SP(G, (i, j))$.

Based on $SP(G, (i, j))$ another two unique sets can be defined: a set containing all the nodes from which the paths in $SP(G, (i, j))$ originate, called $In(i, j)$, and the other one consisting of all the nodes where these paths terminate, called $Out(i, j)$. The nodes i and j are by definition elements of $In(i, j)$ and $Out(i, j)$, respectively. Note that we are using sets, therefore, while several shortest paths could be found in $SP(G, (i, j))$ with the same first and/or last node, no replicas are allowed in $In(i, j)$ and $Out(i, j)$. Now, Bányai et al. went on to define a third set, one that contains the nodes which are present in both of these sets, that is, their intersection: $Int(i, j) = In(i, j) \cap Out(i, j)$.

The $Int(i, j)$ nodes are part of circles, that is, circular paths going through (i, j) , and therefore should not be considered in the calculating of the CD. According to this we define the strict version of the in- and out-set as

$$SIn(i, j) = In(i, j) \setminus Int(i, j) \quad (16)$$

$$SOut(i, j) = Out(i, j) \setminus Int(i, j) \quad (17)$$

With the help of these sets the notion of causality (34) in global signal flow through the edge (i, j) can be understood as: i) $SIn(i, j)$ holds the nodes which are causes of the flow, ii) $SOut(i, j)$ holds the ones which suffer the effects of the flow, iii) the nodes in $Int(i, j)$ exhibit ‘circular causality’, while iv) all the remaining nodes in the graph that are not part of the $In(i, j) \cup Out(i, j)$ set, are not related in causal terms to the signal flow through the given edge. To continue, by representing the cardinality of a set with $|\cdot|$, for a given edge the relative size of its strict in- and out-sets are

$$RIn(i, j) = \frac{|In(i, j)|}{|In(i, j) \cup Out(i, j)|} \quad (18)$$

$$ROut(i, j) = \frac{|Out(i, j)|}{|In(i, j) \cup Out(i, j)|} \quad (19)$$

The CD is then defined as

$$CD(i, j) = RIn(i, j) - ROut(i, j) = \frac{|SIn(i, j)| - |SOut(i, j)|}{|In(i, j) \cup Out(i, j)|} \quad (20)$$

Note that this definition is equivalent with Equation 3, given in the Introduction, because $|SIn(i, j)| - |SOut(i, j)| = |In(i, j)| - |Out(i, j)|$, since the same overlap is present in both terms, and thus it drops out in the difference. To reiterate, a positive CD value indicates a signal flow through the given edge, that has more sources than sinks, while a negative CD indicates the opposite case, that is, the flow having more sinks than sources. Based on this property, it is reasonable to say that an edge with positive CD is convergent, an edge with zero CD is balanced, while one with negative CD is divergent. An illustration as to why the overlap in the in- and out-sets should be excluded from the calculation could be the extreme case of a graph consisting of a single oriented circle, where the overlap is maximal, and the absolute value of the CD is the smallest possible (34).

3.2.2 Node-reduced convergence degree

As I mentioned in the Introduction, while defining the CD as an edge-based measure is desirable, it is useful to have a node representation of the CD. For this end Bányai et al. (34) defined the node-reduced CD (nrCD) and the so called CD flow. The nrCD is in fact four different normalized sums for every node, calculated based on the four possible combinations arising from the negative and positive CD values of the incoming and outgoing edges. If we consider a $n \times n$ graph, denoting the left-continuous Heaviside

step function (which gives 1 for $x > 0$, and zero otherwise) as $H(x)$, the degree of a node as Γ , and otherwise using self-explanatory indices, the four node-reduced CD components of a node i are

$$nrCD_{in}^-(i) = \frac{1}{n-1} \sum_{j \in \Gamma_{in}(i)} H(-CD(j, i)) CD(j, i) \quad (21)$$

$$nrCD_{in}^+(i) = \frac{1}{n-1} \sum_{j \in \Gamma_{in}(i)} H(CD(j, i)) CD(j, i) \quad (22)$$

$$nrCD_{out}^-(i) = \frac{1}{n-1} \sum_{j \in \Gamma_{out}(i)} H(-CD(i, j)) CD(i, j) \quad (23)$$

$$nrCD_{out}^+(i) = \frac{1}{n-1} \sum_{j \in \Gamma_{out}(i)} H(CD(i, j)) CD(i, j) \quad (24)$$

The normalization factor $n - 1$ is simply the number of possible edges of the given node in the graph. This leads to the node-reduced diagram representation of a graph, exemplified by Figure 7 in the Introduction. As an interesting side note, according to Bányai et al., extending this plot by similarly calculated terms for the overlaps in the in- and out-sets (see the previous chapter) as the z-axis, we get a totally unique (although somewhat redundant) three-dimensional representation of the given graph.

The node-reduced CD equations can be thought of as a decomposition of the CD flow, which is then defined by Bányai et al. as the difference between the sum of the incoming CD values and the sum of outgoing CD values. The authors also discuss the ability of this index to characterize the hierarchical order of the nodes in a graph. During my work I found this hierarchical classification to be more consistent if I normalize the sums in the CD flow definition. This is important, since in a graph where the lack of edges is not distributed uniformly (for example in our case, V1 has significantly fewer connections than higher-order areas), the summing of in- and out-CDs gives a skewed representation. Thus, when I refer to the CD flow (Φ) in the following, I will implicitly use the normalized definition of the CD flow, which is

$$\Phi(i) = \frac{1}{d_{out}(i)} \sum_{j \in \Gamma_{out}(i)} CD(i, j) - \frac{1}{d_{in}(i)} \sum_{j \in \Gamma_{in}(i)} CD(j, i) \quad (25)$$

where d_{in} and d_{out} denote the in- and out-degree of the given node.

3.2.3 Beyond shortest paths

While the convergence degree is deeply rooted in the concept of shortest paths, one might notice that it holds no explicit assumption about the actual method for finding them. Nor

does it limit the number of shortest (or even near-shortest) paths present between two nodes; in fact, as I mentioned previously, there can be multiple paths of the minimum length. Conceptually, the CD simply takes the pairs of nodes assumed to communicate through an edge. It is also to be noted that in the work done around the CD until now, only binary graphs were considered. The same applies here as in the previous case: the CD algorithm does not hold any inherent limitations to an extension into the domain of weighted graphs, as far as we have a viable method for finding the weighted shortest paths. These observations present an opportunity to ‘make things more real’, that is, to apply the CD to a weighted graph and to relax the strict (and for the brain unrealistic) measure of considering only the absolute shortest paths. In the following I will describe the steps taken to achieve these goals.

As I mentioned in the Introduction, while shortest paths pervade the network scientific literature, they also pose some challenges in their application to real-world networks, especially semi-metric ones. One such problem arises when we apply the classical Dijkstra algorithm. In particular, it does not take into account the number of edges when calculating the total cost of a path, only the cost assigned to them. A relatively easy workaround to this problem is to transform the inverted weights simply by adding a tuning parameter α proposed by Opsahl et al. (31). The shortest path between nodes i and j is then

$$d(i, j) = \min \left(\frac{1}{w_{ik}^\alpha} + \dots + \frac{1}{w_{lj}^\alpha} \right) \quad (26)$$

where w_{xy} is the weight of the edge (x, y) , and the exponent α is the tuning parameter. If α is one, we get back the standard results; with the extreme value of zero, the binary shortest paths are computed. But if we take a value smaller than one, but not zero, a trade-off is being made between edge costs and number of jumps, which can be adjusted with respect to the node delay for the given graph.

Another problem of the strict definition of shortest paths is the fact, that most real-world networks do not optimize solely for communication efficiency, necessarily resulting in mathematically suboptimal solutions. It can also be said, that the classical shortest path-finding algorithms do not take into account all factors needed for real-world efficient communication; for example a very unbalanced edge betweenness centrality (with only a few superpopular edges) can lead to network congestion, or traffic jams as an analogy.

One way to somewhat spread out this ‘superfocused’ configuration could be to consider multiple short(est) paths instead. That is, use Yen’s algorithm (which in turn builds on Dijkstra’s), to find the k -shortest paths between two nodes. This is in line with a more relaxed and much more robust assumption about communication efficiency, that is, two nodes can communicate through several channels, and if one fails (e.g. either by congestion or by some physical malfunction) the others are still available (28). I will give the exact

tuning parameters I used for both α and k in the Results chapter, also explaining in detail the reasoning behind the specific choices.

4 Results

4.1 A modified shortest path structure

For my analysis I constructed a binary version of the 29×29 anatomical network based on the FLN matrix, essentially by applying a left-continuous unit step function (keeping the zeros and reducing all non-zero FLN values to unity). I also defined a cost for the weighted version of the 29×29 network, considering solely the length and strength of a projection. More precisely, I calculated the cost of an edge as the product of the projection distance in millimeters (the ‘length’ of the edge) and the multiplicative inverse of the FLN (the ‘width’ of the edge). Thus for a given connection its distance is directly proportional, while its FLN value is inversely proportional to the cost of traversing it. Note that I did not incorporate the SLN or any other explicit hierarchical information in the cost.

In deconstructing the shortest path structure of the weighted graph I found that simply applying Dijkstra’s algorithm does not do justice to its semi-metric nature, as it yields unrealistic paths by taking long detours (sometimes up to 20 jumps in my analysis) in the pursuit of minimum cost instead of taking the direct, though more expensive route (Figure 17). This strategy might be appropriate in the case of going on a budget vacation abroad, but it is certainly not realistic for a network of cortical areas with complex inner structure. The signal arriving into an area has to pass through a myriad neuronal somas, axons and synapses, all of which impose a delay on its flow that cannot be neglected respective to the relatively quick propagation of action potentials along interareal axonal bundles.

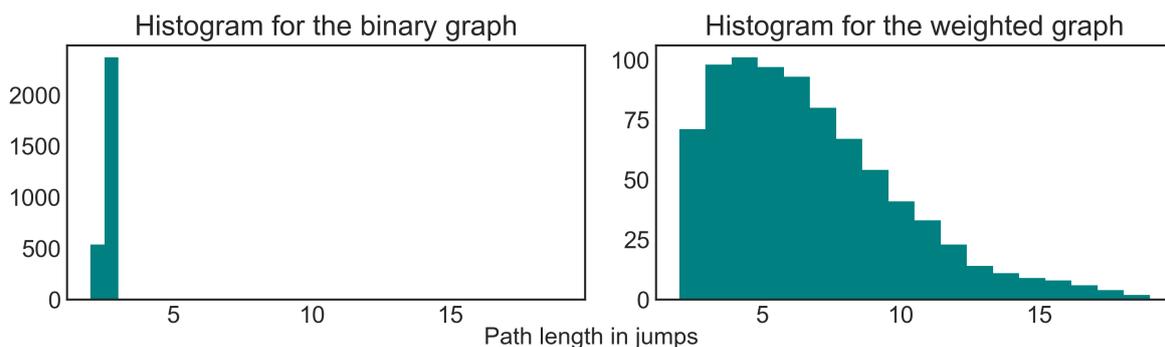


Figure 17: Histograms for the binary and weighed graphs showing the distributions of shortest path length in jumps. In the binary case the maximal length is three, while in the weighted graph it is 18. This is due to the explicitly cost-minimizing nature of the classical approach.

Another problem that I found with the classical approach for finding weighted shortest paths is purely numerical. In fact, binary graphs can be understood as a subtype of weighted ones, where the path costs are ‘binned’, that is, they are *integers*, instead of *real*

numbers. Since integers provide much smaller variability in a given interval, binary path lengths will fall into only a few ‘categories’, that is, many of them will have the same length. This leads to the already mentioned fact, namely that in a binary graph usually there are several shortest paths connecting any given node-pair. In weighted graphs on the other hand, since the weights, and therefore costs are real numbers, it is extremely rare that two path lengths coincide, resulting in a shortest path structure that relies on a few very popular edges, imposing an unrealistic burden on them. Also, this leads to most of the edges getting excluded from the communication on the graph; real brains could not afford this level of excess. Both of these features are especially clear if we look at the edge betweenness (EB) values of the two graphs (Figure 18).

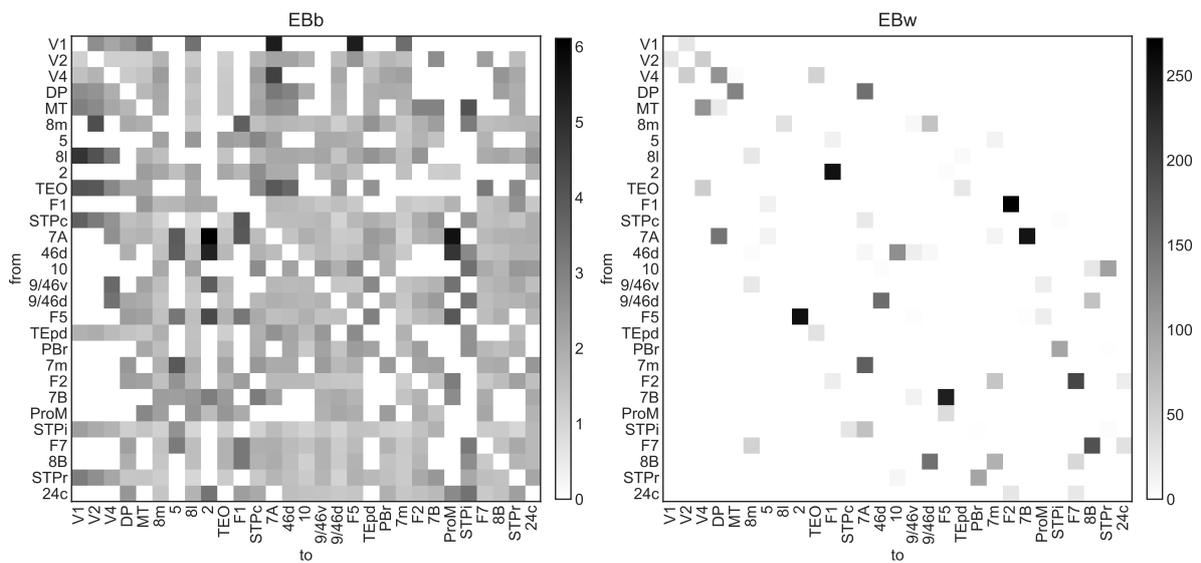


Figure 18: Edge betweenness matrices for the whole 29×29 network in the binary (EBb) and weighted case (EBw). The binary network shows a relatively balanced shortest path distribution where most of the edges contribute to the signal flow. The weighted case on the other hand shows an extremely skewed shortest path distribution relying on a few super-busy edges, while most of the others are simply excluded from the signal flow. This is clearly not in line with real brain networks. Note that (assuming only a single shortest path for every node-pair) the EB values show the exact count of shortest paths comprising a given edge.

To solve the above problems I first employed the α tuning parameter introduced in the Materials and methods. To find the ideal value for alpha I formulated an optimization problem: minimize both the number of zero EBw values (edges excluded from the signal flow) and the highest EBw (the burden on the most popular edge), while keeping alpha the highest possible (to maximize the realism brought by the empirical weights). The left panel of Figure 19 shows the optimization process, which resulted in the exact value of 0.07 for the alpha, assigning much more cost to jumps than to traversing edges. Any value

greater than that brings us back into the realm of nonsensical detours, while lower values discard the significance of weights.

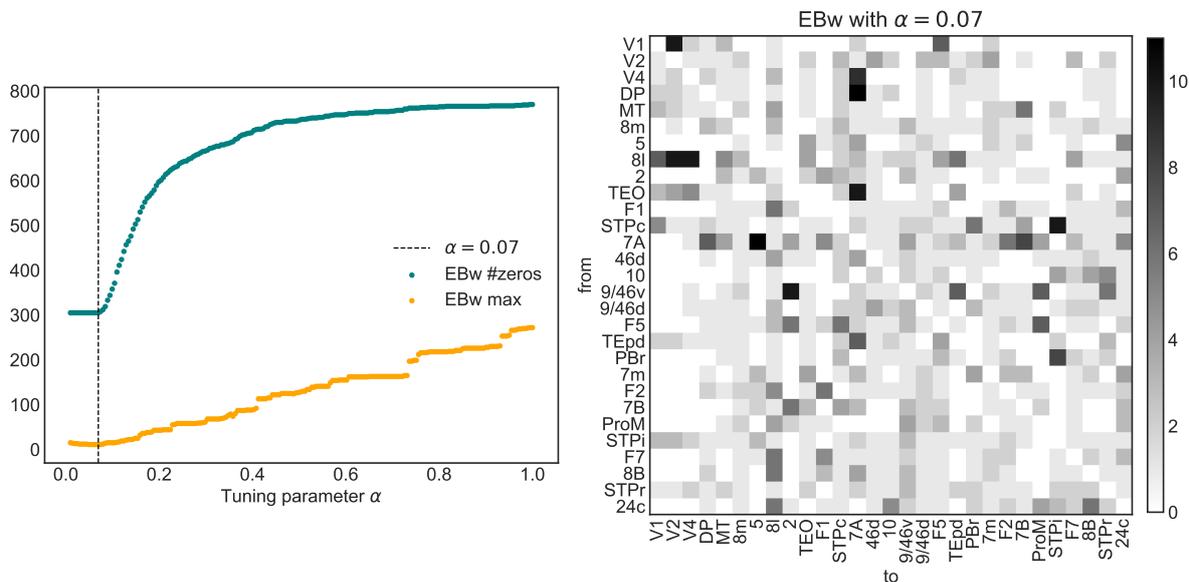


Figure 19: Optimization of the parameter α . Left: the number of unused edges (those with a zero EBw value) and the exclusive popularity of a few edges (the maximum of all EBw values) are minimized, while keeping the significance of the weights (alpha) at the highest value possible. Right: the shortest path structure resulting from the ideal alpha. The maximum is still higher than the one seen in the binary case, but much more reasonable than without the alpha. The number of zeros is the lowest possible (the white cells are all absent edges), although most edges support only a single shortest path (i.e. the edge itself connecting the two nodes at its ends).

The resulting shortest path structure is much closer to what we would expect for a real-world network, but the number of shortest paths is still less than a third of those found in the binary case (812 against 2903, to be precise). This can be observed in the right panel of Figure 19, as a prevalence of edges that support only a few shortest paths, and a handful of those who are 4-5 times more popular. In fact, at this point we can already get moderately meaningful results for the weighted convergence degree (CDw), although due to the small number of shortest paths, and the large number of edges with only one shortest path (i.e. the edge itself), the CD will give a lot of zero values (both the in- and out-set having only one element). This is not the case for the CD computed on the binary graph (CDb) which shows very few zero values, nor it is in line with previous findings that show a very hierarchical picture based on the CD, with minimal balanced edges.

To address these concerns, I ventured to further improve the weighted model by considering not a single, but multiple shortest paths for each node-pair. By including k -shortest paths in the shortest path structure of the graph, we also incorporate a degree of ro-

bustness (in the face of malfunction or damage), which is an observed feature of many real-life networks, including the brain (25). The optimization of k turned out a little bit less straightforward of an objective than it was in the case of α . Investigating the statistics of the EBw in response to changes in k I found that its maximum is unreliable at this scale due to the discrete nature of graphs, while the mean and variance does not change significantly. What does change though is its kurtosis, which really should not be surprising since more alternative shortest paths reduce the burden on the most popular edges by distributing the communication among other edges, thus decreasing the outliers and with them the kurtosis (Figure 20). In fact the kurtosis of the EB distribution is significantly lower in the binary case.

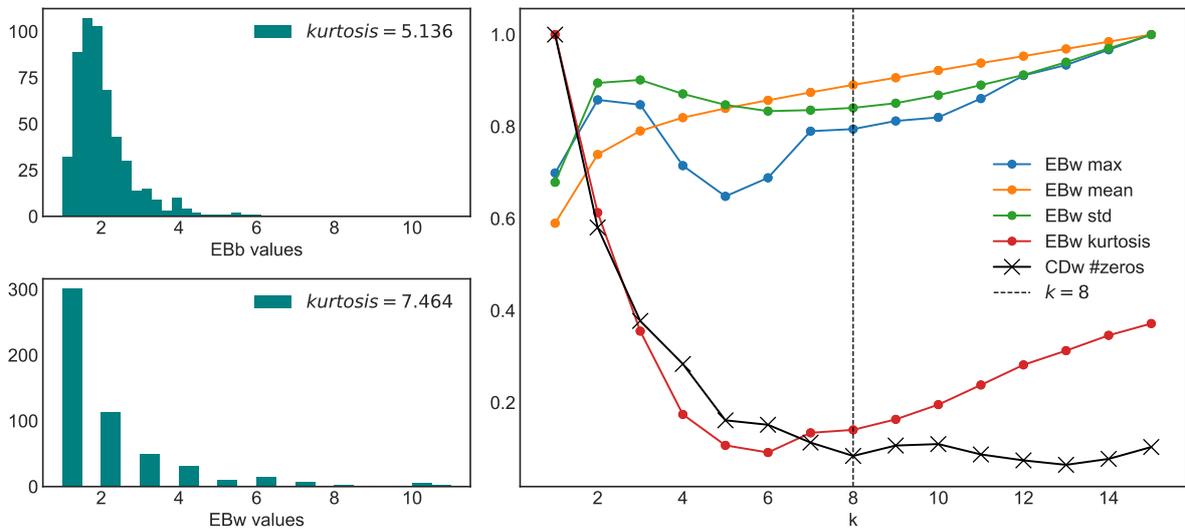


Figure 20: Optimization of the parameter k . Left: histograms of the edge betweenness in the binary (top) and weighted (bottom) case, in the latter already with α included but without k . Note that there are fewer bins in the weighted case, due to the fact that for most node-pairs only a single shortest path exists, therefore the edge betweenness is an integer value, in contrast with the binary case where it is a fraction. Right: normalized EBw statistics and also the normalized number of zeros in the computed CDw, all as a function of k , that is, the number of considered shortest paths for each node-pair. The maximum of the EBw fluctuates due to the small scale (this is supported by the combination of the monotonically increasing mean and the little peak in the standard deviation). The kurtosis falls to a minimum around $k = 6$ and starts to increase again, while the number of zero CDw values reaches a minimum around a $k = 8$ and does not change significantly. Jointly minimizing the kurtosis of the EBw and the number of zeros in the CDw results in an optimal parameter value $k = 8$.

Therefore I settled for a joint minimization of the kurtosis and the number of zero CDw edges, resulting in the exact value of 8 for the k . With α and k chosen this way I arrived at a shortest path structure that is somewhat similar to the binary case (i.e. it has several alternative but similarly short paths between any two nodes), but also takes

into account the actual empirical characteristics of cortical projections (in the form of the empirical cost metric).

As a side-note, I also experimented with several other methods for finding alternative paths instead of the k -shortest paths approach. In particular I tried replacing k with another parameter, e.g. σ , which would be used as a cost limit. This way instead of finding a fixed k number of paths, their actual count would vary, only their cost would be limited (for example as the minimal cost plus σ , or times σ). In the end I found that the actual method does not change the end results in any especially significant way. Computationally the k -shortest paths algorithm is much more efficient and therefore easier to test for large batches of different parameter configurations, so ultimately I stuck with this method, although as a future direction the σ approach could be put to further testing.

4.2 Analysis of the CD flow

I used the shortest and ‘relaxed’ shortest path structures to compute the convergence degree for the binary and weighted graph. Their joint distribution shows a highly significant positive correlation, but it is also clear that the weighted case exhibits a more differentiated picture compared to the binary case, which has a more discrete distribution due to the smaller number of shortest paths (Figure 21).

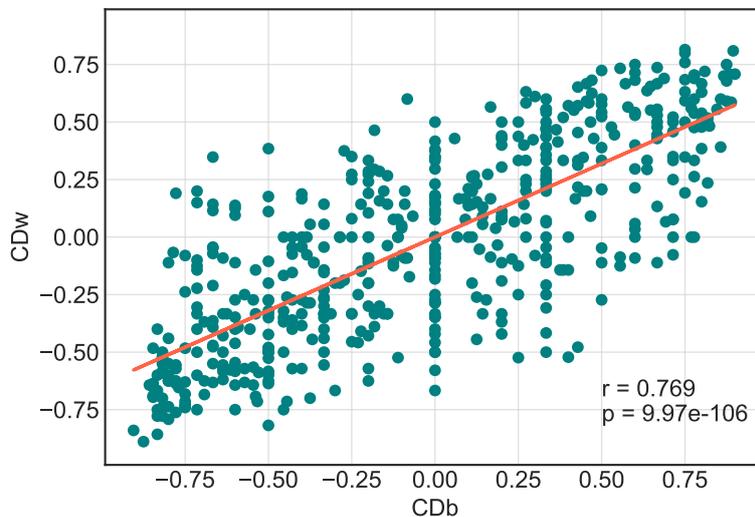


Figure 21: Linear correlation between the binary and weighted CDs (CDb and CDw, respectively). The vertical stacks of points are due to the fact that the binary graph has a relatively sparse shortest path structure. The low cardinality of the in- and out-sets of an edge leads to a ‘lower resolution’, i.e. a higher probability for a zero CD values and non-zero ‘rounded’ fractional values, like -0.5. The weighted graph shows a more ‘fine-grained’ picture, where most edges have a unique CD value.

Although it was not done in previous works, visualizing the CD of a graph via an adjacency matrix has the advantage of clearly showing the distribution of convergent and divergent edges across the nodes. Perhaps even more importantly it is quite intuitive: one does not have to look long to find patterns jump out (see Figures 22 and 23), which can then be tested via more specific visualization methods. In particular, a clear asymmetric distribution can be observed, with mainly convergent edges in the lower triangle and divergent edges in the upper triangle of the matrix. If we compare the CD matrices to the SLN matrix (Figure 13 for the 8×8 subnetwork and Figure 31 in Appendix F for the 29×29 version), this asymmetry corresponds well to the notion of Négyessy et al. (33) that convergent edges match with feedback projections and divergent edges with feedforward projections.

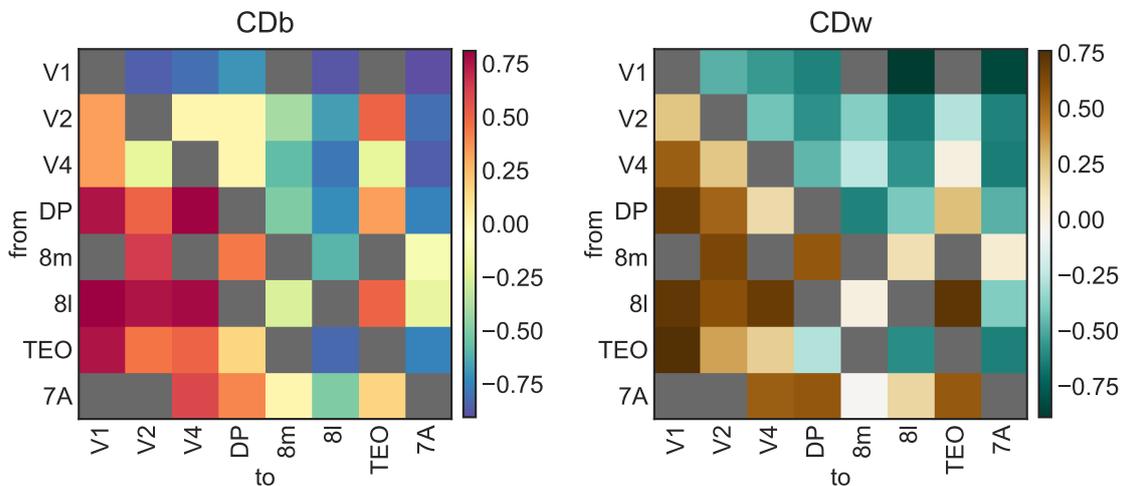


Figure 22: Matrices representing the CD values for the binary (left) and weighted (right) version of the 8×8 subnetwork. Just as for the SLN matrix in Figure 13, both CD distributions are highly asymmetric. The areas are also in the same hierarchical order as in Figure 13, therefore it shows that negative CD values coincide with FF connections (upper triangle from the diagonal), while positive CD is seen for FB connections (lower triangle). It is also clear just by looking at the color distributions, that area TEO should be lower, while frontal eye field areas 8m and 8l should be higher in the CD-based ranking. Gray cells are NaNs, where there is no edge in the graph. See Figures 33 and 34 in Appendix F for the entire 29×29 matrices.

The asymmetrical organization of the CD matrix is perhaps even more striking if instead of the SLN-based hierarchy we reorder the rows and columns according to the CD flow (Figure 23 for the weighted and Figure 35 in Appendix F for the binary case). These reordered matrices show a smooth transition from extremely convergent to extremely divergent connections, through edges that have more neutral CD values. Remarkably, an even more interesting feature can be observed by looking not for the existing, but for

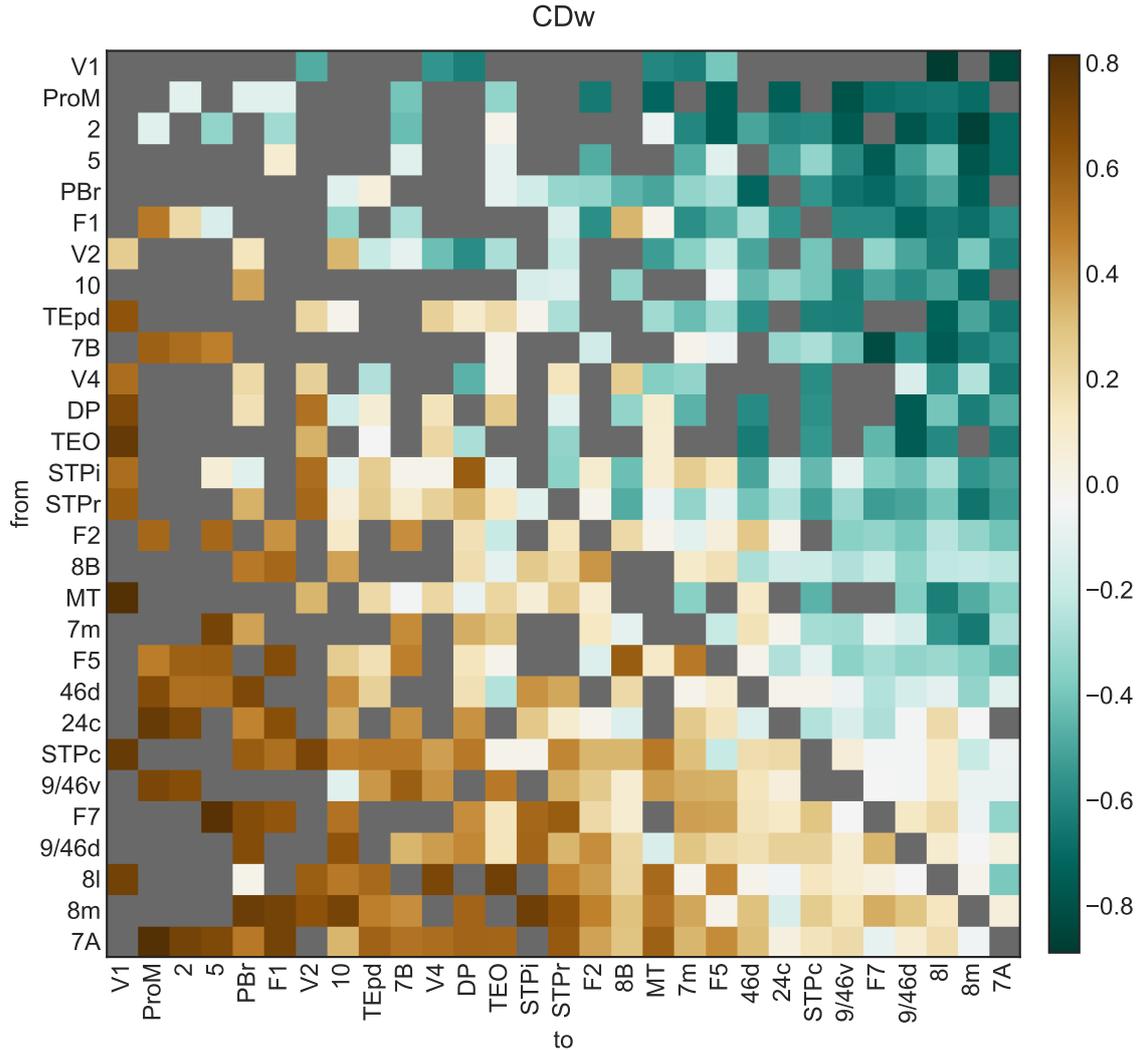


Figure 23: CD values for the weighted 29×29 graph, reordered by the CD flow, revealing several interesting details about the hierarchical organization of the signal flow structure. Perhaps most striking is the almost perfect gradient, perpendicular to the diagonal, moving from extremely convergent (dark brown) to extremely divergent (dark blue) edges, with the zero CD edges mostly aligning in the middle, along the diagonal. Note that the hierarchically ordered SLN did not show this level of smooth separation (see Figure 31 in Appendix F for the SLN matrix of the whole graph for comparison). Another clear feature of the hierarchy based on the CD flow is a densely connected cluster of higher-order areas in the bottom right quadrant, in stark contrast with lower areas which exhibit a much sparser connection pattern. This might correspond to the empirical observation, that less differentiated areas tend to form more connections (18), resulting in a rich-club (19). This is also not seen in the hierarchically ordered SLN matrix. Note that the number of non-edges is exactly the same as in the FLN matrix. See Figure 35 in Appendix F for a reordered version also of the binary CD for the whole graph.)

the non-existent edges. In the CD flow-ranking a clear tendency can be discerned for higher-order areas to have significantly more connections than their lower counterparts.

This clusterization is closely resembling the empirical idea described in the Introduction, that based on complex architectural principles (of which the SLN is only one observation) the differentiation of cortical regions can be distinguished, and that one corollary of this differentiation is that less differentiated (higher-order) areas tend to have more projections (9, 18), forming a rich-club (19). In fact, based on the CD flow a densely connected core of higher-order areas presents itself, that is not seen in the SLN-based hierarchically ordered SLN matrix (see Figure 31 in Appendix F).

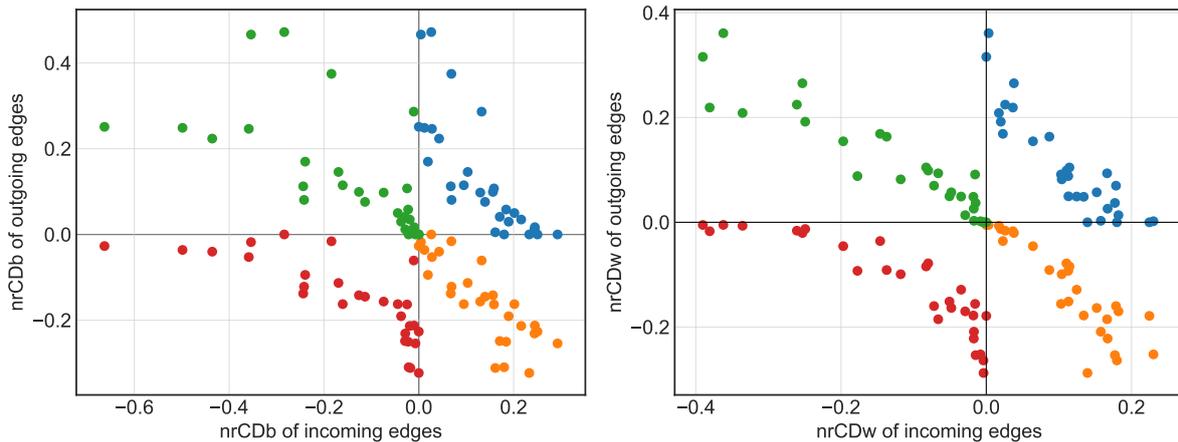


Figure 24: nrCD plots for the entire 29×29 graph, in the binary (left) and weighted (right) case. Note that every node is present in all four quadrants (perhaps somewhat contrary to quick intuition), quantifying different aspects of the CD character of the area the given node represents. On first sight it is clear that the weighted version is much more consistent than the binary. Focusing on the $(-,+)$ and $(+,-)$ quadrants, the nodes align clearly to show a negative correlation between sink (upper left) and source (lower right) attributes. Area V1, which is on the bottom end of the CD flow hierarchy (totally source), would be present here close to the origin in the $(-,+)$ and furthest from the origin in the $(+,-)$ quadrant. The relay attributes of the nodes represented in the $(-, -)$ and $(+, +)$ quadrants show hyperbolic-like shapes, signifying a minimized relay character and therefore a markedly hierarchical organization of the signal flow structure. Note that these are the normalized nrCD values, in contrast to the plot presented in the Introduction; also, the axes are not the same, the weighted case shows much fewer outliers.

Following the example of previous publications I also computed the node-centric CD measures for both the binary and weighted graph. Figure 24 shows the node-reduced CD distributions which returns the characteristic shape found first by Négyessy et al. (33) and then by Bányai et al. (34). The ‘sinkness’ (upper left) and ‘sourceness’ (lower right) quadrants show a clear negative correlation, while the other two ‘relay’ quadrants show hyperbolic-like distributions. Notably, the weighted version of the nrCD plot shows a much cleaner distribution with fewer outliers.

Figure 25 shows the average incoming and outgoing CD values of nodes, which are ranked based on the CD flow. The observable hourglass shapes also correspond well to

those found by Négyessy et al. According to the CD flow ranking, this shape corresponds to a gradient from source to sink nodes, in the bottom-up direction.

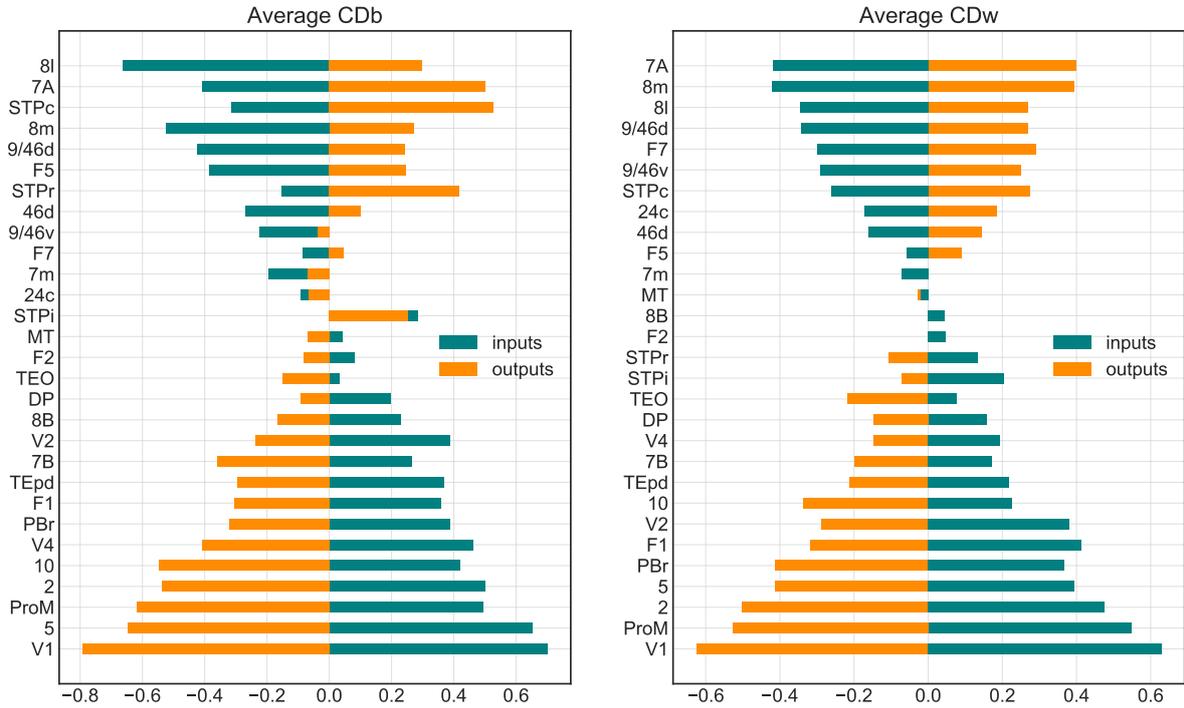


Figure 25: Average incoming and outgoing CD values for the entire binary (left) and weighted (right) graph. The areas are ordered by the CD flow from the bottom-up. The same hourglass-shapes can be seen as was reported by Négyessy et al. (33), with mostly source areas residing in the bottom half and sink areas in the top half. The CD flow is similar for the binary and weighted graphs, with a few erratic areas (for example V2 and V4 change places in the binary version).

4.3 Correlations of topology, anatomy and dynamics

Next, I looked for a relationship between the topological, anatomical and functional hierarchies of the 8×8 subnetwork, namely the CDb/CDw, the SLN and the DAI/mDAI. To re-emphasize, neither one of the anatomical or functional hierarchical indices were included in the graphs on which I computed the convergence degrees. Neither did I tune the arbitrary parameters of the weighted model to match either one of them in any deliberate way. Therefore it was quite astounding that a significant correlation showed up between both versions of the CD and the SLN (Figure 26), as well as between both CDs and the DAI/mDAI (Figure 27). The edge betweenness of the relaxed shortest path structure does not show any of these correlations (for the $EBw \times DAI/mDAI$ relations see Figure 28).

To be precise, the correlations with the CDs are negative, according to the inverse relation between the CD and FF/FB connections described by Négyessy et al. Specifically,

positive CD values represent convergent edges, which in turn correspond to feedback communication, that is, an SLN value close to zero; and the other way around for divergent and feedback connections. Accordingly, to make the comparisons clearer I use the additive inverse of the CD (denoted as invCD) in the correlation plots in Figures 26 and 27. This does not change the results in any way, it only serves the purpose of a more consistent visualization.

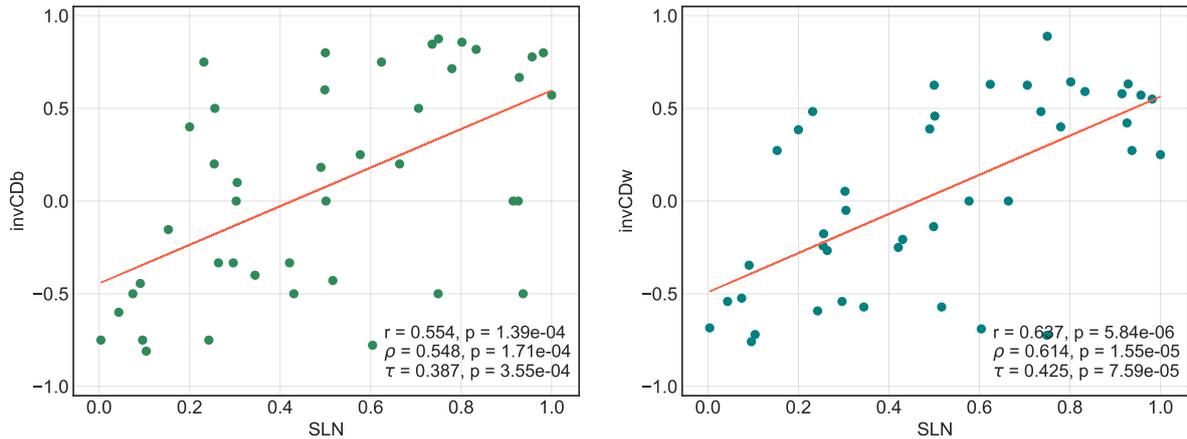


Figure 26: Correlations between the SLN and the inverse CD for the binary (left) and weighted (right) 8×8 subgraph. The legends show Pearson’s r , Spearman’s ρ and Kendall’s τ , with their corresponding p -values. All correlations are significant ($p < 10^{-4}$ for the binary and $p < 10^{-5}$ for the weighted case). Note that except for following the example of Mejiás et al. in not including SLN values of the two erratic extremes of zero and one, no outlier exclusion was done in any of the plots in this work (except perhaps for the accidental partial covering of a point with the legend).

Although the correlations do appear in the binary case as well as in the weighted case, they are by 1-2 orders of magnitude more significant for the latter. Since the weighted graph has two arbitrary parameters I wanted to see how these affect the relationship between the CDw and the SLN/mDAI. To test the dependency of these correlations on α and k I calculated them as the joint function of these parameters (Figure 29).

For the SLN it is clear that by increasing α (i.e. decreasing the cost of the number of edges with respect to their metric length) the correlation falls off quickly, most probably due the shortest path structure becoming ‘superfocused’. Apparently k just by itself cannot solve this problem, but it is also clear that if we settle for the ideal α that I found by optimizing for the edge betweenness, increasing k also increases the correlation. Notably, the parameter values I found ideal (Figures 19 and 20) are not the best choices if we want to maximize the topology-structure or topology-function correlations.

According to the left plot in Figure 29, to achieve a maximal correlation in the CDw \times SLN case, an α of 0.1 and a much higher value for k (specifically 19, if we consider a maximum of 20, but possibly even higher) would fit the most. In fact, in my tuning pro-

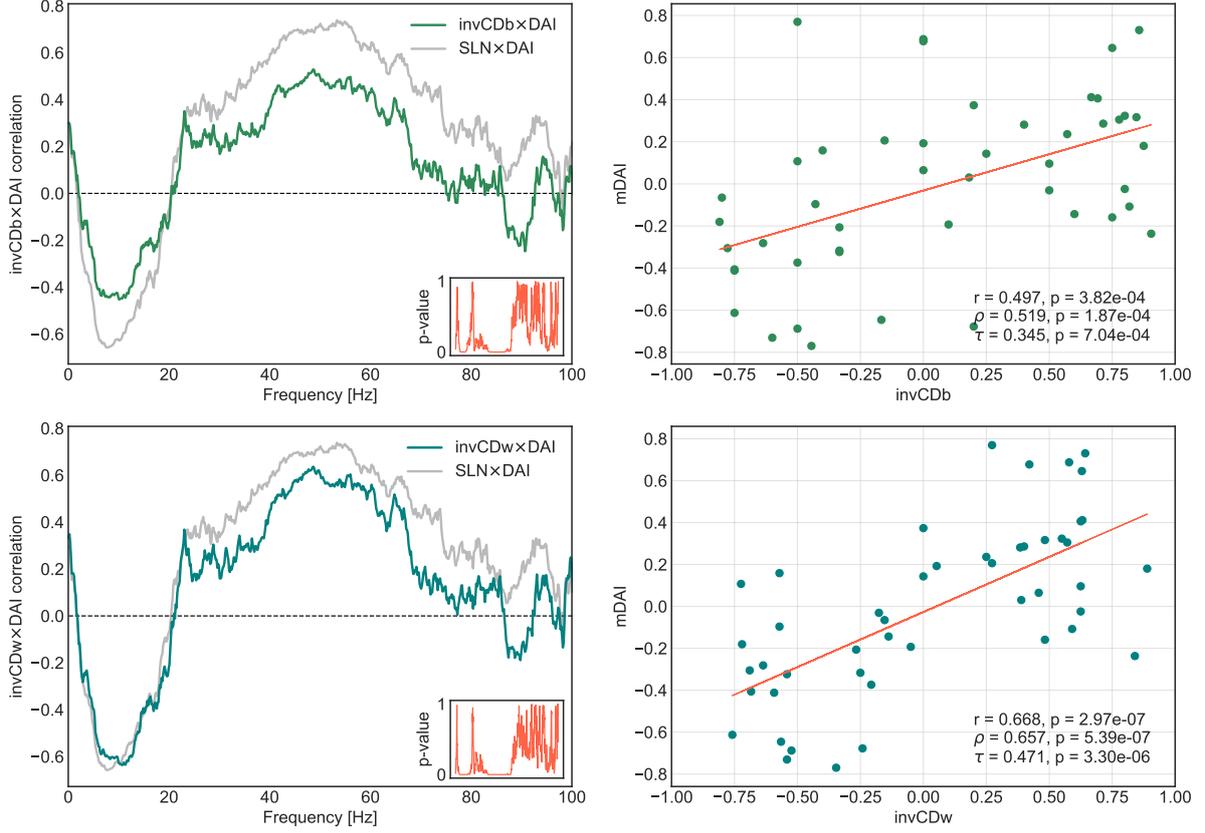


Figure 27: Relationship between the DAI and the inverse binary CD (top) and the inverse weighted CD (bottom) for the 8×8 subgraph. The layout is the same as with the SLN in Figure 16, seen in the left plots in light gray for comparison. Left: Spearman rank correlations for the invCDs and the DAI as functions of frequency. The insets show the p-values; just as with the SLN, there is a significant negative correlation in the alpha-band and a significant positive correlation in the gamma-band for both invCDs. Remarkably, the weighted CD shows almost the exact level of correlation in the alpha-band (for feedback connections) as the SLN, falling short only in the gamma-band (for feedforward connections). Right: correlation between the invCDs and the mDAI. The legends show the correlation coefficients and their corresponding p-values in the same manner as in Figure 26. All correlations are significant ($p < 10^{-4}$ for the binary and $p < 10^{-6}$ for the weighted case).

cedure for k , illustrated in Figure 20 I did not even consider such a high value as possible (only $k \in [1, 15]$). The dependence is not the same for the SLN and mDAI. According to the right side of Figure 29 the correlation between the invCDw and the mDAI shows a maximum for the parameter values $\alpha = 0.22$ and $k = 2$. Further testing might illuminate the exact nature of the dependency for the two cases of correlations.

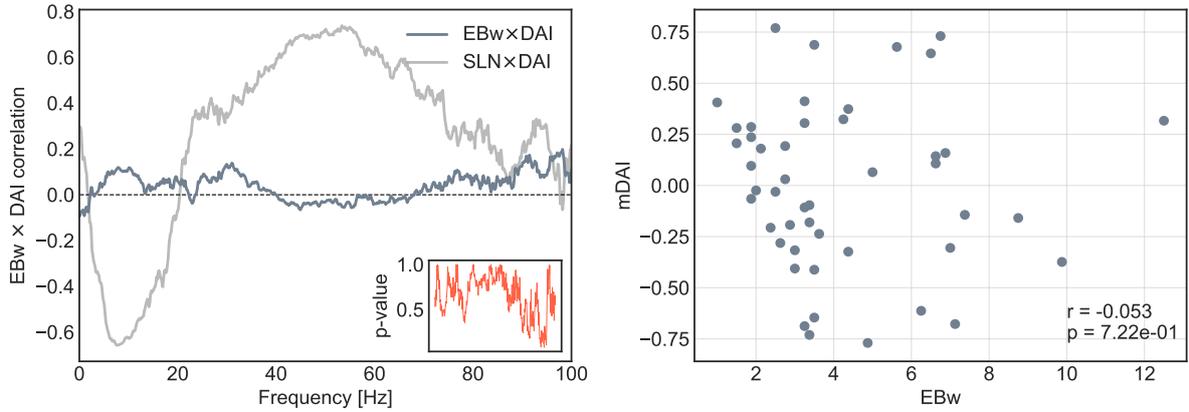


Figure 28: Relationship between the DAI and the edge betweenness of the modified shortest path structure (EBw) for the 8×8 subgraph. No significant correlations can be seen.

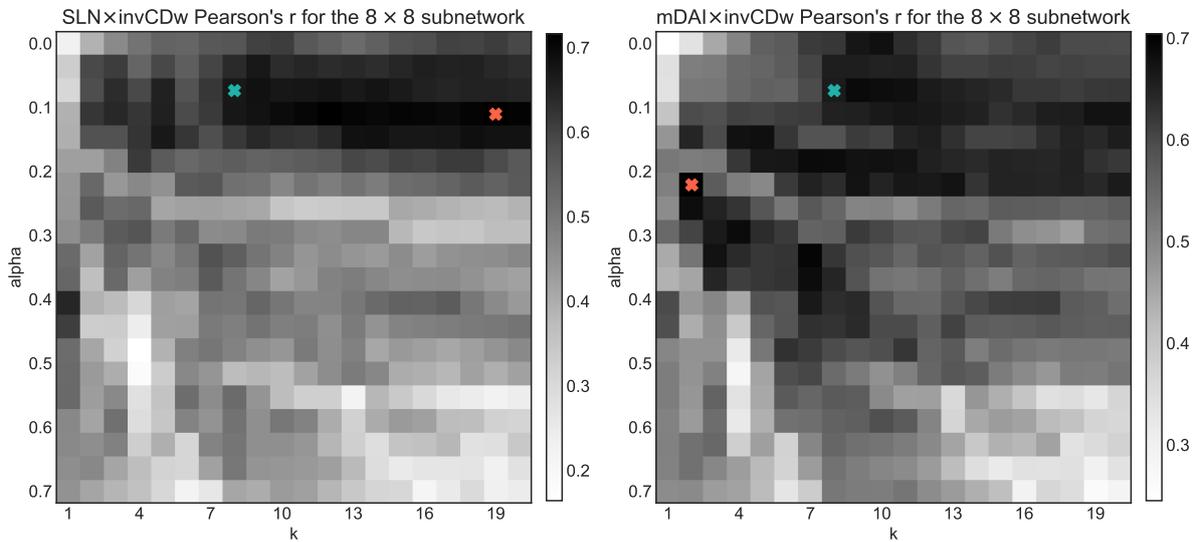


Figure 29: The tuning parameter dependencies of the linear correlations for the invCDw with the SLN (left) and with the mDAI (right) for the 8×8 subnetwork. The blue markers on both plots show the optimized α and k that I found based on the EBw and CDw, while the red markers are the absolute maxima of the 2D arrays. The exact profiles show interesting similarities and also differences in the correlation patterns. Further testing should be employed.

5 Discussion

The work presented here had two main stages. First I had to implement the multi-scale dynamic model adapted from Mejias et al. (23), and recreate their results to test the implementation. During this long and taxing phase I learnt a lot about i) stochastic differential equations and their numerical integration (1, 22, 24), ii) the frequency-domain conversion and analysis of (simulated) electrophysiological time series data (35) and finally iii) about the numerical estimators of directed brain connectivity, especially the conditional spectral Granger causality analysis (36). At the end of this stage I had an empirically plausible dynamic model implementation of the network of cortical areas, showing anatomical (7) and functional hierarchy (17). Additionally, I reproduced most of the findings published by the original authors of the model (23), thus verifying most of their statements.

The second phase was the investigative part of my project. Here I based my analysis on the topological measure of signal flow, namely the convergence degree (CD), introduced by Négyessy et al. (33) and Bányai et al. (34). I successfully extended the CD into the domain of weighted graphs, increasing its potential. To achieve this I employed two extra parameters for relaxing the strict and (for cortical networks) unrealistic assumptions of the classical shortest path definitions (28, 31) in which the CD is rooted. To my best knowledge this combination of the methods have not yet been used in network neuroscientific research. In fact, as of now most published studies of cortical graphs analyze binarized models, even if an empirical weight is actually available (e.g (21)).

I tested and tuned both parameters in an empirically plausible way, relying on the edge betweenness to obtain a realistic, ‘relaxed’ shortest path structure. This has two main differences from the one given by the classical algorithms. First, it takes into account the semi-metric nature of the cortical network, in particular by considering the cost not only of traversing edges but also of passing through nodes (31). This is especially relevant in the brain, since neuronal signals flowing through distinct cortical areas are delayed by multiple synapses and possibly even altered/compromised, making the more direct connections much more efficient (28). Second, by finding not a single, but multiple shortest paths leading from one node to another, the communication backbone of the model network assumes a level of robustness that is seen in real cortical functioning (25), but is missing from the output of the classical algorithm (28, 31). The resulting improved signal flow structure combines the virtues of a purely binary and a purely weighted approach to achieve a realistic map for cortical communication.

I computed the CD profile based on the weighted, relaxed shortest path structure and also for a binarized version of the graph for reference. The correlation of the binary and weighted CD shows that the improved weighted model is more powerful in terms of

resolving fine differences in the signal flow structure. Analysing the CD profile of current anatomical data (7) for both the binary and the weighted case, highly similar distributions were found as in previous studies (33, 34), which were based on older datasets (6). Remarkably, the CD-based ranking of the network nodes, especially the improved weighted case revealed a densely connected component of higher-order areas, in line with prior experimental findings about the rich-club structure of the brain (9, 18, 19). In contrast, the anatomical hierarchy of the model network, namely the SLN does not show such a clusterization.

For the first time it was also shown here that the CD is a viable starting point for the characterization of the structure-function relationship of the cerebral cortex by serving as a communication model (28). In particular, for an 8-area subnetwork previously investigated in the context of visual control (17, 23) the topological signal flow hierarchy (represented by the CD) exhibits a significant positive correlation with the anatomical hierarchy (SLN index, (7)) and also with the functional hierarchy (mDAI index, (17)). The correlation could be seen with both the binary CD and the weighted CD, although notably it was more significant in the latter case. The topological signal flow structure and the level of functional interdependencies are separated by several layers of complex organization and the relationship is far from trivial (28), therefore this result could have exceptional importance. I also tested the dependence of the correlations on the parameters used in computing the weighted CD and found that the dependence is different in the cases of anatomical and functional hierarchy.

Although serious computational limitations exist, these results should be subjected to further, more rigorous testing. The 8 areas comprising the 8×8 subgraph have been selected by Bastos et al. (17), and Mejias et al. (23) based their model partially on the findings of this experiment. They also included in their analyses of functional hierarchy only this subset of areas. This was done first of all due to limitations in computational capacity: in computing the conditional spectral Granger causality, the inclusion of more areas increases the size of the spectral cross matrix (and along with it the running time of the analysis) quadratically. In simple terms calculating the functional connectivity for more than 8 nodes gets out of reach very quickly. Although I have already made relatively spectacular results in optimizing the code I used, much more work is needed for satisfactory effectiveness (see Appendix D for more details on my implementation).

Therefore taking the obvious next step and extending the analysis to the entire network requires a more powerful computational basis, and is not feasible on my current setup. Testing on the other hand can be done (although in a very time- and memory-consuming way) by completing the analysis for randomly sampled 8×8 subgraphs. My preliminary results on this front call for caution in deducting global properties based on the results for the selected 8 areas. In particular, these preliminary findings suggest that the correlations

with the mDAI are not consistent across the network, fluctuating for the random subsets. Remarkably, the correlation strength varies not only between the CD and mDAI, but also between the SLN and mDAI. To re-emphasize, previous studies did not publish any results for other areas apart from the selected 8 (17, 23).

Based on my preliminary tests the correlation between the CD and the mDAI is more consistent than the one between the CD and the SLN. Additionally, for samples where the $\text{invCD} \times \text{mDAI}$ correlation was lower, coincidentally the $\text{SLN} \times \text{mDAI}$ correlation was also lower. Note that these statements are based on a very small sample size (due to the limitations in acquiring the mDAI), and much more testing should be done before making any more decisive claims.

The line of research that I ventured to extend here has many goals. Unraveling the relationship between cortical structure and emergent function would lead to more efficient diagnostic methods based on limited information; for example non-invasive imaging techniques that can be safely employed in the case of human patients (26, 37). In the long run, modelling of this kind may also serve as (at least a partial) substitute for animal trials (38). Ultimately, besides the simple curiosity-based motivation for understanding ourselves, this knowledge would also manifest as further potential in the creation of brain-computer interfaces, a technological development that could perhaps define our future as a species. I look forward to work towards such goals.

References

- (1) Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nature Neuroscience*, *20*, 340–352. <https://doi.org/10.1038/nn.4497>
- (2) Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138. <https://doi.org/10.1038/nrn2787>
- (3) Borges, J. L. (1975). On exactitude in science, In *A universal history of infamy*. London: Penguin.
- (4) Abbott, L. F. (2008). Theoretical neuroscience rising. *Neuron*, *60*, 489–495. <https://doi.org/10.1016/j.neuron.2008.10.019>
- (5) Kandel, E. R., Barres, B. A., & Hudspeth, A. J. (2013). Nerve cells, neural circuitry, and behavior. In E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of neural science* (pp. 21–38). The McGraw-Hill Companies.
- (6) Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1–47. <https://doi.org/10.1093/cercor/1.1.1>
- (7) Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., & Kennedy, H. (2014). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*, *522*(1), 225–259. <https://doi.org/10.1002/cne.23458>
- (8) Olson, C. R., & Colby, C. L. (2013). The organization of cognition. In E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of neural science* (pp. 392–411). The McGraw-Hill Companies.
- (9) Hilgetag, C. C., & Goulas, A. (2020). 'hierarchy' in the organization of brain networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1796), 20190319. <https://doi.org/10.1098/rstb.2019.0319>
- (10) Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in Cognitive Sciences*, *8*(4), 143–145. <https://doi.org/https://doi.org/10.1016/j.tics.2004.02.004>
- (11) Marr, D., Ullman, S., & Poggio, T. A. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.
- (12) Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, *160*, 106–154. <https://doi.org/10.1113/jphysiol.1962.sp006837>
- (13) Amaral, D. G., & Strick, P. L. (2013). The organization of the central nervous system. In E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, & A. J.

- Hudspeth (Eds.), *Principles of neural science* (pp. 337–354). The McGraw-Hill Companies.
- (14) Sporns, O. (2013). Network attributes for segregation and integration in the human brain (Macrocircuits). *Current Opinion in Neurobiology*, *23*(2), 162–171. <https://doi.org/10.1016/j.conb.2012.11.015>
- (15) Tononi, G., Sporns, O., & Edelman, G. M. (1994). A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, *91*(11), 5033–5037. <https://doi.org/10.1073/pnas.91.11.5033>
- (16) Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*, 695–711. <https://doi.org/10.1016/j.neuron.2012.10.038>
- (17) Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., De Weerd, P., Kennedy, H., & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*, 390–401. <https://doi.org/10.1016/j.neuron.2014.12.018>
- (18) Hilgetag, C. C., Beul, S. F., van Albada, S. J., & Goulas, A. (2019). An architectonic type principle integrates macroscopic cortico-cortical connections with intrinsic cortical circuits of the primate brain. *Network Neuroscience*, *3*(4), 905–923. https://doi.org/10.1162/netn__a__00100
- (19) van den Heuvel, M. P., & Sporns, O. (2011). Rich-club organization of the human connectome. *Journal of Neuroscience*, *31*(44), 15775–15786. <https://doi.org/10.1523/JNEUROSCI.3539-11.2011>
- (20) van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, *111*(40), 14332–14341. <https://doi.org/10.1073/pnas.1402773111>
- (21) Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, *104*(24), 10240–10245. <https://doi.org/10.1073/pnas.0701519104>
- (22) Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., & Friston, K. (2008). The dynamic brain: From spiking neurons to neural masses and cortical fields. *PLOS Computational Biology*, *4*(8), 1–35. <https://doi.org/10.1371/journal.pcbi.1000092>
- (23) Mejias, J. F., Murray, J. D., Kennedy, H., & Wang, X.-J. (2016). Feedforward and feedback frequency-dependent interactions in a large-scale laminar network of the primate cortex. *Science Advances*, *2*(11). <https://doi.org/10.1126/sciadv.1601335>

- (24) Cowan, J. D., Neuman, J., & van Drongelen, W. (2016). Wilson–cowan equations for neocortical dynamics. *The Journal of Mathematical Neuroscience*, *6*. <https://doi.org/10.1186/s13408-015-0034-5>
- (25) Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*, 186–198. <https://doi.org/10.1038/nrn2575>
- (26) Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*, 353–364. <https://doi.org/10.1038/nn.4502>
- (27) Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, *45*(2), 167–256. <https://doi.org/10.1137/s003614450342480>
- (28) Avena-Koenigsberger, A., Misic, B., & Sporns, O. (2018). Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, *19*, 17–33. <https://doi.org/10.1038/nrn.2017.149>
- (29) Newman, M. E. J. (2001). Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, *64*, 016132. <https://doi.org/10.1103/PhysRevE.64.016132>
- (30) Newman, M. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, *27*(1), 39–54. <https://doi.org/https://doi.org/10.1016/j.socnet.2004.11.009>
- (31) Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*(3), 245–251. <https://doi.org/https://doi.org/10.1016/j.socnet.2010.03.006>
- (32) Ercsey-Ravasz, M., Markov, N. T., Lamy, C., Van, D. C., Knoblauch, K., Toroczka, Z., & Kennedy, H. (2013). A predictive network model of cerebral cortical connectivity based on a distance rule. *Neuron*, *80*, 184–197. <https://doi.org/10.1016/j.neuron.2013.07.036>
- (33) Négyessy, L., Nepusz, T., Zalányi, L., & Bazsó, F. (2008). Convergence and divergence are mostly reciprocated properties of the connections in the network of cortical areas. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1649), 2403–2410. <https://doi.org/10.1098/rspb.2008.0629>
- (34) Bányai, M., Négyessy, L., & Bazsó, F. (2011). Organization of signal flow in directed networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2011*(06), P06001. <https://doi.org/10.1088/1742-5468/2011/06/p06001>
- (35) Prerau, M. J., Brown, R. E., Bianchi, M. T., Ellenbogen, J. M., & Purdon, P. L. (2017). Sleep neurophysiological dynamics through the lens of multitaper spectral analysis. *Physiology*, *32*(1), 60–92. <https://doi.org/10.1152/physiol.00062.2015>
- (36) Wen, X., Rangarajan, G., & Ding, M. (2013). Multivariate granger causality: An estimation framework based on factorization of the spectral density matrix. *Philo-*

- sophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1997), 20110610. <https://doi.org/10.1098/rsta.2011.0610>
- (37) Douw, L., van Dellen, E., Gouw, A. A., Griffa, A., de Haan, W., van den Heuvel, M., Hillebrand, A., Van Mieghem, P., Nissen, I. A., Otte, W. M., Reijmer, Y. D., Schoonheim, M. M., Senden, M., van Straaten, E. C. W., Tijms, B. M., Tewarie, P., & Stam, C. J. (2019). The road ahead in clinical network neuroscience. *Network Neuroscience*, 3(4), 969–993. https://doi.org/10.1162/netn_a_00103
- (38) Aradi, I., & Érdi, P. (2006). Computational neuropharmacology: Dynamical approaches in drug discovery. *Trends in Pharmacological Sciences*, 27, 240–243. <https://doi.org/10.1016/j.tips.2006.03.004>
- (39) Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents — eeg, ecog, lfp and spikes. *Nature Reviews Neuroscience*, 13, 407–420. <https://doi.org/10.1038/nrn3241>
- (40) Seth, A. (2007). Granger causality (revision #127333). *Scholarpedia*, 2(7), 1667. <https://doi.org/10.4249/scholarpedia.1667>
- (41) Mejias, J. F. (n.d.). *Large-scale laminar model of macaque cortex*. <https://senselab.med.yale.edu/ModelDB/showmodel.cshtml?model=249589&file=/Mejias2016/readme.html> (Accessed: 23.05.2020)
- (42) Denovellis, E. (n.d.). *Spectral connectivity package*. https://github.com/Eden-Kramer-Lab/spectral_connectivity (Accessed: 23.05.2020)

Appendices

A Retrograde tract-tracing

Several entirely different methods exist for mapping neural connections, each with distinct advantages and disadvantages. Retrograde tract-tracing is an invasive method for precisely identifying interareal projections. It involves injecting a tracer into a cortical area, which is taken up in the synaptic clefts to the axon terminals of the presynaptic neurons. There, ‘hijacking’ the normal microtubule-based axonal transport the tracer travels back (hence *retrograde*; anterograde tracers also exist) to the cell body, which is the origin of the projection. There it can be detected (e.g. via fluorescent microscopy), making it feasible to count the projecting neurons thus quantifying the strength of the interareal connection (7).

B Electrocorticography

Classical electroencephalography (EEG) measures the electrical activity of the brain with external electrodes placed on the scalp. The recorded signal is a distorted and attenuated version of the local field potential (LFP), filtered through the soft and hard tissues between source and recording electrode. Electrocorticography (ECoG) is done by bypassing the bothersome non-neural tissues and placing the electrodes directly on the surface of the brain; it is an invasive method which necessitates the surgical opening of the skull, therefore it is done mostly on animal subjects or humans in clinical environments, during or preceding operations. Note that in contrast with depth electrodes, during ECoG the electrodes are not directly inserted in the neural tissue. The resulting recordings have relatively high temporal (approx. 5 ms) and spatial (approx. 1 cm) resolution (17, 39).

C Causality analysis

Granger causality (GC) quantifies the directed influence between two processes, based on the ability to predict future values (40). In particular, if the accuracy of predicting future values of process X is improved by including information about process Y, with respect to the prediction considering solely values of X itself, then we say that Y ‘Granger causes’ X. Mathematically this is formulated by comparing the prediction error (i.e. variance) terms of two autoregressive models, one with both X and Y included and one with only X. For a given process X, the autoregressive model of order n is an equation between the predicted value of X and the sum of n terms comprising n past values of X, each one with a unique coefficient, and the error term. In case of more than two processes this can

be distorted by indirect influences, therefore a conditioning can be made to exclude the effect of processes beyond the pair investigated.

The analysis can be transformed from the time-domain to represent the frequency-based influences between pairs of processes. Notably, although Granger first published his method in 1969, it is an open topic ever since. For this work I implemented the non-parametric version of the spectral GC, which estimates the influences from the cross spectral matrix of the processes, without an order parameter. I based my implementation mainly on (36) but I used other sources too to gain a moderately deep understanding of the method.

D Technical notes on the implementation

All work was done on my not-especially-formidable but sturdy and time-tried laptop, sporting nothing more and nothing less than 8 GBs of RAM and a 2.6 GHz CPU with four cores (I did not employ the GPU in the computations for the time being). I wrote almost all my code from scratch in Python. Dr. Jorge F. Mejias and Prof. Xiao-Jing Wang were kind enough to answer my plea at a desperate moment in this adventure and provide me with their Matlab code. In fact, as I had to find out, their code is publicly available (although in a somewhat hidden place) (41). Unfortunately (or perhaps fortunately, since I undoubtedly gained much by doing things myself), by that point I had already reverse engineered almost the entire model and implemented it in Python (which I absolutely prefer to Matlab in any case). Nonetheless, scouring their code was helpful in verifying my home-brew solutions, and also illuminated some hidden details and coding tricks which I had not come up with by myself. Lately I also found out that there exist also a publicly available Python implementation of the model, but it is incomplete (lacks the entire large-scale model), and in my humble opinion is written in a very obscure style. With due respect, mine is simply better.

I wrote all my code in a Jupyter notebook framework, relying heavily on the public libraries NumPy, SciPy, Pandas and NetworkX. The plots have been made with Matplotlib. Besides these libraries I built all my routines from scratch, in most cases in a functional paradigm. For the computation of the conditional version of the spectral Granger causality I have not found any available Python-repositories. The closest one that works well is the *spectral connectivity* package made by Eric Denovellis (42), which has the pairwise version of the spectral Granger causality analysis (SGCA), and also many other useful bi- and multivariate autoregressive-model-based connectivity measures, but not the conditional SGCA. Also, it is brutally memory-greedy, using up hundreds of gigabytes at a time for parallel manipulating gigantic matrices, making it impossible to scale up the analyses.

Therefore I rewrote most of his object-oriented framework to run in a much more efficient way (generators can do magic: I achieved a speedup on the order of x100, reducing a normal running time from days to a mere hours), enabling the running of multiple trials and the investigation of different parameter configurations in moderately reasonable time. In particular I use his (improved) multitaper method for time-to-frequency domain transformation to obtain high-resolution power spectra of the simulated time series data, which is especially needed for the causality analysis. For proper parametrization of the multitaper analysis I referred to the description in (35). Ultimately, I also implemented my own non-parametric conditional spectral Granger causality estimator (all my results were made with this), based on the procedure described in (36). Although on this front the battle is still waging (the scalability of the analysis is still very much constrained), I look forward to making publicly available all my code on my github profile (github.com/celestialteapot).

E Parameter values for the dynamic model

Most of the way I followed Mejias et al. (23) in choosing the right parameters for the model. The time constants in case of all levels of the model are $\tau_{L2E} = 6 \text{ ms}$ and $\tau_{L2I} = 15 \text{ ms}$ for the supragranular excitatory and inhibitory populations, respectively, and $\tau_{L5E} = 30 \text{ ms}$ and $\tau_{L5I} = 75 \text{ ms}$ for the infragranular populations. The lower values in the former case lead to the gamma-band oscillations, while the higher values in the latter case result in the lengthening of the period, therefore achieving the desired alpha-band activity. The strengths (i.e. standard deviations) of the Gaussian random variables in the noise terms are $\sigma_{L2} = 0.3$ and $\sigma_{L5} = 0.45$ for L2/3 and L5/6, respectively; they are the same for excitatory and inhibitory populations. The layer- and population-specific connection strengths are $J_{EE} = 1.5$, $J_{EI} = -3.25$, $J_{IE} = 3.5$, $J_{II} = -2.5$, $J_{52} = 1$ and $J_{25} = 0.75$, where the second indices indicate the source population and the first is the target (a convention used by Mejias et al.). The interareal layer-specific connection strengths are $J_{FF1} = 1$, $J_{FB1} = 0.1$, $J_{FB2} = 0.5$, $J_{FB3} = 0.9$ and $J_{FB4} = 0.5$. For transforming the FLN $c_1 = 1.2$ and $c_2 = 0.3$ was used.

I ran all of my simulations with a dt value of 0.2 ms , corresponding to a sampling frequency of 5 kHz . For most cases of laminar simulations a total time period of $T = 30 \text{ s}$ was generated, with a transient of 5 s . For the interareal and large-scale model much longer time periods are needed, which I chose as $T = 180 \text{ s}$ with a transient of 20 s . The causality estimations were done on downsampled time series, taking the mean for every 25 values; this does not distort the data in any significant way. For most spectra running mean smoothing with various reasonable window lengths were applied to account for the omnipresent noise.

F Adjacency matrices

Here I include all the large 29×29 matrices that did not fit the main text.

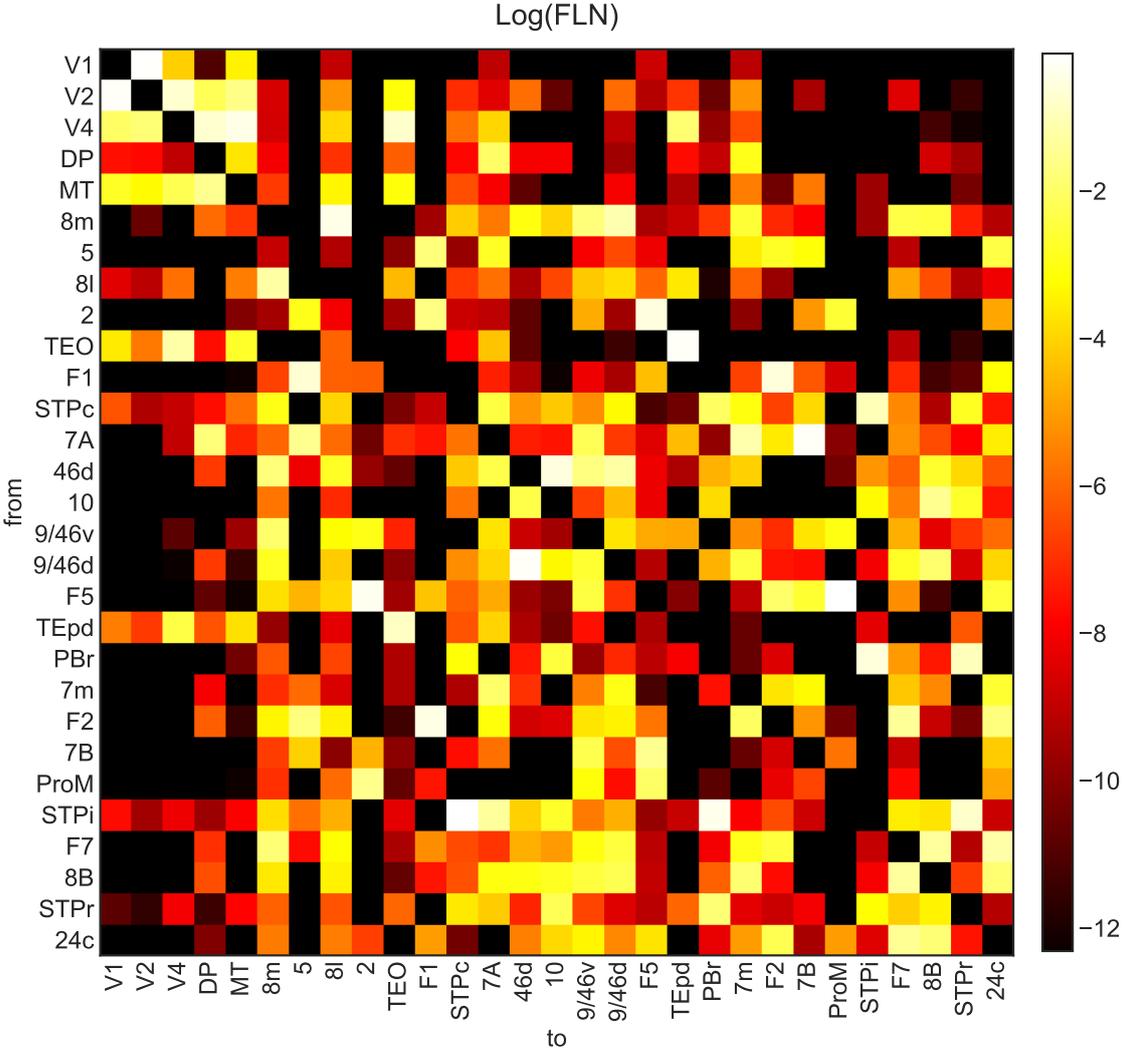


Figure 30: FLN values for the entire network. Nodes are ordered by the SLN-based anatomical hierarchy. Recreated following (23); data from core-nets.org (7).

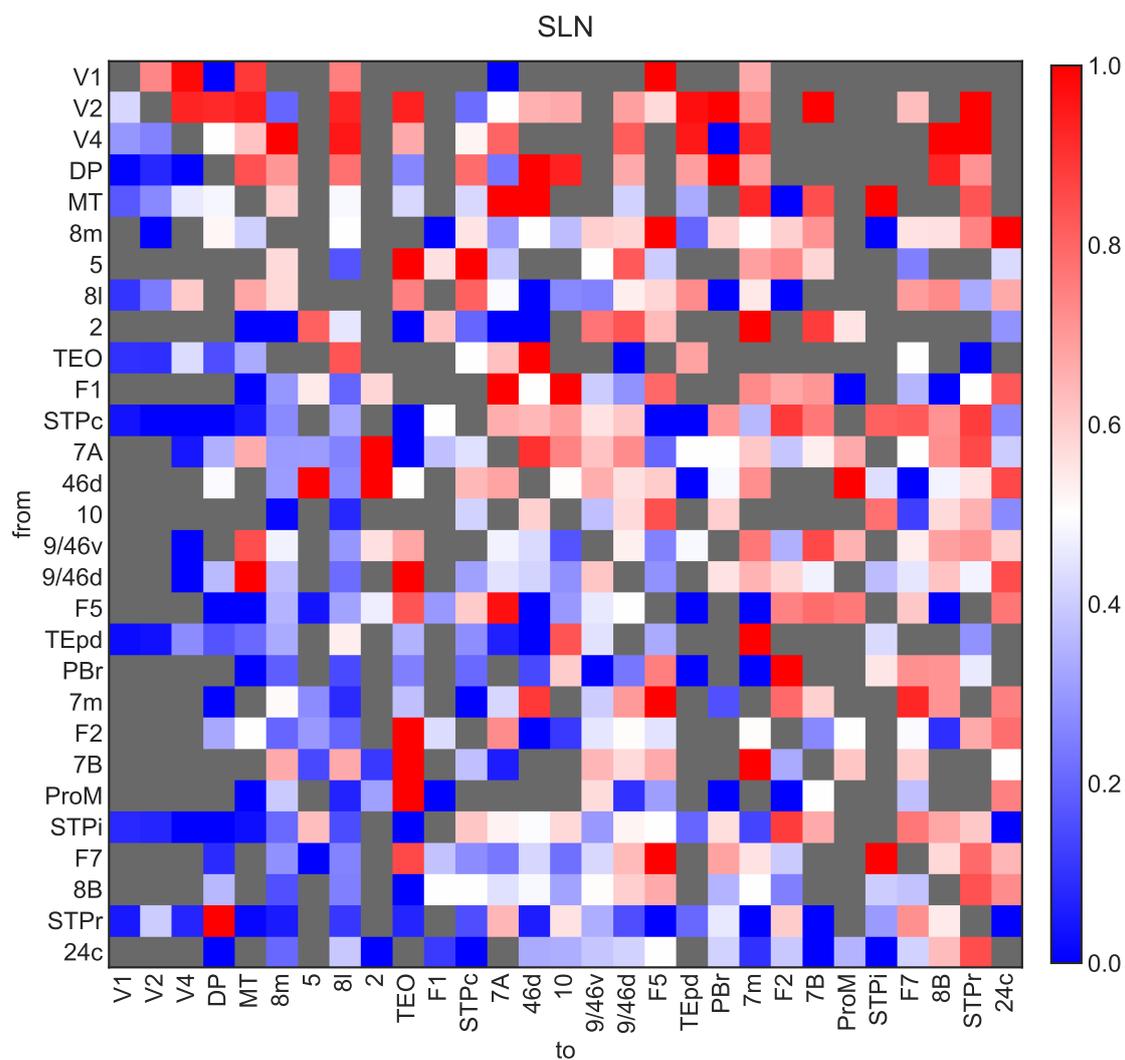


Figure 31: SLN values for the entire network. Nodes are ordered by the SLN-based anatomical hierarchy. Recreated following (23); data from core-nets.org (7).

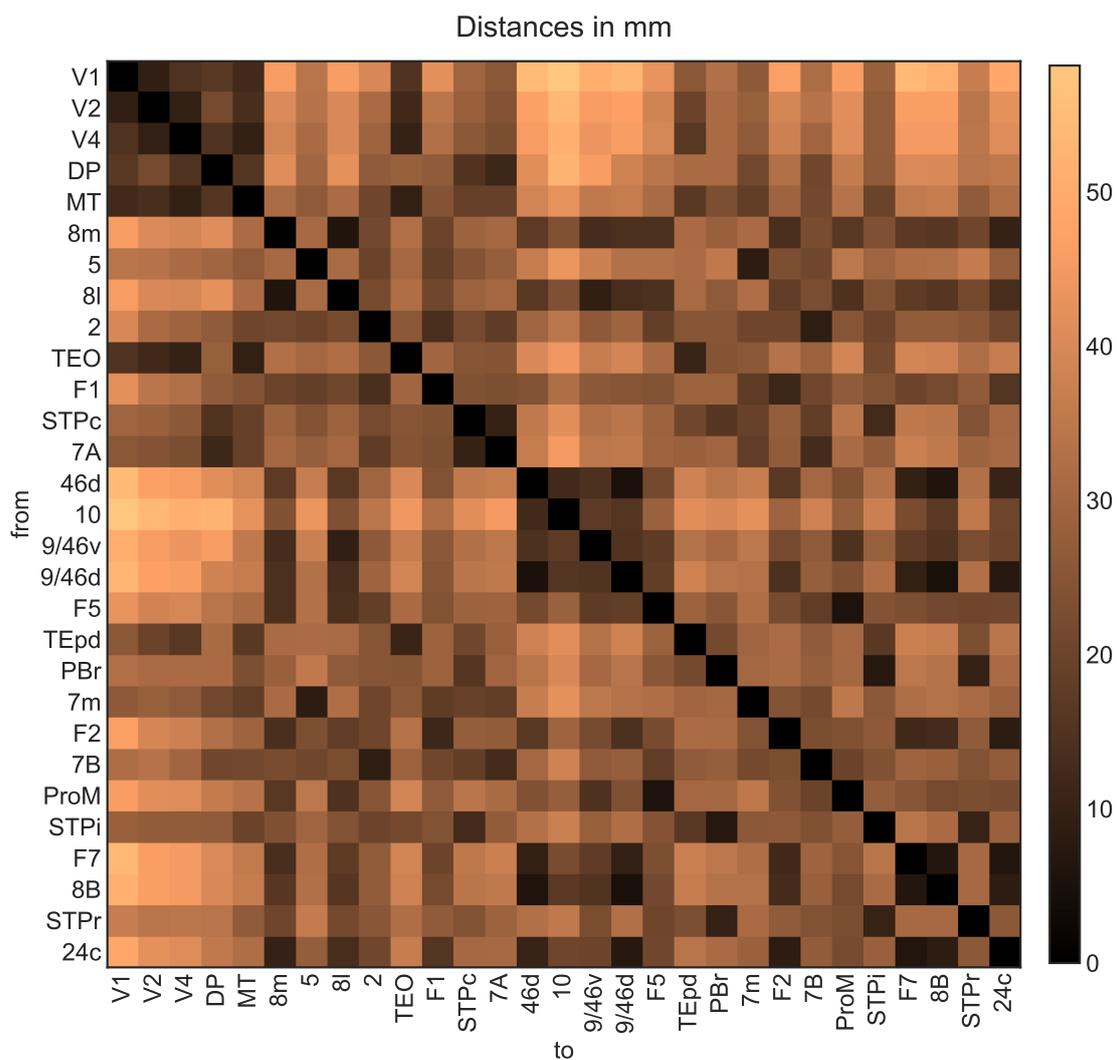


Figure 32: Distances in mm for the entire network. Nodes are ordered by the SLN-based anatomical hierarchy. Recreated following (23); data from core-nets.org (7).

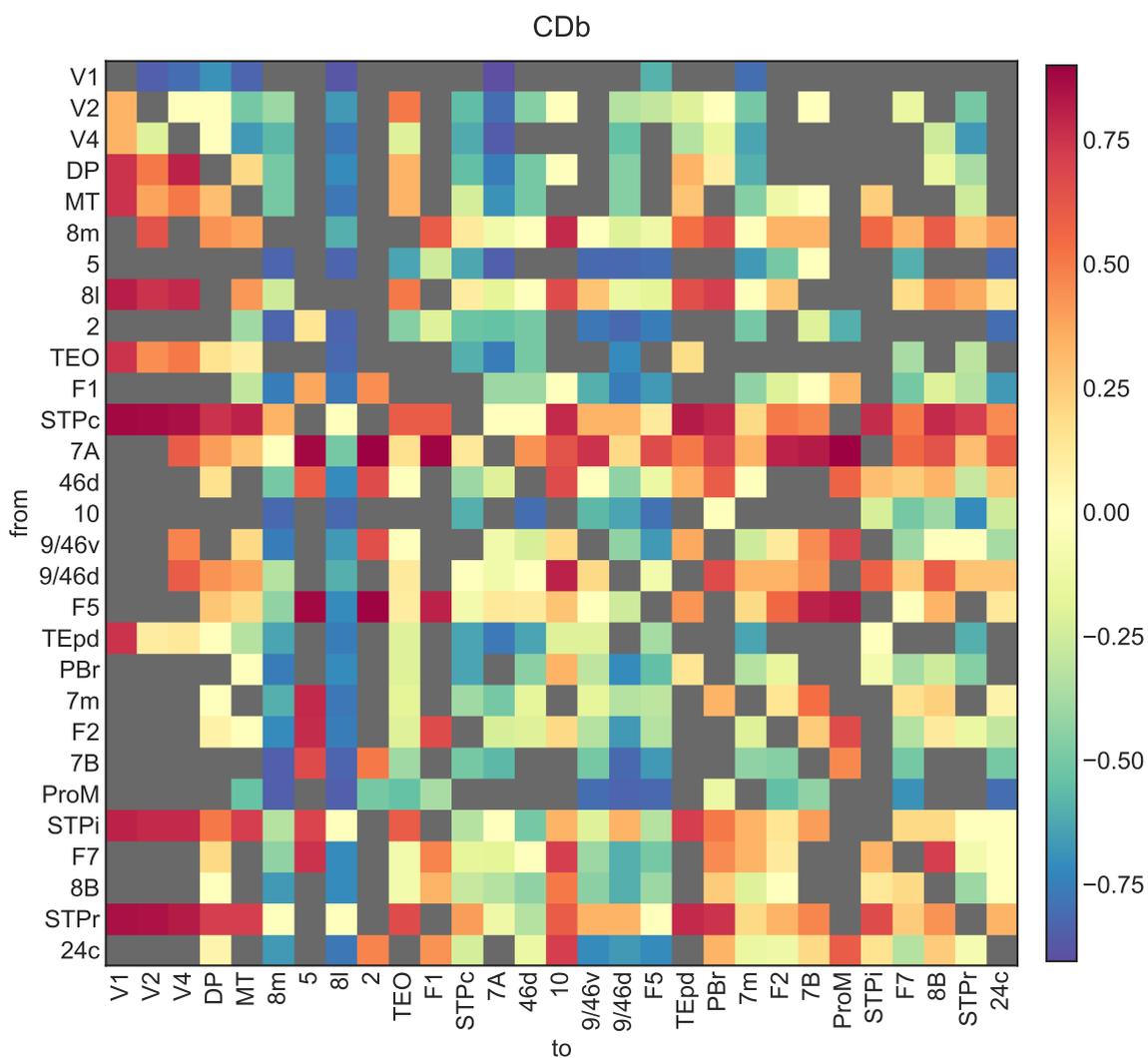


Figure 33: CD values for the binary graph. Nodes are ordered by the SLN-based anatomical hierarchy.

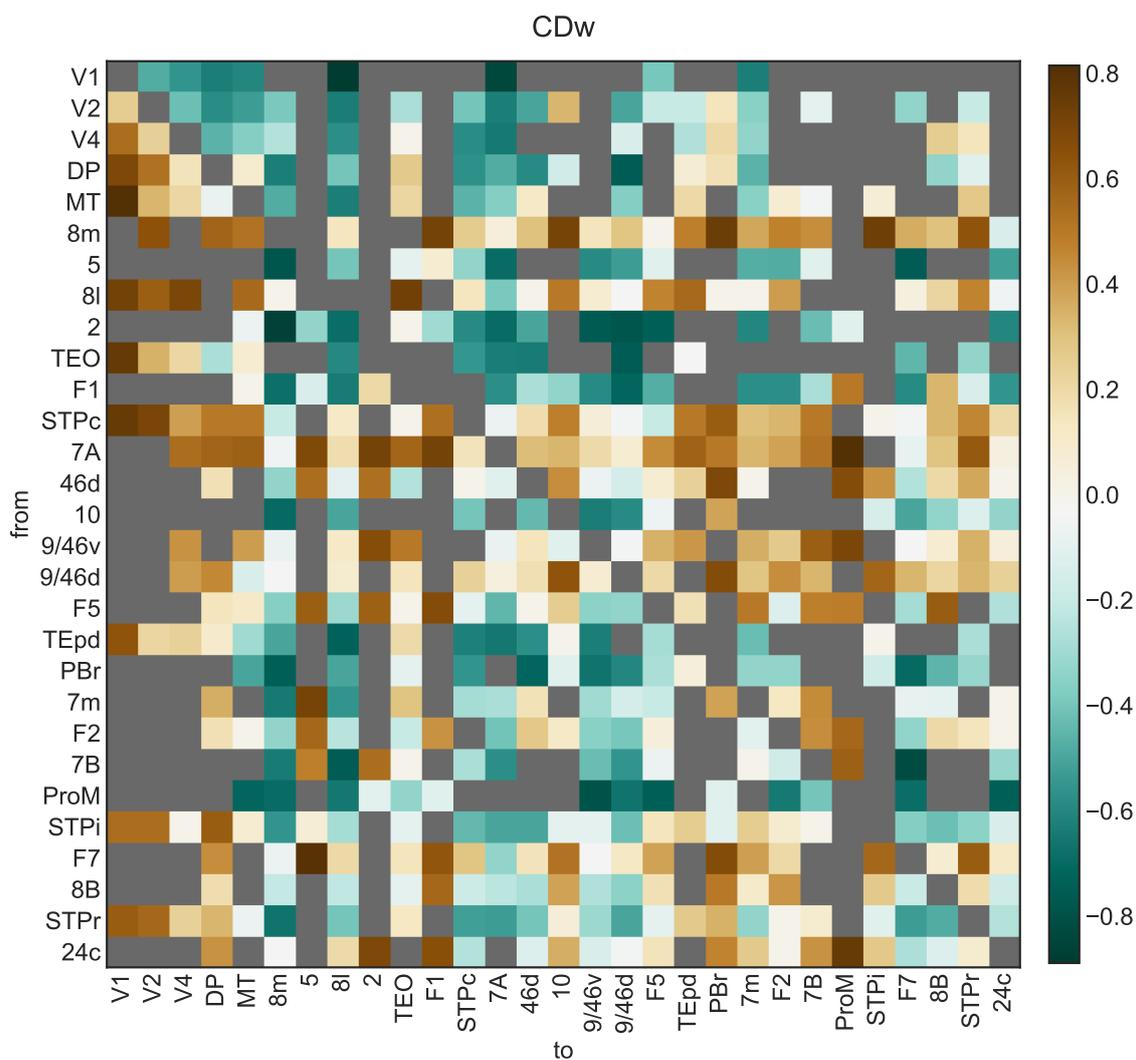


Figure 34: CD values for the weighted graph. Nodes are ordered by the SLN-based anatomical hierarchy.

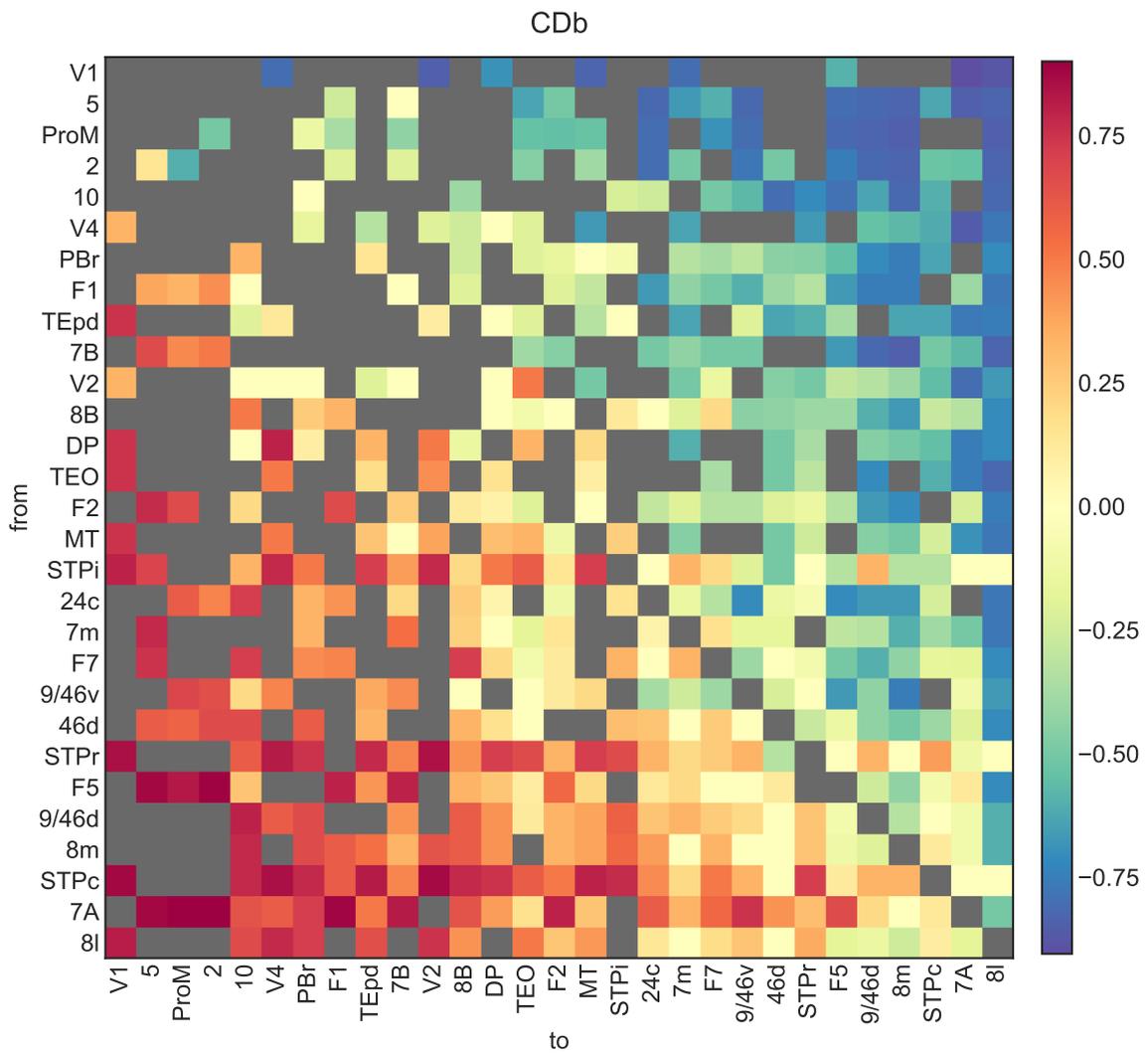


Figure 35: CD values for the binary graph, reordered by the CD flow ranking.