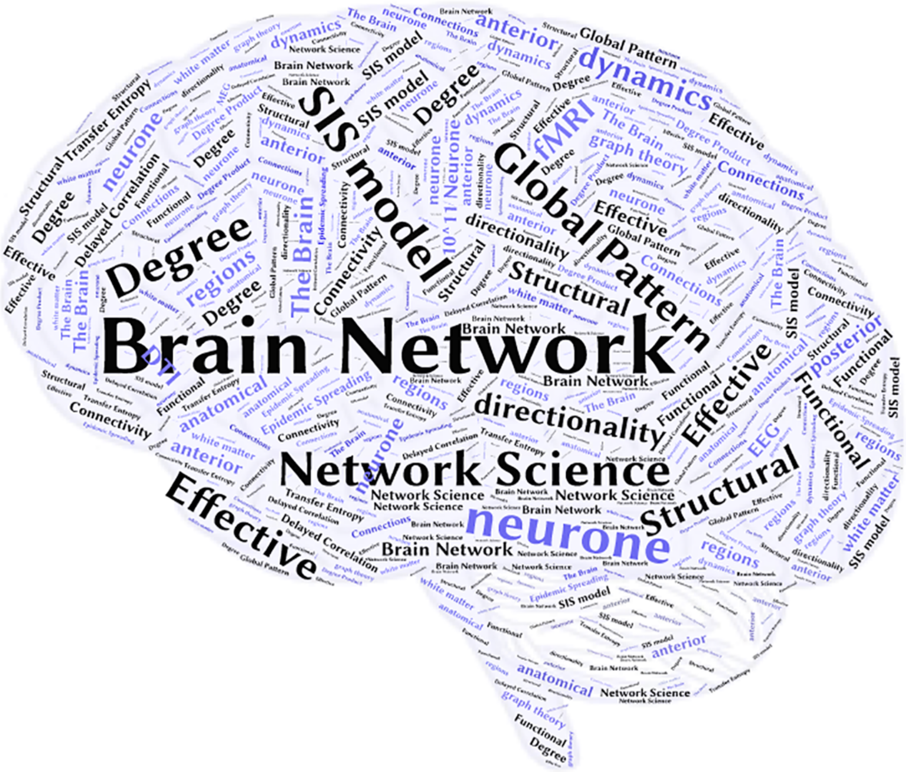
Applying the epidemic spreading model on the brain network to explain effective connectivity

Xiangyu Zhou

Master of Science Thesis



Network Architectures and Services

PVM 2016-087

Delft University of Technology

Faculty of Electrical Engineering, Mathematics and Computer Science
Telecommunications

Applying the epidemic spreading model on the brain network to explain effective connectivity

Xiangyu Zhou
4418654

Committee members:

Supervisor: Prof. Dr. Ir. Piet Van Mieghem
Mentor: M.Sc. Jil Meier
Members: Prof. Dr. Ir. Marco Loog
Dr. Arjan Hillebrand

August 25, 2016
M.Sc. Thesis No: PVM 2016-087

Abstract

Network science studies a complex system as a network to capture the connectivity patterns and topological features. Different network topologies have been observed to shape dynamic spreading processes on the network in various ways, while the exact relationship is complicated and not yet fully understood. Epidemic models are often applied to describe the spreading process on a network and to facilitate studying the dynamic interactions. We apply a simple epidemic model, the Susceptible-Infected-Susceptible (SIS) model, on the structural brain network to explore the topological properties that drive the dynamic processes. A recent study examined the transfer entropy of empirical data and observed a dominant posterior-anterior spreading pattern in the brain. In both transfer entropy and delayed correlation measures, we show that hubs are more sending information to the network than lower degree nodes. With our continuous-time simulations, we also found the empirically-observed posterior-anterior global pattern. Based on our results, the brain topology of hubs mainly located at the back of the brain seems to be responsible for the emergence of the global pattern.

Keywords:

SIS epidemic spreading model, brain structural network, functional connectivity, effective connectivity, delayed correlations, transfer entropy, directionality, global pattern

Contents

Chapter 1

Introduction

1.1 Background

Network science is widely used to study complex systems and problems across various fields [1]. The network of interconnected neurones in the brain and the food web of predators and preys in an ecosystem are examples of biological networks. Man-made networks include the telephone infrastructure networks, railway roads, power grids, the well-known Internet of physically interconnected computer networks, and the World Wide Web (WWW) of virtually linked web pages. Social network examples are various human interactions through physical contact and online social platforms.

The network approach conducts macroscopic analysis to reveal the whole picture of a problem and thus has become popular in the study of many complicated systems. To extract the general organisation properties and to disclose the topology of connection patterns, many details in a system are disregarded and only elements and their interactions are considered. For instance, the entities on the Internet are modelled as computers in general, regardless of differences in operating systems, transmission bandwidths limited by local service providers, as well as geographic landscapes. Upon the resulting simplified topology of connection patterns, routing protocols can be designed to conduct communications through shortest paths effectively and efficiently. However, due to the geographically dispersed locations, the Internet topology will not always appear in regular shapes such as lattices or grids. The term complex network is then used to indicate a network with complicated topology structures that are not as simple as networks of regular shape or uniformly random structure but with a large size of elements and connections [2].

Topology analysis is helpful in allocating limited resources effectively to achieve the best overall outcome of the network. For example, the topology of railway roads may show a transportation railway located in the centre that is crucial for connecting many different places. Then high priority of reconstruction resources should be assigned to ensure the centre railway can function properly in connecting various locations.

Apart from the static topology, there are often dynamic processes conducted upon the network topology, such as the trains running on railways, communication packets transmitting on the Internet and information propagating among social groups. Even if the same rumour is spreading among two different social groups, the outcome of how fast and how widely the rumour is disseminated among each social group can be largely different depending on their topologies. The topology patterns have crucial impacts on shaping different dynamic behaviours and outcomes. Knowledge of the topology can also help to prevent undesired consequences. Particularly during the outbreak of global epidemic diseases, effective and efficient treatment and preventions are crucial in controlling the disease spreading in a population network. During the SARS outbreak in 2003 and the Ebola virus outbreak in December 2013 for example, the lack of topology information as well as spreading predictions had been a significant obstacle to the disease control. Hence, one important question is to study the relationship of a topology and the outcome of a dynamic process.

One important application of network analysis is to facilitate the understanding of healthy brain operations as well as brain disorders such as Parkinson's disease, epilepsy and Alzheimer's disease. The brain is already a very complex system, not mentioning the dynamics of neuronal signals. Research has shown that studies on the brain can benefit from modelling the brain as a network, eliminating intricate details and only focusing on the general properties and dynamics patterns [3]. In recent years, relationships between the anatomical brain network and brain disorders have been found with the help of network analysis [4]. Therefore, studying how the brain topology influences the dynamics is helpful to improve the understanding of how the healthy brain operates. Once the influence is specified for the healthy brain, we can further study the injured brain by making comparisons with the operation mechanism in the healthy brain and eventually improve the treatment effectiveness.

1.2 Motivation

Modern network science mainly explores complex network topology properties and influences on dynamic processes. Typical examples of spreading dynamics on complex networks are rumours spreading on the social network, epidemic diseases spreading among populations and computer viruses propagating on the Internet. The step-by-step spreading details during a dynamic process is usually untraceable due to the large data size of a complex network. However, with the help of the epidemic models, overall outcomes of the spreading process can be studied [5]. For example, after a certain time of spreading how large is the infected population, which parameters lead to a massive epidemic outbreak and the speed of disease spreading. Applying epidemic models to real networks provide rough estimations of the dynamic process. Together with the underlying network topology, interesting properties such as nodes that are in charge of large outbreaks can be mathematically studied. Limited resources could then be applied to the most effective node in order to prevent further large outbreaks, especially in the case of epidemics global outbreak such as the Ebola in 2014.

To study the dynamics outcome influenced by the network topology, various epidemic spreading models [6] are proposed to characterise different types of dynamics that are not limited to epidemic disease spreading. The basic Susceptible-Infected-Susceptible (SIS) [7] model can be used to describe simple spreading processes. Motivated by studying the topological influence on shaping brain dynamics, we apply the SIS epidemic model on an anatomical brain network to simulate the brain dynamics. We especially analyse the influence between pairs of elements by adding a time delay variable to the time series correlation and calculating the so-called delayed correlations. Based on the delayed correlations, we then investigate on the emergence of global patterns.

1.3 Thesis Outline

Following the background on applying network science to study the brain, Chapter 2 provides mathematical notations on network science and useful metrics to analyse network topologies. Chapter 3 illustrates the Susceptible-Infected-Susceptible (SIS) model from the exact Markovian method to the N-Intertwined Mean Field Approximation (NIMFA), covering important concepts of epidemic threshold and meta-stable states. A basic introduction to the continuous-time SIS simulator (SISS) program is also included. Chapter 4 applies the SIS epidemic model to simulate brain dynamics. Simulation results on the brain connectivity patterns are presented in comparison with related results from a previous literature. The delayed correlation is described in Chapter 5 to study the patterns of a dynamic process on the network. Analysis of the simulation results of delayed correlations on a brain network as well as unique properties of the delayed autocorrelation are presented. While Chapter 5 mainly investigate pairwise properties locally, Chapter 6 demonstrates the emergence of a global pattern from the analysis on the delayed correlation. Transfer entropy is implemented to benchmark the simulation results. Finally, in Chapter 7, limitations on our work and conclusions are given, as well as an outlook to future work.

Chapter 2

Network Science

Real systems such as the World-Wide-Web, social networks or biological systems are composed of interacting elements with different local properties that as a whole would form the global feature of the system. In order to examine the global properties of the whole system, network theory depicts a system as a network and simplifies the analysis by ignoring individual micro-differences of each element. This chapter introduces the graph theory and basic metrics to study network properties. The mathematical expressions in this report follow the notations from books [8] and [9]. While network science is a multidisciplinary subject involving mathematics, statistics, engineering and many other fields, here we only introduce metrics that are closely related to our work. For example, Markovian theory and graph spectra are not illustrated here in detail, but comprehensive explanations are provided in books [8] and [9].

2.1 Graph Theory

Traditional network science studies simple and practical problems such as making a round trip through some places without repetition or distinguishing regions on a map using the minimum number of colours as networks. Graph theory provides tools and measurements to study general network properties. Graph theory was first used by mathematician Leonhard Euler to solve the Königsberg 7 bridge problem in 1736 [10]. The problem was to make a round trip through the Prussian city of Königsberg while traversing the seven bridges in the city once and only once. Euler modelled the parts of the city as nodes connected by seven bridge links. Reformulating the problem as an abstract graph consisting of only nodes and links shed light on mathematical analysis method on a set of similar problems.

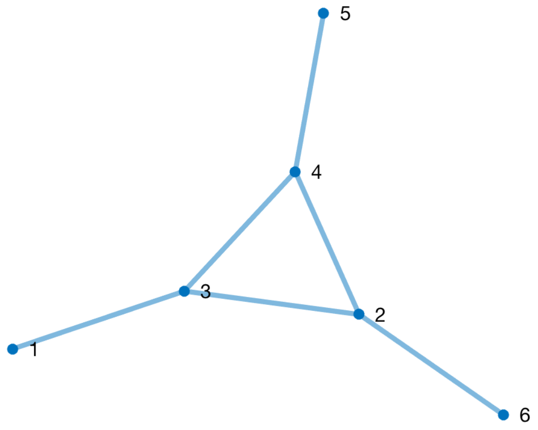


Figure 2.1: An example of an undirected unweighted graph $G(6,6)$.

In graph theory, elements in a system are regarded as nodes of a network, and the interactions of these elements are illustrated as connections of links. A network is described as a graph $G = (N, L)$ where N is the number of nodes in the network, and L is the number of links. A weighted value is assigned to the link to describe different contexts, such as the frequency of communications or the real spatial distance between two nodes. The links can also have directions, depicted as arrows pointing from a source node towards a target node. However, mutual interactions characterised as undirected links are usually used to describe the presence of connections. A general network throughout this thesis consists of only undirected and unweighted links.

There are different network topologies and in Figure 2.2 we show some typical topology patterns: (a) shows an example of a path graph $G(5,4)$, the simplest type of topology where there are only back and forth connections on a line; (b) shows an example of 2-dimension square lattice G(9,12); (c) shows an example of branching tree G(11,10), starting the branching process from a root node and there is no loop; (d) shows an example of star graph G(9, 8). There is a centre node located in the middle, and others are leaf nodes. Every leaf node is connected to the centre node and no links between leaf nodes exist; (e) shows an example of ring graph G(8,8). It is the simplest type of regular graph where a node only connects to the left and right nearest nodes; (f) shows an example of complete graph G(5,10). Every node is connected to every other node, so the number of links equals $L = \frac{N(N-1)}{2}$. The maximum possible links in any graph with node N is equal to the number of links in the complete graph, where the term complete implies a full connection of all possible links.

Adjacency matrix

The adjacency matrix A transforms a graph into matrix notation. An element in the adjacency matrix a_{ij} for undirected networks denotes the presence of a link between two nodes i and j, and

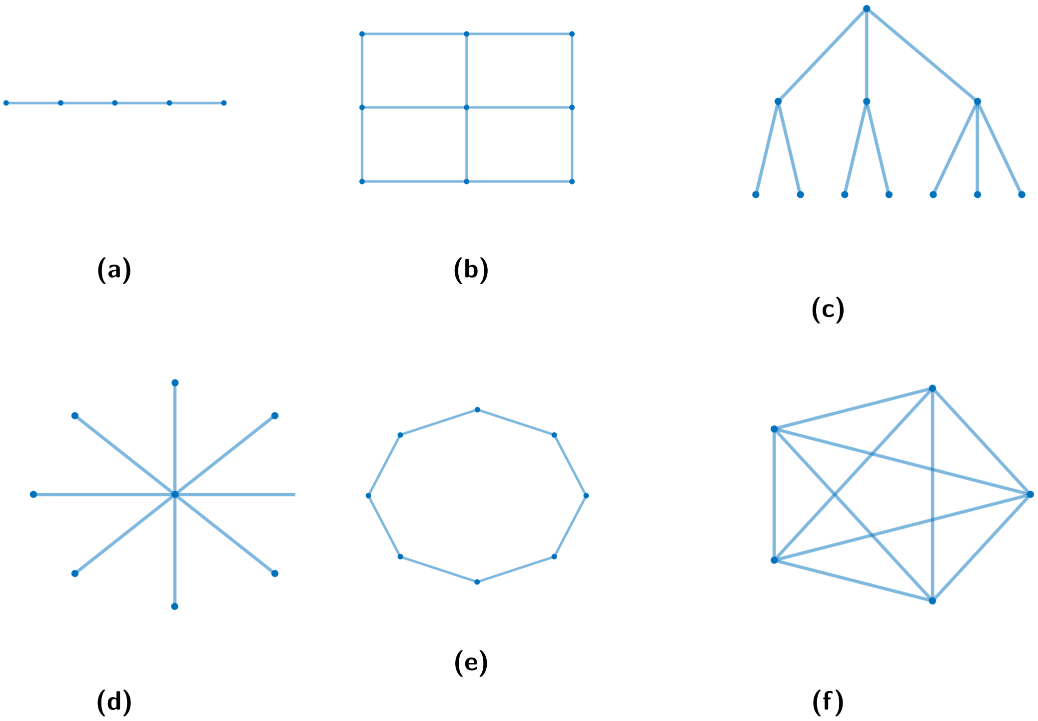


Figure 2.2: Examples of typical topologies. (a) A path graph, (b) a 2D lattice, (c) a branching tree, (d)a star graph, (e) a regular ring graph and (f) complete graph.

is a binary number such that

$$a_{ij} = \begin{cases} 1, & \text{if node } i \text{ is connected to node } j\text{, } j \neq i \\ 0, & \text{otherwise.} \end{cases}$$

Self-loops are not allowed and therefore the diagonal entries a_{ii} are defined as 0. The adjacency matrix of the example shown in Figure 2.1 is given as

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

where the adjacency matrix is symmetric, meaning the network is an undirected network with bi-directional connections between nodes.

Degree

If there is a link between nodes i and j, we consider them as neighbours. The degree of node i is the number of neighbours and is denoted as

$$d_i = \sum_{j=1}^{N} a_{ji},$$

where $0 \leq d_i \leq N-1$ since we have N nodes in the network. The degree of a node is the simplest measure of importance in the network, for example, changes regarding high degree nodes are expected to have a huge impact on the network since more neighbours are influence than the influence by changes of low degree nodes.

Distance metrics

A **walk** of length k from the node i to j includes a succession of k links/hops from the node i to j, which could include loops. The number of k-hop walks can be computed by A^k, where an element $(A^k)_{ij}$ represents the number of k-hop walks from the node i to j. A **path** of length k from the node i to j is walks of length k without loops and the **shortest path** is the path consisting of the smallest number of links/hops from the node i to j. The **hopcount** H_{ij} is used to describe the number of hops in the shortest path from the node i to j.

The **eccentricity** of a node i is the furthest length of shortest paths from node i to all other nodes, which is denoted as

$$\epsilon_i = max_{j \in N}(H_{ij}).$$

The **diameter** of a graph G is the length of the longest shortest path in the graph, i.e. the largest eccentricity. A network with a small diameter can transmit information among nodes more efficiently than a network with a larger diameter. The efficiency can be clearly observed if information is delivered only through the shortest path on a network, where small diameter indicates nodes are near to each other. Deliveries could be faster if then considering the shortest paths.

Clustering Coefficient

For an arbitrary node i in a network, clustering coefficient $c_G(i)$ is defined as the ratio of links between neighbours of node i (denoted as γ) to the maximum possible links between neighbours of node i, which is written as

$$c_G(i) = \frac{\gamma}{d_i(d_i-1)/2}.$$

The maximum links between the d_i neighbours of node i is equal to the links of a complete graph with node number $N = d_i$. For the centre node in a star graph (see Figure 2.2(d)), the clustering coefficient equals 0 since there are no connections between leaf nodes. So the clustering coefficient measures how often neighbours of a node form a triangle with this node as clusters or small groups. The (global) clustering coefficient of a graph G is denoted as the average of the clustering coefficient of individual nodes

$$c_G = \frac{1}{N}\sum_{v=1}^{N} c_G(v).$$

Closeness

The closeness of a node i measures the reciprocal of the sum of shortest hopcount to every other nodes and is denoted as

$$C_i = \frac{1}{\sum_{j=1,j\neq i}^{N} H_{ij}}.$$

In a network with a low diameter (or high clustering coefficient), there are small hopcounts between pairs of nodes and the reciprocal results in a high closeness.

Eigenvalue and eigenvectors

The eigenvalues of adjacency matrix A are a set of scalars $\lambda = \lambda_1, \lambda_2, ...$ while the corresponding eigenvectors are represented in columns $x = [x_1, x_2, ..., x_k, ...x_N]^T$. One eigenvalue scalar λ_k is corresponded with one eigenvector column x_k and they satisfy the equation

$$Ax = \lambda x.$$

A symmetric adjacency matrix produces real eigenvalues which are often sorted. In a descending sort sequence, the largest eigenvalue is denoted as λ_1, which is also called the spectra radius of the graph. The principal eigenvector x_1 is the one corresponding to the largest eigenvalue λ_1 which satisfies the equation

$$Ax_1 = \lambda_1 x_1. \tag{2.1}$$

The largest eigenvalue is also an important parameter to estimate the lower bound of an epidemic threshold [5].

The basic eigenvalue equation regarding the largest eigenvalue λ_1 (Equation 2.1) for a node i can be written as [9, Equation 1.3]

$$\lambda_1 (x_1)_i = (Ax_k)_i = \sum_{j=1}^{N} a_{ij} (x_1)_j, \tag{2.2}$$

where the summation in the last term is taken on neighbours of node i since $a_{ij} = 1$ only when node i and j are connected and sum up eigenvector centralities of node i neighbours.

Then, the eigenvector centrality of node i regarding the principal eigenvector x_1, denoted as x_{1i}, is defined as the i^{th} component of the principal eigenvector x_1 belonging to the largest eigenvalue λ_1. The eigenvector centrality x_{1i} can be computed from Equation 2.2 by diving the λ_1 in all terms, which gives

$$(x_1)_i = \frac{(Ax_1)_i}{\lambda_1} = \frac{1}{\lambda_1} \sum_{j=1}^{N} a_{ij} (x_1)_j. \tag{2.3}$$

In this way, the eigenvector centrality provides another ranking measure of the node's importance in a network which can be interpreted as 'weighted degrees' [9]

2.2 Network Models

The network science started with analysis on simple networks of small size that could be easily computed and all nodes or links have similar properties, such as a regular or purely random topology. With the fast technology development in the 20th century, many real networks such as the Internet, power grids, telephony infrastructures and transportation networks are evolving rapidly to immense size with complex connections. Analysis of these networks often shows complex properties that do not occur in simple networks: enormous network size that is infeasible to do computation on the whole network, irregular and non-homogeneous properties for nodes and links and even dynamically evolving topologies. To distinguish from simple networks and highlight the complexity, networks of very large size and complex connections are categorised as complex networks [11]. Empirical studies on complex networks have found two essential features. One type of complex feature is the small-world property found by Watts and Strogatz in 1998 [12] that on average, arbitrary node pairs can be reached in relatively short paths of six hops. Another type is the scale-free property found by Barabási and Albert in 1999 [13] that many large size networks show a power law degree distribution. In the following paragraphs, we give a brief illustration on two simple networks: the regular and purely random network, and two complex models: the small-world and scale-free network.

Regular network

In a regular network all nodes have the same degree r, which means every node is connected to r-nodes. An example is shown in Figure 2.3(a) on a graph with 20 nodes oriented on a circle. Every node has 4 neighbours: two connections to the nearest nodes from left and right, and two connections to the nodes separated by 3 other nodes on the circle both from left and right.

Random network

In a random network, the probability of a link existing between any pair of nodes is p. Probability $p = 0$ indicates all nodes are independent and there is no connection in the network, while $p = 1$ results in a complete graph, as shown in Figure 2.3(c).

Small-world network

An important complex network model is the small-world network proposed by Watts and Strogatz in 1998 [12]. As shown in the paper, the small-world network can be generated by randomly rewiring the links in the regular graph by a certain value of randomness factor. Therefore, small-world properties are often observed in between two extremes, the regular network and the random network. The small-world network is named after the discovery of 'six-degree separation' in the social network, indicating that the world is small since we can reach any strangers within an averaged small number of intermediate friends [14]. The short distance between overall node pairs also holds for large size small-world networks while a small network with only 20 nodes is shown as an illustration.

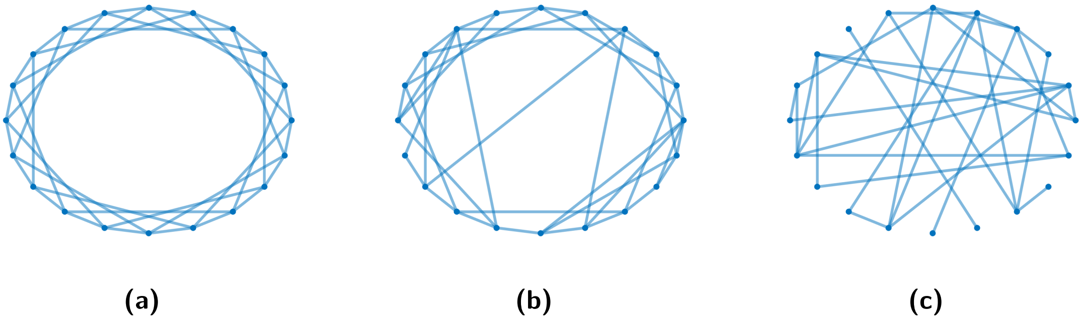


Figure 2.3: An illustration of (a) a regular network, (b) small-world network and (d) random network. The small-world network is generated from the regular network with a randomness factor 0.1. In three cases, N=20 and nodes are fixed on a circle to show changes in connections.

Scale-free network

In a large size network, individual nodes may be hard to show the whole picture of the network and extensive details of nodes are difficult to collect. A global network property that scientists usually use to characterise large size network is the degree distribution, denoted as $Pr[d_i = k]$. The degree distribution measures the probability of a node i having degree k, or the fraction of nodes having degree k in the network. The Faloutsos brothers found the Internet and many other real networks follow a power-law distribution [15] and is denoted as

$$Pr[d_i = k] \propto k^{-\alpha}, \alpha \in (2.2, 2.5),$$

which is also called a scale-free network. A famous generation of a power-law network is using preferential attachment proposed by Barabási and Albert in 1999 [13] : starting with some initial connected nodes m_0, new nodes are added to the network one by one. Each new node is connected to the existing m nodes ($m \leq m_0$) with a probability proportional to the node's current degree. This method results in a 'preference' of new nodes connecting to high degree nodes, shown as an example in Figure C.

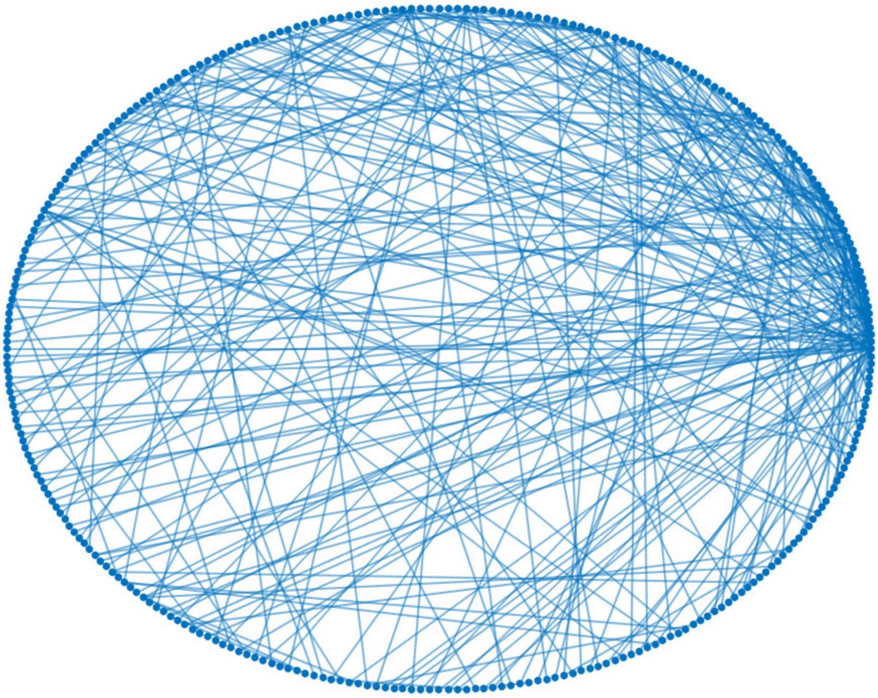


Figure 2.4: A power-law distributed network, with $\alpha = 2.3$ and $N = 300$. All nodes are drawn on a circle. The nodes located at top right corner is highly densely connected.

Chapter 3

Epidemic Spreading Models

Epidemic models transform conceptual understanding of dynamic processes in networks into computable results. In order to expose detail properties of the dynamic process in a network, simple models are preferred so that desired properties would not be concealed in the complexity of the model. To start with, general concept of the Susceptible-Infected-Susceptible (SIS) and Susceptible-Infected-Removed (SIR) models are introduced, including the definition of meta-stable regime which will be used in later chapters. Then equations for solving the 2^N-state Markov model and N-Intertwined Mean-Field Approximation (NIMFA) method will be presented.

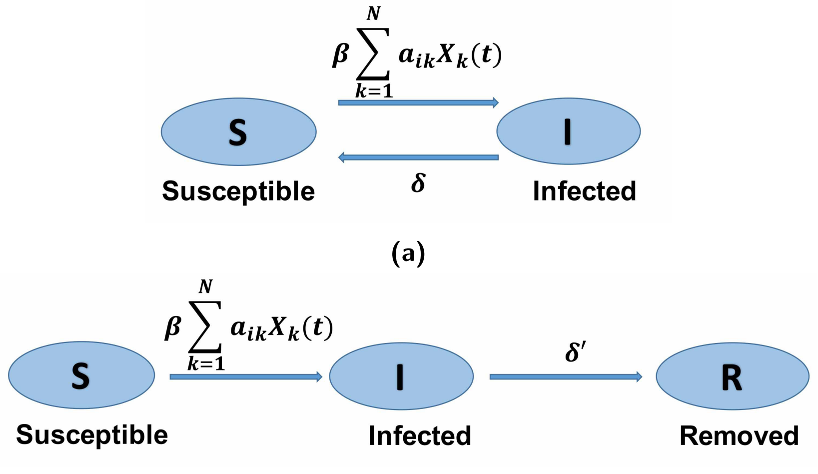
3.1 The SIS model

An epidemic disease spreading on the links of a network yields two basic states of a node: [16]

- Susceptible status S, if a node is currently healthy and may become infected if in contact with infected nodes;
- Infected status I, if a node is infected and become contagious to other healthy nodes.

as shown in the schematic graph in Figure 3.1(a). The status of a node i at time t is often described as a Bernoulli random variable $X_i(t) \in \{0, 1\}$, where 0 represents the healthy and susceptible status, and 1 represents the infected status [8].

A healthy node can become infected by interacting with infected neighbours. Once a node gets infected, the curing process is automatically triggered. The curing process is modelled as a Poisson process of curing rate δ which is independent of the spreading process. The average curing time



(a)

(b)

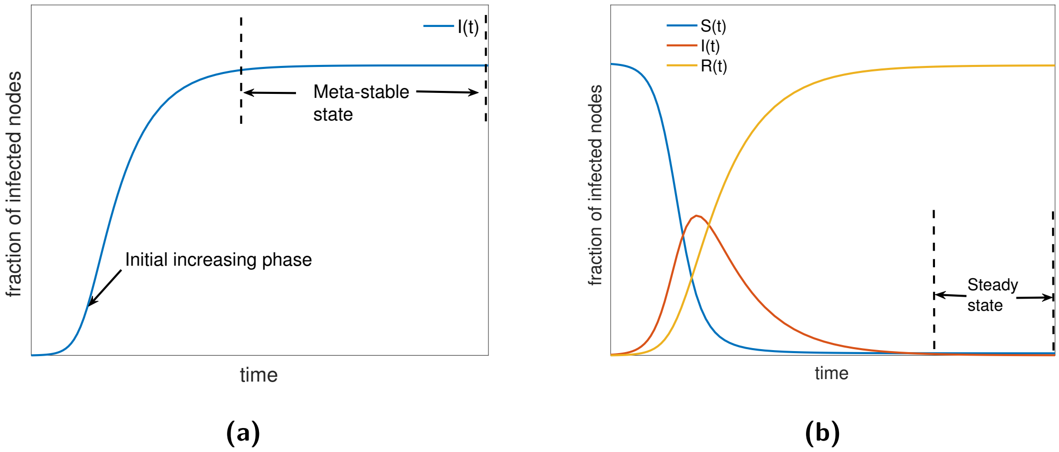
Figure 3.1: Schematic graph of (a) SIS and (b) SIR model. In both graphs, the term $\sum_{k=1}^{N} a_{ki} X_k(t)$ is the number of infected neighbours of a node, β is the spreading rate on a link. In the SIS model, δ is the curing rate while in the SIR model δ' is used to distinguish the curing rate from that in the SIS model.

is measured as $1/\delta$ unit time. An infected node can spread to neighbouring nodes before being cured. In the spreading or infection process, an infected node spreads the disease to neighbouring nodes through their connections. The spreading on each link is independent of the spreading on other links and is modelled as a Poisson process with a spreading rate β.

The average spreading time on a link is then measured as $1/\beta$ unit time. The number of infected neighbours of node i is computed as $\sum_{k=1}^{N} a_{ki} X_k(t)$, where $a_{ki} = 1$ if node k is a neighbour of node i and $X_k(t) = 1$ if node k is infected. The infection rate imposed on a node i will be the sum of rates from all infected neighbours $\beta \sum_{k=1}^{N} a_{ki} X_k(t)$. This summation is due to the linear feature of Poisson processes that the sum of two Poisson processes is still a Poisson process, and the new rate is the sum of two original rates. Since the infection rate is proportional to the number of infected neighbours, a node with more neighbours generally is infected more frequently.

The effective spreading rate in a network is defined as $\tau = \beta/\delta$, which influences the average fraction of infected nodes in a long term. A high value of τ indicates a large epidemic outbreak and large average fractions of infected nodes can be observed. On the contrary, the epidemic dies out exponentially fast. The critical epidemic threshold (τ_c) is usually used to quantify where a spreading process changes from the early die-out phase to the massive infection phase. An approximation method for the epidemic threshold is introduced later.

Figure 3.2(a) shows the fractions of infected nodes of an example SIS model with high effective spreading rate. Initialised by a small number of infected nodes, the fractions of infected nodes increases exponentially fast during the initial phase of an epidemic outbreak. Then a stable level of infected fractions is observed for a long time and the system is in a relatively stationary phase,



(a) **(b)**

Figure 3.2: Fraction of infected nodes of (a) SIS and (b) SIR model.

which is often characterised as the meta-stable state. In this thesis where we only consider the SIS model, we define the meta-stable state as the later half of simulation time for simplicity. However, after a very long time, there exists a steady state where all nodes are healthy (disregarding self-infections). The steady state of a SIS model is only possible to observe after a very long time around the scale of 10^7 unit times [17] and therefore, the meta-stable state is usually observed due to feasibility.

An extensive review of various epidemic models can be found in [16]. A short illustration of the Susceptible-Infected-Removed (SIR) model is presented in comparison of the schematic graph (Figure 3.1) and the infected fractions (Figure 3.2) from those in the SIS model. The simple SIR model removes the cured nodes from the network so that they would not be involved in the infection again and as a result, the network reaches the steady state more quickly than the SIS model. The steady state in the SIR model can be easily observed within a period of time that is achievable both in simulations and real experiments. In the case of removed nodes joins the network and infection process again, a Susceptible-Infected-Removed-Susceptible model is usually used. More complicated scenarios could be considered by other models derived from the basic SIS and SIR models. To name a few, the Susceptible-Infected-Recovered-Susceptible (SIRS) model [18] characterises newborn susceptible individuals coming into a population (recovered R to susceptible S process) and the Susceptible-Exposed-Infected-Removed (SEIR) model [19] describes diseases with an incubation process (the exposed E state) before symptoms appear.

We only apply the SIS model in this thesis work to remove exhaustive details and reduce computation complexity, so as to provide a clear view on the global patterns of the whole network in later chapters. Therefore, next section only illustrates the mathematical analysis based on the SIS model which is used throughout this thesis work.

3.2 The 2^N-state Markov Chain

Node states in the SIS model are shown in the schematic graph (Figure 3.1). The infection and curing processes are both Poisson processes independent of history status, which can be studied as a two-state continuous Markov process. The infinitesimal generator describing transition rates between different states is given as [8]

$$Q_i(t) = \begin{bmatrix} -q_i(t) & q_i(t) \\ \delta & -\delta \end{bmatrix}, \tag{3.1}$$

and the term $q_i(t)$ describes the transition rate from a 0-state to 1-state, denoted as

$$q_i(t) = \beta \sum_{k=1}^{N} a_{ki} X_k(t). \tag{3.2}$$

A network of N nodes with each node having 2 states, yields a 2^N-state Markov Chain. For example, a 3 node network consists of 2^3 combinations of node states and the 8-state Markov Chain is illustrated in the state diagram shown in Figure 3.3.

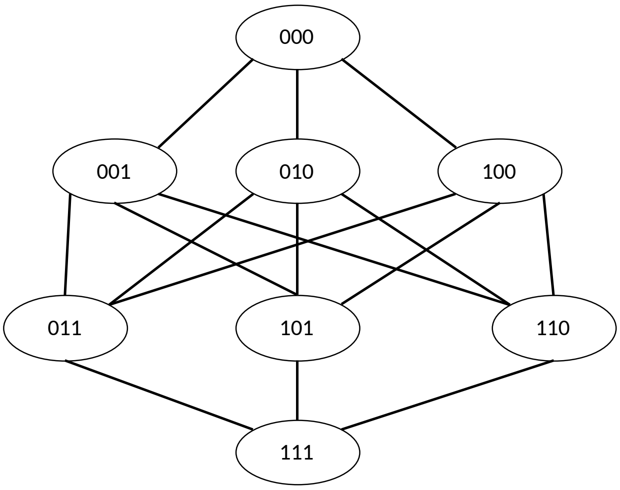


Figure 3.3: 2^N-state Markov state transition diagram with $N = 3$.

In the general 2^N-state Markov model, the term $q_i(t)$ in Equation (3.2) is a random variable due to the Bernoulli term $X_k(t)$. While there should be no random variables in the Markov theory, the mean-field approximation is applied to average over all possible cases

$$E[q_i(t)] = \beta E[\sum_{k=1}^{N} a_{ki} X_k(t)]. \tag{3.3}$$

For Bernoulli random variables, the expectation is numerically equal to the probability of the state being equal to 1 so we have

$$E[X_k(t)] = Pr[X_k(t) = 1]. \tag{3.4}$$

In Equation (3.3) there are 4 combinations of states of a_{ki} and $X_k(t)$: 00, 01, 10 and 11. The total combinations for all nodes k of each node i will be 4^{N^N}, which is a large number of equations to solve increasing exponentially with network size of node N. Therefore, the 2^N-state Markov model is often infeasible to solve.

3.3 N-Intertwined Mean-Field Approximation (NIMFA)

The N-Intertwined Mean-Field Approximation (NIMFA) [5] method provides a feasible solution to the SIS model. The governing equation describing the SIS spreading process is defined in the form of a differential equation

$$\frac{dX_i(t)}{dt} = -\delta X_i(t) + (1 - X_i(t))\,\beta \sum_{k=1}^{N} a_{ki} X_k(t), \tag{3.5}$$

and the mean value is computed as

$$\begin{aligned}\frac{dE[X_i(t)]}{dt} &= -\delta E[X_i(t)] + (1 - E[X_i(t)])\,\beta \sum_{k=1}^{N} a_{ki} E[X_k(t)] \\ &= -\delta E[X_i(t)] + \sum_{k=1}^{N} a_{ki} \{E[X_k(t)] - E[X_i(t)]E[X_i(t)]\}.\end{aligned} \tag{3.6}$$

The NIMFA approximation reduces the complexity of calculating 2^N-state joint probabilities of the term $E[X_i(t)X_k(t)] = Pr[X_i(t) = 1, X_k(t) = 1]$ [5] by assuming independence

$$E[X_i(t)X_j(t)] = E[X_i(t)]E[X_k(t)] \tag{3.7}$$
$$= Pr[X_i(t) = 1]Pr[X_j(t) = 1]. \tag{3.8}$$

Denoting the expected value $E[X_i(t)$ as $v_i(t)$, Equation (3.6) becomes

$$\frac{dv_i(t)}{dt} = -\delta v_i(t) + (1 - v_i(t))\beta \sum_{k=1}^{N} a_{ki} v_i(t), \tag{3.9}$$

and the matrix form is

$$\frac{dV(t)}{dt} = \beta AV(t) - diag(v_i(t))\,(\beta AV(t) + \delta u), \tag{3.10}$$

where vector $V(t) = [v_1(t) \quad v_2(t) \quad \ldots \quad v_N(t)]^T$, $diag(v_i(t))$ is the diagonal matrix with elements $v_1(t), v_2(t), \ldots v_N(t)$, and u is a all-one vector.

Lower-bound on the critical epidemic threshold

The NIMFA equation provides an estimation on the critical epidemic threshold τ_c by solving the vector form differential Equation (3.10) [5] as

$$V(t) \leq e^{(-\delta I + \beta A)t} V(0), \tag{3.11}$$

where vector element $v_i(t)$ tends to be zero exponentially fast if eigenvalues of $-\delta I + \beta A$ (depending on eigenvalues of the adjacency matrix$A$) are negative. And since the eigenvalues are assumed to be arranging in descending order, the largest eigenvalue $\lambda_1(A)$ being negative can guarantee all eigenvalues for $-\delta I + \beta A$ are negative. Therefore if we have

$$-\delta + \beta\lambda_1(A) < 0, \tag{3.12}$$

then we can obtain

$$\tau_c > \frac{1}{\lambda_1(A)} = \tau_c^{(1)}, \tag{3.13}$$

where the first-order NIMFA threshold is denoted as $\tau_c^{(1)}$. The NIMFA provides a lower-bound $\tau_c^{(1)}$ of the critical epidemic threshold τ_c.

3.4 SIS Simulations

The SIS epidemic spreading model is applied throughout this thesis. The underlying simulation tool is the continuous-time SIS Simulator (SISS) program provided by the work in[20]. The SISS is a java program which takes the network adjacency matrix as an input and runs the SIS epidemic spreading process on top of the network. The program then outputs the binary time series of node states $X_i(t)$ of all nodes for the whole simulated time periods (see Figure 3.4).

Various other simulations are also feasible with the help of the SISS. Apart from the performed SIS simulation function, the SISS contains many more functions of other models and parameters that are not used in this thesis but could be utilised in future work. Here we only explain the outline of the SIS simulation procedure which is used in this thesis.

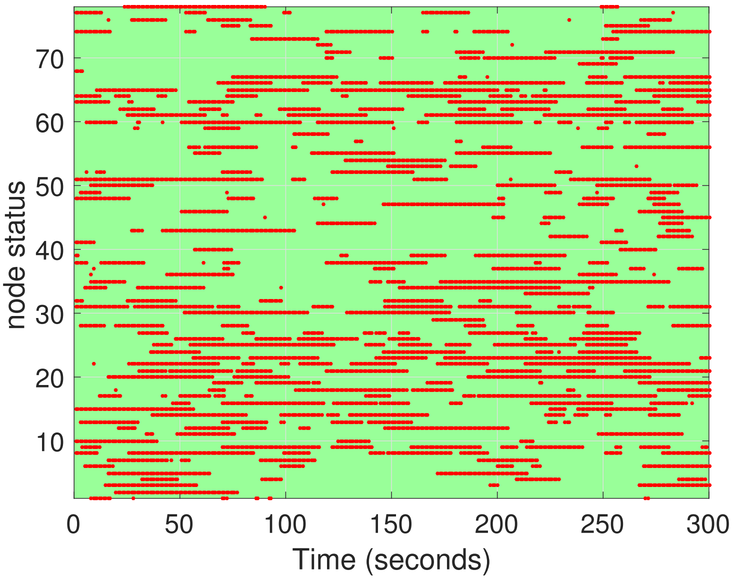


Figure 3.4: An example of output time series from the SISS. The red colour indicates an infected state and green colour indicates the healthy state. The unit time from the simulator is defined as seconds in this thesis work. The underlying network is a 78-node anatomical brain network (described in Chapter 4). The SIS epidemic spreading is initialised by 15 random infected nodes. The effective spreading rate is $\tau = 0.2$ while the lower bound of the critical epidemic threshold is $\tau_c^{(1)} = 0.01$.

To simulate the continuous time feature, the simulator interpolates event tickets onto a timeline. The SIS model has two type of events, the spreading and curing event illustrating the spreading and curing processes respectively. An event is represented in the form of a ticket with time information and the indicated node index. An infected node always generates a ticket for the curing event with an exponential time of rate δ, simulating the curing process. At the end of the curing time, or in other words the node is cured, the node state converts from infected to susceptible. An infected node also generates a spreading event ticket to simulate the infection spreading process of an exponential time of rate β. The spreading event imposes infection on healthy neighbouring nodes. An infected node can only generate spreading event tickets before the curing time is reached.

In the example of the outputs, binary time series, is shown in Figure 3.4. The underlying topology network is a 78-node structural brain network whose properties are described in Chapter 4. Time series describe the simulated dynamics regarding node states at every time unit. Although the SISS conducts continuous-time simulations, time series are sampled for the discrete data process using Matlab. A high sampling rate (or more samples per unit time) provides the sampled data closer to the continuous real data with little losses. However, in our study of SIS dynamics, a moderate sampling rate is of sufficient accuracy since the time a node remains at current status is a Poisson distributed time and it is not a transient time. Our analysis does not focus on the precise time when the node changes between S and I states, and high sampling rate will produce bulk size data which requires much more data processing time. It is noteworthy that the simulated time has the same unit as the Poisson rates δ and β. For simplicity, we take 'seconds' as our time unit in this thesis and later analysis on dynamic processes especially in Chapter 5 and 6 are based on these time series.

Real networks often work around a critical threshold where the system is in a transition phase between massive and scarce infections [21]. In order to find the critical spreading rate τ for a network, we average the fraction of infected nodes at meta-stable state for multiple simulation realisations so as to remove extreme coincidence. The multiple simulation realisations are dynamics of the same spreading rate β and curing rate δ conducting on the same underlying network topology. Figure 3.5 shows the simulation results in finding the critical spreading rate τ_c. The critical threshold is defined at averaged 1% of infected nodes at the meta-stable state is achieved [21]. In the example shown, the critical $\beta = 0.09$ with a fixed $\delta = 0.5$, which yields the critical spreading rate $\tau_c = 0.18$.

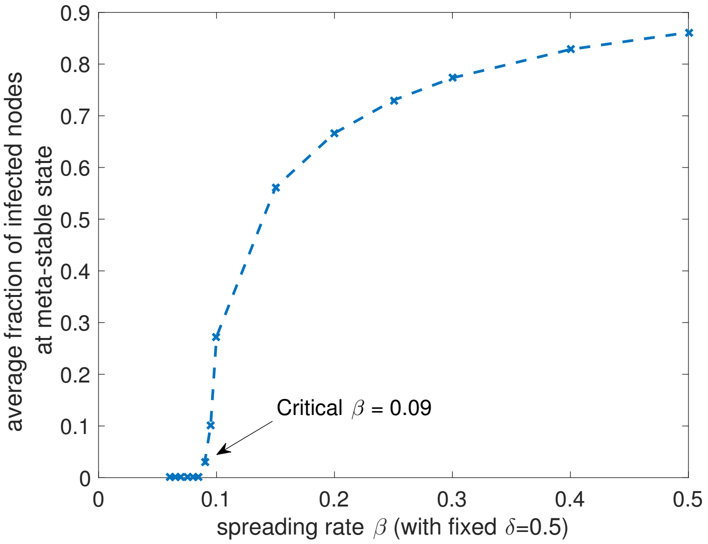


Figure 3.5: Effective spreading rate τ. The critical τ is usually defined at 1% nodes activated or infected in the meta-stable state. The figure is averaged over 10 simulation realisations on the 78-node structural network (see Chapter 4).

Chapter 4

Brain Networks

The concern of this thesis is constrained to the human brain, the most complicated case among all different species. In the rest of the thesis, the brain refers to the human brain if not specified otherwise. This chapter justifies a simple model can be used to analyse general patterns of brain connectivities.

4.1 Network Modelling of the Brain

The brain is the most important organ in a body but also the most complex system of human knowledge and therefore studying how the brain operates is of highly research interest. The brain performs centralised control over the body by electrical communications between 10^{11} neurones located on the surface of the brain [22]. The surface structure is called the cerebral cortex and consists of about 10^{15} neuronal connections upon which electrical signals can transmit. The cerebral cortex, therefore, plays a critical role in controlling the behaviour of an individual, such as in emotions, thinking, intelligence, and movements.

Study on the cerebral cortex has been limited by data acquisition that only animal brains were measured in the past due to the harm on subjects of invasive chemical tracing. With the development of neuroimaging techniques, non-invasive methods can be applied to gather human brain data. Brain functions which co-activate different brain regions can be measured by functional magnetic resonance imaging (fMRI), revealing the actual functioning on the underlying anatomical connections. The underlying anatomical structure can be measured by diffusion MRI and a special form of (structure) MRI, called diffusion tensor imaging (DTI). The electroencephalography (EEG) measure the electric flow between neurones and magneto-encephalography (MEG) measures the magnetic fields generated by the electrical currents of neurone activities.

Various data acquisition technologies enhanced the study of the brain. With the help of a combination of non-invasive technologies, many projects, such as the Human Connectome Project [23] and Human Brain Mapping Project [24], are working on a comprehensive mapping of the brain. However, due to limitations in massive data storage and computation, it is still difficult to analyse the brain as a whole. As a consequence, brain data needs to be integrated to a smaller scale in order to perform data processing and there is always a trade-off between comprehensive details and feasibility.

To gain an overall view of the brain, neuroscientists integrate groups of neurones into a small number of regions according to the anatomical structure and functions. Two popular integration levels are:

- Voxel-level: A neuroimaging analysis technique is to statistically average a million neurones in a local area to be a voxel. A voxel can be understood as an element of 3-dimensional space of an object, which is similar to the 2-dimensional pixel element of an image. The voxel-level analysis projects the brain as a whole network while also provides detailed properties of the brain [25].
- Region of Interests (ROIs): The brain can be divided into macro-scale cortical regions according to anatomical architectures and cell organisations. The underlying brain network used throughout this thesis work represents connections between 78 cortical regions of some high-level functions[26]. The nodes representing major anatomical regions are defined by the Automated Anatomical Labeling (AAL) [27]. Measured by the diffusion tensor imaging (DTI) tractography, major connections observed in all 80 healthy subjects are mapped as links to form a network of macroscopic anatomical regions. As a result, the anatomical brain network is also called the structural network.

The complex topology results in even more complicated dynamics that is hard to analyse. However, research has found simple models that do not impose extra complexity from the model often performs better to feature backbone properties and global patterns [28]. A simple SIS model has been applied to simulate the dynamics on the brain network in discrete-time and connectivity patterns are found analogous to the patterns from empirical data [21].

Following the same purpose of getting a rough but overall view of the brain, we apply the simple SIS model with a more realistic continuous-time simulation performed. The SIS model provides a good match of the two neurone states, activated to signalling and excitable, to the infected and susceptible states. A neurone that is not signalling information is at the excitable state which can be interpreted as the susceptible state. A neurone that is activated and sending electrical signals can be mapped as a node at infected state. An activated neurone sending signals is considered as the spreading process, and after some time the neurone restores to excitable state which is considered as the curing process.

However, one feature of the SIS model, however, needs to be carefully considered to confine the modelling to actual brain dynamics. We choose the critical spreading rate due to the fact that the brain is believed to be working around the critical threshold [29]. While a spreading process around the epidemic threshold is in between of highly possible outbreak to the whole network and rapidly dies out, the simulation realisations contain many rapid die out dynamics. On the contrary, the brain dynamics never stops until the brain dies. The brain is considered to be working in the meta-stable state. Therefore, in all of the following analysis steps, the rapid die out dynamic realisations from the simulator should be removed, and only perform analyses on outbreak realisations consisting of a long meta-stable state.

4.2 Connectivity Patterns

The structural brain topology largely influences the dynamics outcome since spreading process is conducted on the connection links. There are two major connectivity patterns often observed from the dynamics on the structural network, the functional connectivity and the effective connectivity. Here we present analysis and comparisons of the connectivity patterns.

Structural Network

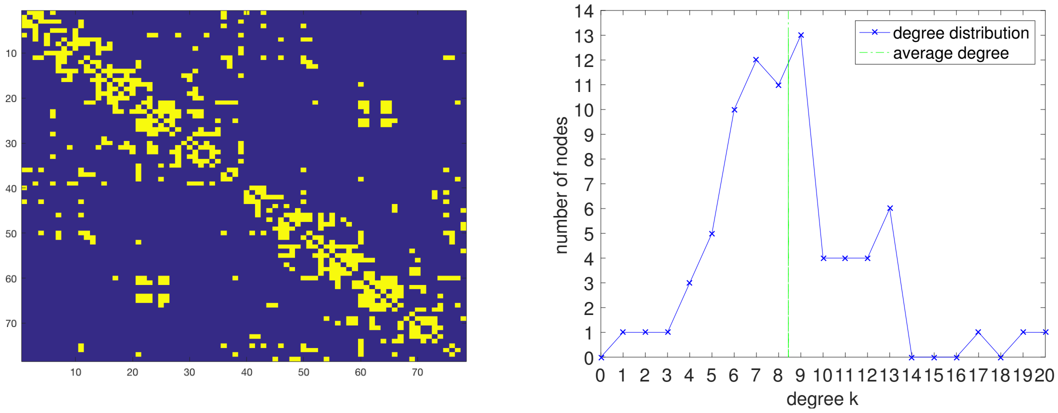
The underlying anatomical brain network describing the regional connections upon which brain dynamics and communications occur is often referred to as the structural network. The link weight is rather referred to as the structural connectivity. As mentioned in the previous section, the anatomical network under the AAL atlas will be taken as the underlying structural network throughout the whole work. A mapping of the AAL regions to the 78 nodes is provided in Appendix A and the adjacency matrix of the structural network is shown in Figure 4.1(a).

The structural network contains 78 nodes and 329 links. The degree distribution of this structural brain network is shown in Figure 4.1(b) and the average degree is $E[D] = 8.43$. The eccentricity of the nodes ranges between 4 and 6, and the diameter of the network is 6. Since the diameter is rather small, efficient spreading can be anticipated. Considering the highest degrees, there are 3 hubs in the structural network: node 21, 23 and 60. From the network science perspective, the anatomic brain network also shows the small-world and high clustering properties [30, 31] and thus falls into the complex network category as explained in Chapter 3. The NIMFA lower bound for the epidemic spreading threshold is $\tau_c^{(1)} = \frac{1}{\lambda_1} = 0.095$ with the largest eigenvalue $\lambda_1 = 10.47$.

Simulation parameters

For the rest of the thesis, the following simulation parameters are taken as default if not specified otherwise:

- Underlying network: the 78-node anatomical brain network, also named as the structural network interchangeably in this thesis;



(a) **(b)**

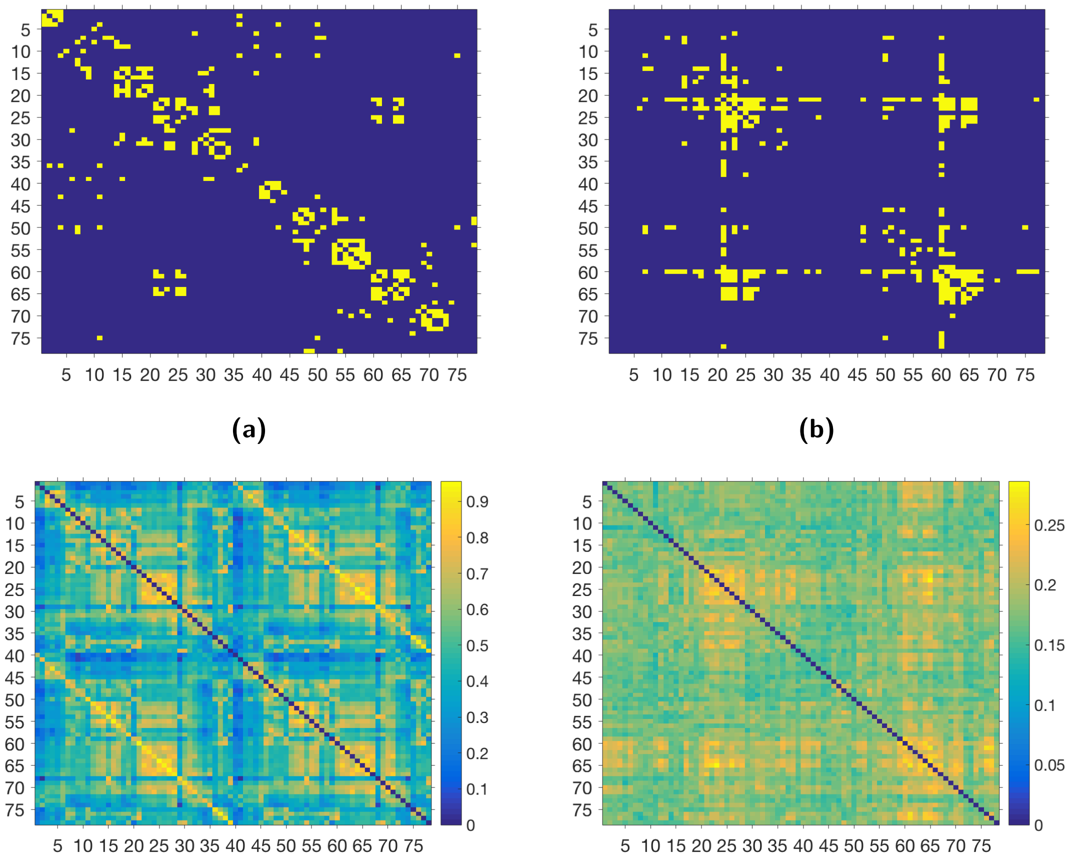
Figure 4.1: (a) The adjacency matrix of the structural network, describing the connections between AAL regions. The yellow colour represents a connection, i.e. $a_{ij} = 1$. (b) Structural network degree distribution with average degree 8.43.

- The network consists of nodes $N = 78$ and links $L = 329$;
- The average degree of the network is $E[D] = 8.43$;
- The largest eigenvalue is $\lambda_1 = 10.47$;
- The NIMFA lower-bound on the critical epidemic threshold is $\tau_c^{(1)} = \frac{1}{\lambda_1} = 0.095$;
- The SIS epidemic spreading process is initialised by 15 infected nodes which are randomly chosen in every simulation realisations;
- The curing rate is fixed as $\delta = 0.5$, and the default spreading rate is $\beta = 0.1$, resulting in an effective spreading rate $\tau = 0.2$ which is slightly above the lNIMFA lower-bound threshold;
- The default simulation time period is 4096 unit time, and the default unit time is taken as seconds;
- The simulated results are sampled as rate 0.1, meaning taking 10 samples per second and one sample represents 0.1 seconds;
- The delayed correlations (Chapter 5) is computed in the meta-stable state which is defined as the range from $t = 2048s$ to $4096s$, the later half of the simulation time;
- Simulation results such as the delayed correlations are averaged over 100 realisations.

Functional Connectivity

Functional connectivity are often measured as macro-scale regional dynamics in the brain and is defined as the correlation between regional communications. The dynamic process is modelled as the SIS spreading process, thus the functional connectivity is the pairwise correlation of the SIS time series. Details about correlation are explained in Chapter 5.

Figure 4.2(a) shows the functional connectivity by calculating the correlations of the SIS time series, averaged over 100 runs and applied supra-threshold [explain] to include the same number of links as that in the structural network. The pattern is similar to the structural connectivity matrix.



(a) **(b)**

(c) **(d)**

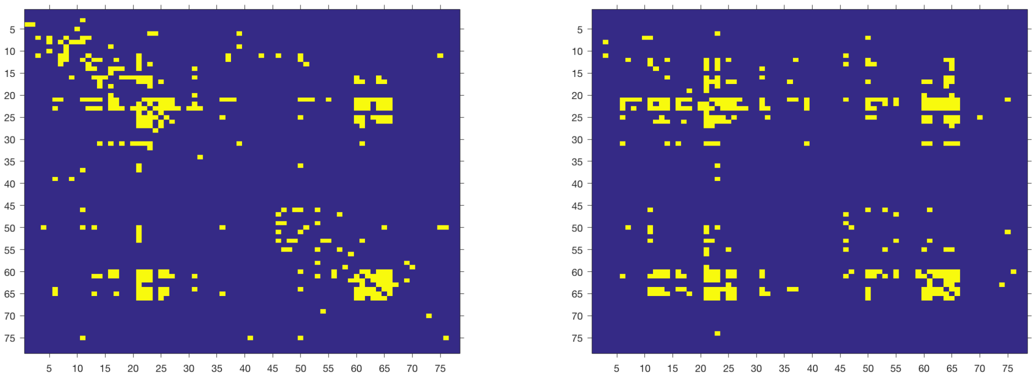
Figure 4.2: (a) Functional connectivity based on SIS simulations. (b) The effective connectivity, which measures the influence of node i to j in next time unit. Both (a) and (b) matrices are made binary by supra-thresholding so that there are the same number of links ($L = 329$) as that in the structural network. (c) Empirical functional connectivity, measured by fMRI and (d) MEG.

Figure 4.2(c) and (d) shows two empirical function connectivity patterns measured in fMRI and MEG from healthy controls. The pattern in fMRI has two slashes around the diagonal, while the SIS model correlations reveal the general pattern of links around the diagonal and two 'blobes' on the off-diagonal direction.

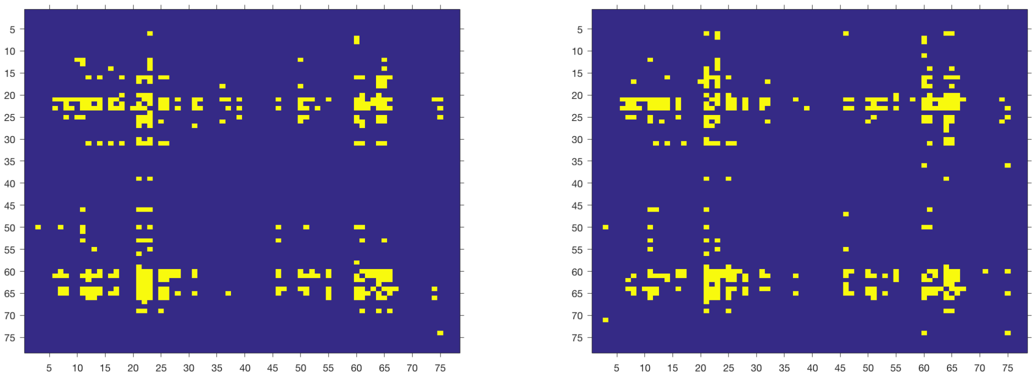
A comparison of correlations under different time lag h is performed in Figure 4.3. Note that this is the matrix presentation of the delayed correlations in Chapter 5. From the matrix presentation, the presence of links does not change significantly in different time delays.

Effective Connectivity

The effective connectivity is often taken to measure the influence of the network dynamics. Sim-



(a) Functional connectivity of time lag 1.0s **(b)** Functional connectivity of time lag 4.0s



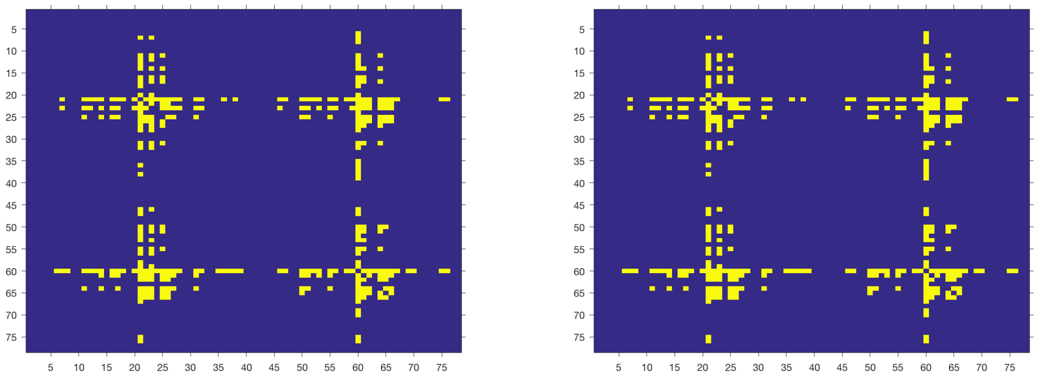
(c) Functional connectivity of time lag 8.0s **(d)** Functional connectivity of time lag 10.0s

Figure 4.3: Delayed correlations (see Chapter 5) displayed in the same way as the functional connectivity. The delayed correlation also applied supra-threshold and displayed in as a binary matrix for comparison with structural brain network. The yellow colour represents a link in the matrix.

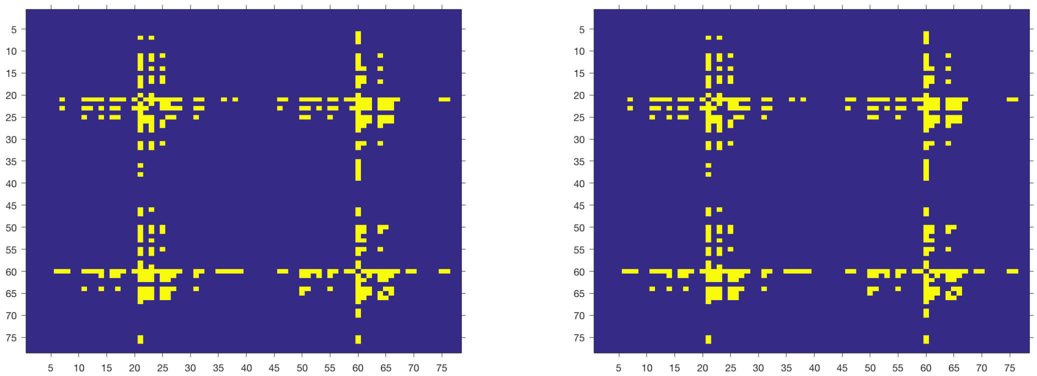
ilar to the definition and calculations from [21], we define the effective connectivity C_{eff} of the dynamics conducting on the structural brain network at delay $h = 1$ as

$$C_{eff} = \frac{Pr[X_j(t+h) = 1 | X_i(t) = 1] + Pr[X_i(t+h) = 1 | X_j(t) = 1]}{2}, \tag{4.1}$$

and the idea is to remove directionality in the time lag and remain mutual connections to compare with the undirected structural network.The effective connectivity is shown in Figure 4.2(b) in matrix presentation, which measures the probability of node i is infected before and following with node j is infected. The effective connectivity showed four 'blobles' of communication flows in one second.



(a) Effective connectivity of time lag 2.0s **(b)** Effective connectivity of time lag 4.0s



(c) Effective connectivity of time lag 8.0s **(d)** Effective connectivity of time lag 10.0s

Figure 4.4: Effective connectivity with different delays. The effective connectivity matrix with different delays also applied supra-threshold and displayed the same number of links as that in the underlying network. The yellow colour represents a link in the matrix.

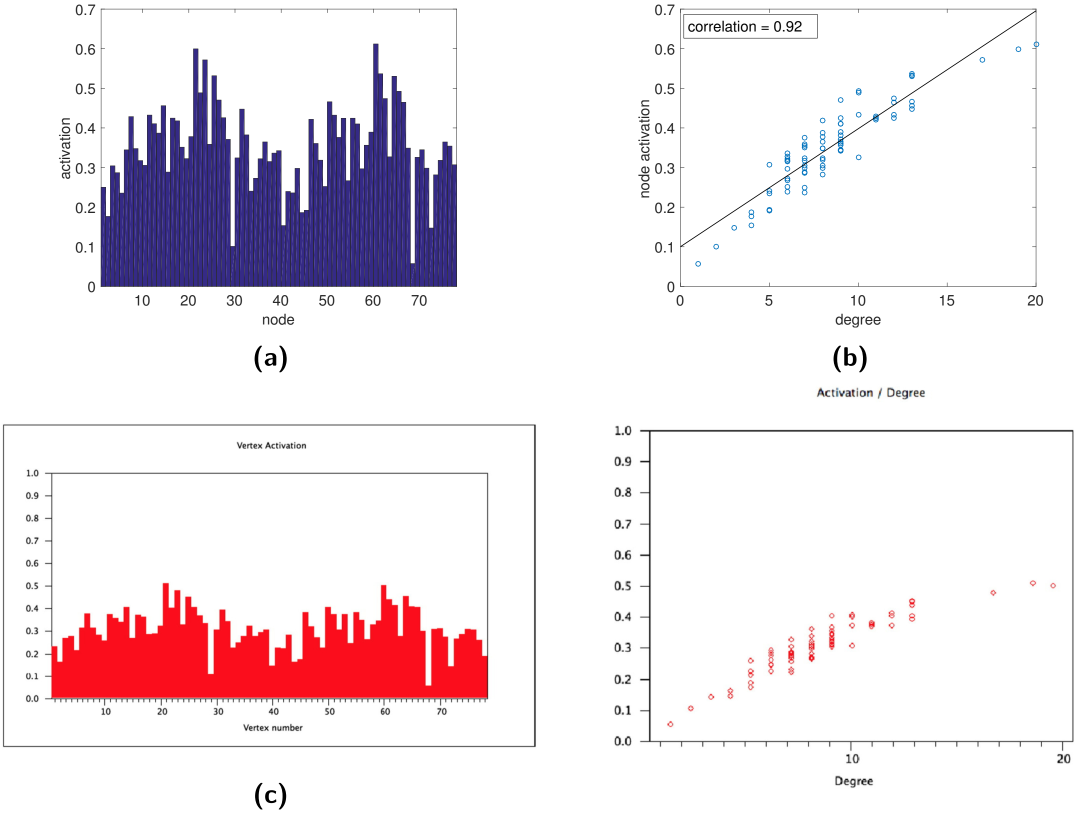
The effective connectivity in Equation (4.1) is computed with other delays shown in Figure 4.4. The connectivity patterns are very similar under different delays.

The effectivity connectivity reveals some interactions related to time delay, which can also further help to explain the structure that shapes the dynamics. For example, in the effective connectivity, many connections around the diagonal disappeared while the four 'blobes' remains visible, comparing to the structural network in Figure 4.2(a). This could be a consequence of the four 'blobes' being connected hubs or rich-clubs. However, the exact relationship between the structural network and the connectivity patterns is not yet clearly explained and more efforts are needed to conduct more extensive research.

Studying the dynamics on the anatomical network provides perception on the formation of the functional network, and thus reveals the mechanisms a brain operates [32]. Researchers have been looking for mathematical functions that can portrait the relationship between the structural and functional network analytically [33, 34]. To show visual relationships, Paper [35] studied the functional connectivity in the brain and verified the feasibility of applying network science to the brain. Furthermore, the paper proposed that analysis on healthy brain reveals the primitive mechanisms from which we can further study the brain disorders.

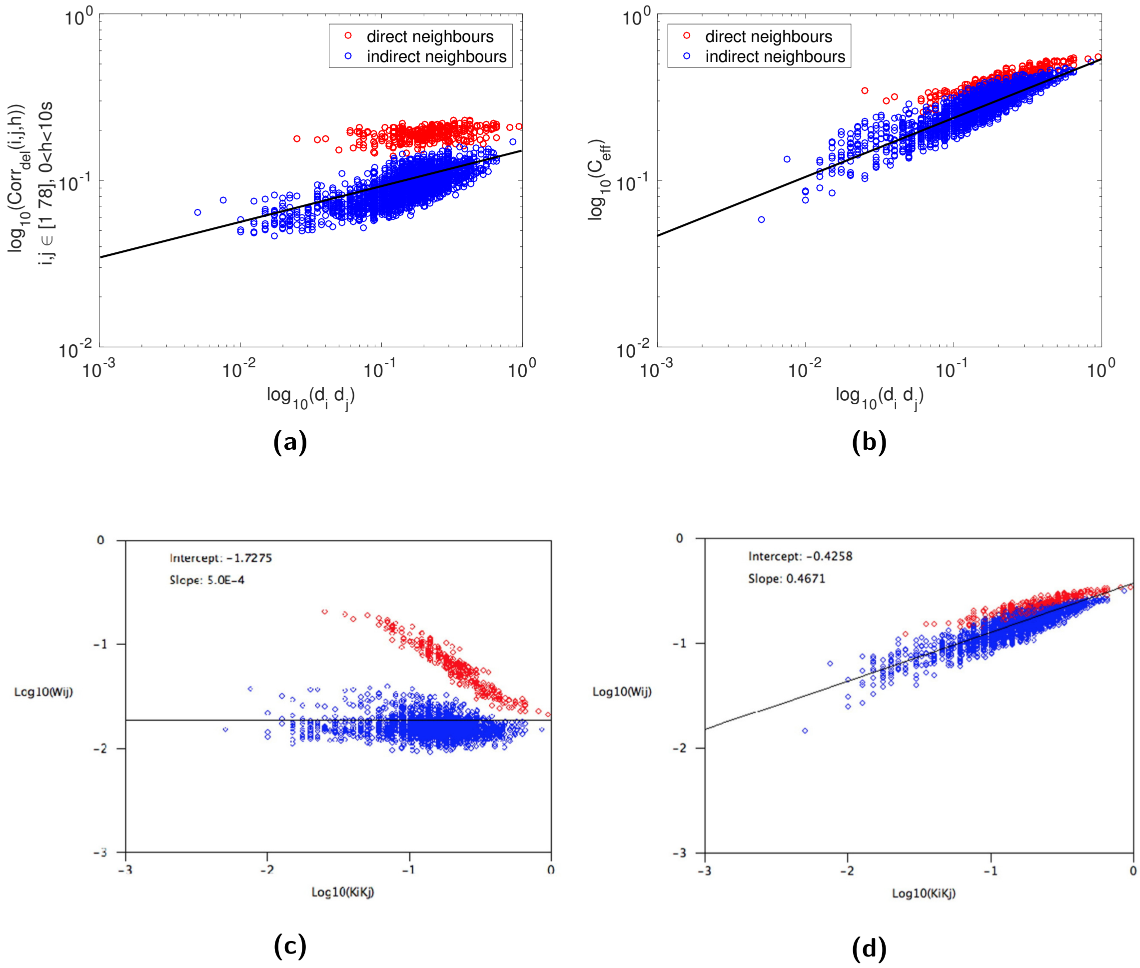
4.3 Confirmation of Simulation Results

Equipped with the continuous-time SIS simulator, simulations with much more accuracy to reality can further verify previous results that have been done in the discrete-time simulator. Referencing to the work in [21], comparisons are performed on the node activation and the degree versus activation plots in Figure 4.5, as well as the degree product versus functional and effective connectivity scatter plot in Figure 4.6. In general, high degree nodes activate for a longer time or more frequently. A rather negative correlation between degree and functional correlation is found due to higher τ than critical τ. A positive correlation between degree and effective connectivity is also confirmed at similar τ. General principals from previous work in discrete-time simulations match with those from the continuous-time simulation.



(a) (b) (c) (d)

Figure 4.5: (a) Node activation time. (b) Degree versus node activation. (c)Node activation time taken from [21]. (d) Degree versus node activation taken from [21]. The dynamics are chosen at similar effective spreading rate, and the underlying structural network has taken the same network used in this work.



(a) **(b)**

(c) **(d)**

Figure 4.6: Degree product versus the (a) functional and (b) effective connectivity in log-log scale. Degree product versus the (c) functional and (d) effective connectivity in log-log scale taken from [21]. The dynamics are chosen at similar effective spreading rate, and the underlying structural network has taken the same as use in this work.

Chapter 5

Delayed Correlations

In this chapter, we will introduce the delayed correlation to explore the pairwise node influence in a spreading process on the structural network. The delayed correlation is an easy and straightforward metric that simply measures the linear correlation as a function of time delay. Simulation results will be provided in order to study the delayed correlation over different time delays.

5.1 Analytic Derivations

Pearson's Correlation Coefficient

The Pearson's correlation coefficient or often short as correlation is generally used to describe the linear relationship of two simple data sets. The Pearson's correlation coefficient of node i and j is defined as

$$corr(i,j) = \frac{Cov(X_i(t), X_j(t))}{\sigma_{Xi}\sigma_{Xj}}, \tag{5.1}$$

which measures the covariance of two data sets dividing the product of standard deviation of each data set. The conceptual formula can be expressed in various forms due to different interpretations on the concept [8]. Based on alternative expressions for covariance and standard deviation, as well as denoting the data sets as time series $X_i(t)$ and $X_j(t)$ of two nodes, an alternative expression for Equation (5.1) that is used throughout our work is written as

$$corr(i,j) = \frac{E[X_i(t)X_j(t)] - E[X_i(t)]E[X_j(t)]}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_j(t)]}}. \tag{5.2}$$

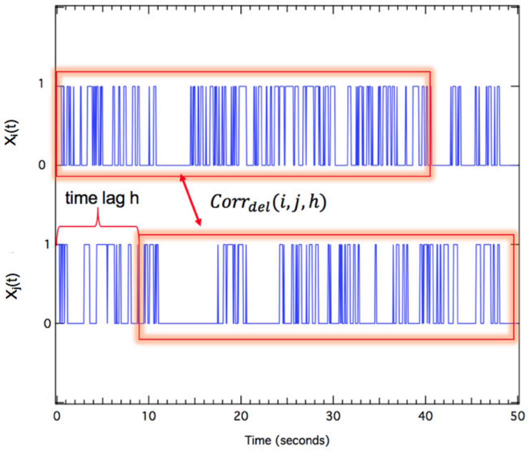


Figure 5.1: Schematic graph for the computation of the delayed correlation. $X_i(t)$ and $X_j(t)$ are time series of two nodes i and j, generated from one simulation and the first 100 seconds are presented here as an example. The length of both time series will be shortened by the length of the time lag in the calculation of delayed correlation to ensure both time series are of the same length.

The linear correlation coefficient is always constrained to the range of $[-1,1]$ [36]: a coefficient equal to 1 means two data sets are linearly and positively related that increase (decrease) of one series involves also an increase (decrease) in the other series; a coefficient equal to -1 means that two data sets are linearly but negatively related that one series increase (decrease) results in another series decrease (increase); a coefficient equal0 means two data sets are not linearly correlated. The 0 correlation coefficient does not lead to independence while independence yields 0 correlation coefficient.

Delayed Correlations

Supplied by the SISS with time series $X_i(t)$ for all nodes in a network, an elementary query would be to investigate the pairwise node correlation and interactions. The pairwise Pearson's correlation coefficient is measured as a function of time delay (or time lag) h to characterise the pairwise relationship in a spreading process [37]. To highlight the time delay variable, the measurement of linear correlation as a function of time delay is named as the 'delayed correlation'. It is noteworthy that our 'delayed correlation' (from the statistics perspective) is the same concept as the 'cross-correlation' from the signal processing field [8].

The delayed correlation of node i and node j with certain time delay h is given as

$$corr_{del}(i,j,h) = \frac{E[X_i(t)X_j(t+h)] - E[X_i(t)]E[X_j(t+h)]}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_j(t+h)]}}, \tag{5.3}$$

and a schematic graph in Figure 5.1 is shown as an example illustration.

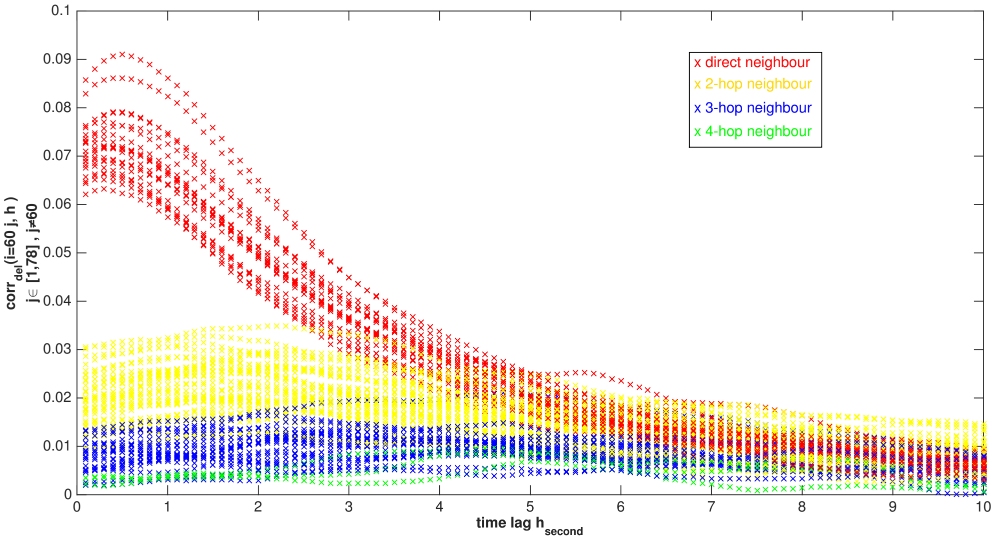


Figure 5.2: Delayed correlations of node 60 to all other 77 nodes as a function of time lag h. Node numbe 60 is the node of highest degree. Different colours indicating delayed correlation of k-hop neighbours of the node 60.

The correlation of two identical data set or a data set and itself with some time shifts is called the autocorrelation which written following our notation is

$$autocorr_{del}(i,h) = corr_{del}(i,i,h) = \frac{E[X_i(t)X_i(t+h)] - E[X_i(t)]E[X_i(t+h)]}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_i(t+h)]}}. \tag{5.4}$$

5.2 Simulation Results

Delayed Correlations

Based on the time series of all nodes generated in Chapter 3, we then calculate the pairwise delayed correlation. For easy visualisation out of mass data on a graph, we explicitly show the delayed correlation of one node to all other 77 nodes in the structural network. The node number 60 has the highest degree and is taken as an example shown in Figure 5.2. The pairwise delayed correlation of other nodes can be found at the end of Appendix C.

Pair wise correlations of three hub nodes 21, 23 and 60 to every other 77 nodes is shown in Figure 5.3. There are some common characteristics observed from both Figure 5.2 and Figure 5.3:

- the delayed correlations at a certain time lag decrease as node distance increases;
- neighbours have higher delayed correlations, which is due to neighbours are connected so the influence spreads very fast between neighbours;

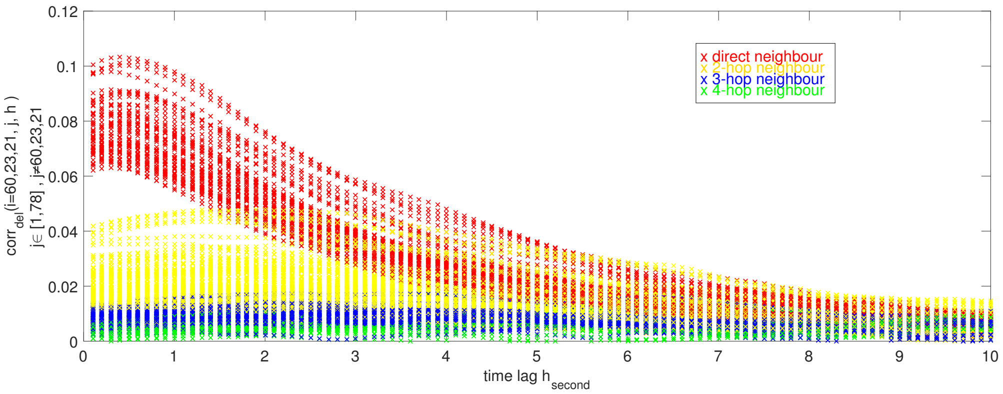


Figure 5.3: Delayed correlations of hubs node 21, 23 and 60 respectively, to all other 77 nodes in the structural network.

- in both figures only 4 hops are seen, because the furthest distance from a hub node to any other node is 4 hops in the structural network, and the diameter of the network is 6 hops;
- within a small delay interval $0 < h < 1$, there is a small increasing phase especially between neighbour nodes;
- the delayed correlation in general is decreasing over time lag h.

From the above observations we found at one delay, the delayed correlations seem to be decreasing as node distances increase and neighbours have higher delayed correlations. This finding agrees with previous knowledge that information are spreading on links. Since neighbours are directly connected, their states or information are more synchronised showing a higher correlations. A node pair with larger distance needs more time in transmission which result in slower synchronisation, therefore delayed correlations decrease accordingly. The simulation results confirmed that nodes have more influence on neighbours because the epidemic is spreading hop-by-hop and thus far away nodes are reached later.

Autocorrelations

A very small delay h in the variance $Var[X_i(t+h)]$ can be considered as no effect if the time series is long and the denominator of autocorrelation can also be considered as a constant value. The numerator of covariance with delay is found to be decreasing over time lag h, and a detailed derivation on the numerator can be found in Appendix B. The autocorrelation decreases over time lag h as shown in Figure 5.4. Within very small time lag interval (for example $h < 2s$ as observed), the autocorrelations can be considered as decreasing exponentially (Figure 5.5). Further, the scatter plot of autocorrelation versus node degree has a negative fitting slope in Figure 5.6. The negative fitting slope can be interpreted as high degree nodes always have lower autocorrelations regardless of the delay. High degree nodes, or hubs, have more neighbours which according to findings on the delayed correlations, should rapidly adapt to each other's states due to fast transmission on direct links. With more adaption to neighbour's states, a node becomes less self-resemblance,

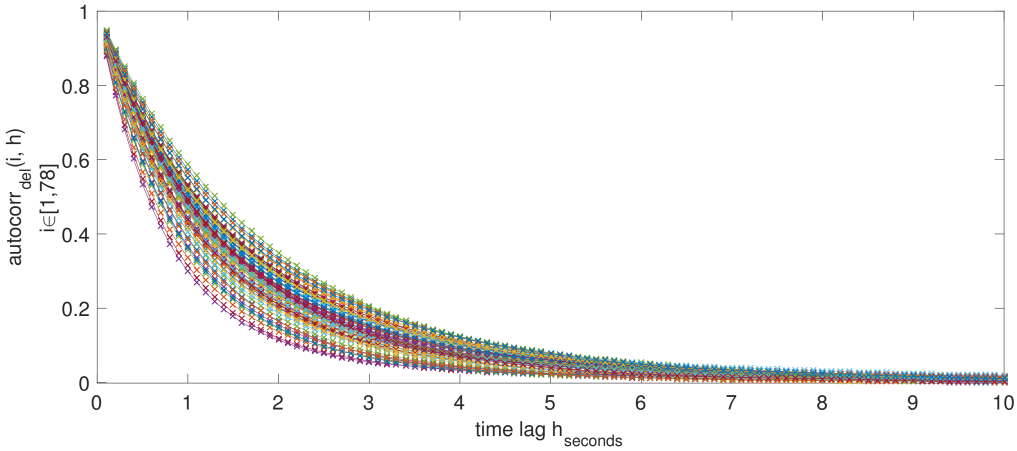


Figure 5.4: Autocorrelations of all nodes over time lag h. The colors represents different nodes.

which results in a lower autocorrelations.

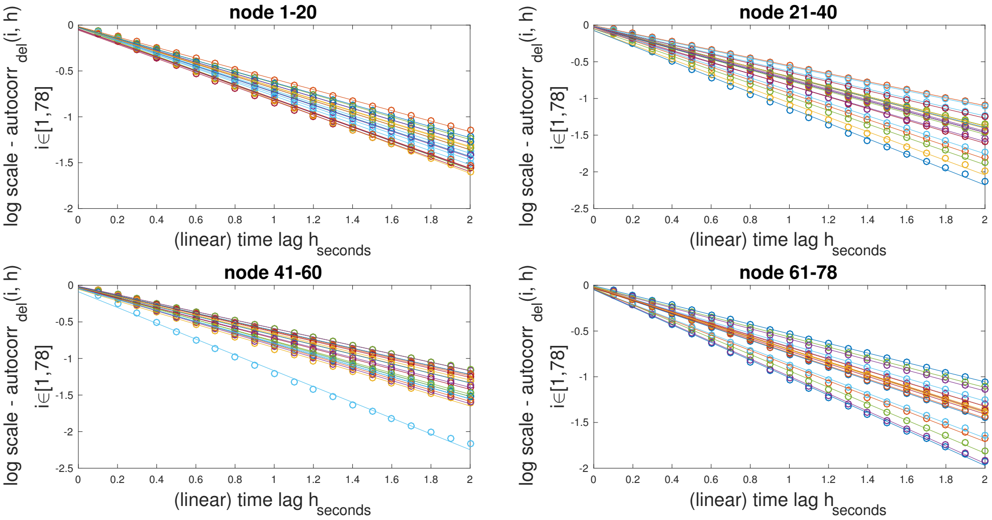


Figure 5.5: Semi-log autocorrelations within very small time lag interval $h < 2 seconds$. For every node, the log delayed autocorrelation is fitted by a line on the linear time lag.

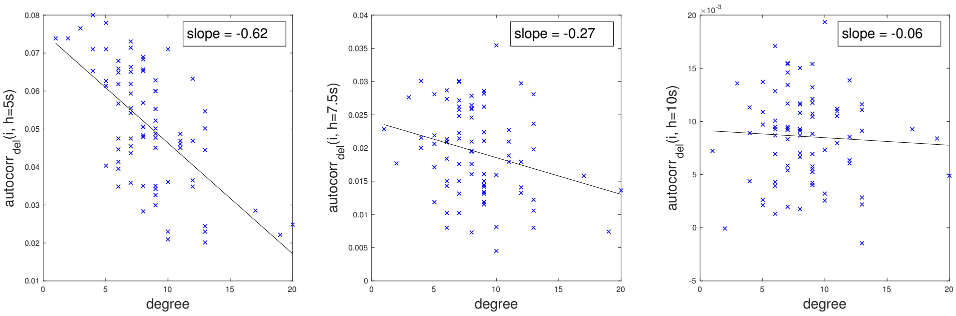


Figure 5.6: Autocorrelations versus degree at $h = 0.1s, 3.5s, 6s$, all three cases have negative fitting slope.

Chapter 6

Emergence of Global Patterns

In the previous chapter, we mainly investigated on the interaction and correlations between pairs of nodes in the structural brain network. In this chapter, we will further study the influence by exploring the direction of spreading on the network.

6.1 Directionality Measured by Delayed Correlations

In order to study the influence and functionality of the brain network, the direction of information flow on the brain network is often analysed from various measurements of MEG, fMRI, EEG and DTI[38, 39, 40] as mentioned in Chapter 4. A recent study has found from the empirical data that brain dynamics often show a posterior-anterior global pattern [41]. We further explore on the directionality using the SIS epidemic model to compare the model analysis with the empirical results.

Chapter 5 explored non-symmetric interactions between two nodes under differently time delays. To further enlarge the directionality, we normalise the delayed correlation to see which direction dominates. Inspired by [41], we compute the normalised directed delayed correlation as

$$dcorr(i,j,h) = \frac{corr_{del}(i,j,h)}{corr_{del}(i,j,h) + corr_{del}(j,i,h)}. \tag{6.1}$$

This normalisation method is valid if the delayed correlations are non-negative and the denominator is non-zero. The correlations of the dynamic time series in the SIS model is proved to be non-negative of the whole spreading period [42]. Our simulation of 4096 seconds shows most delayed correlations are also positive while in case of a negative value, longer simulation period is needed. The directed delayed correlation is then a value between 0 and 1 due to the normalisation, and we interpret node i is more sending towards node j if $dcorr(i,j,h) \geq 0.5$ and more receiving if the smaller than 0.5.

Peviously, we observed an increasing phase around time delay $h = 0.5s$ of delayed correlations before a general decreasing phase (see Figure 5.2 and Figure 5.3). A more spreading process may result in such an increase phase of delayed correlations while a stable system does not generate much changes. Therefore, we examine the directionality at this interesting time delay $h = 0.5s$ and another interesting time delay is introduced later as a benchmark. The resulting directed delayed correlation matrix for the structural brain network is shown in Figure 6.1(b) where the delayed correlations are displayed in matrix form in Figure 6.1(a).

We further take average on the sum of pairwise directed delayed correlation for each node, and display as the directionality index of node i at time delay h as

$$d_{index}(i,h) = \frac{1}{N}\sum_{j=1}^{N} dcorr(i,j,h). \tag{6.2}$$

The directionality index for all nodes in the network is shown in Figure 6.1(c). We characterise the more sending or receiving feature of a node by regarding the node with directionality index $d_{index}(i,h) \geq 0.5$ to be sending node while the reverse as receiving.

The directionality index is interpreted as the information flow of regional communication in the brain and is shown in Figure 6.1(d). A major information flow is from the back to front since more sending nodes (in dark red colour) are observed in the back of the brain, where most hub nodes are located at the back of the brain see Figure 6.2).

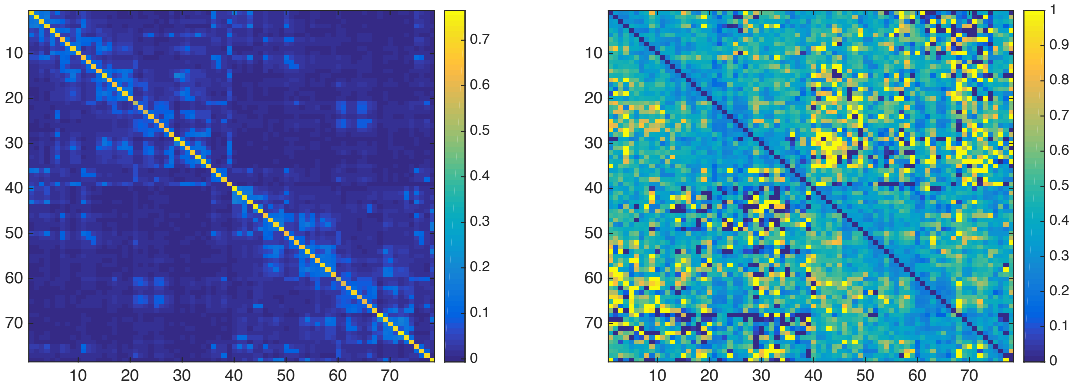
Empirical data often observe a posterior-anterior global pattern which can be quantified by the PAx value was computed in [41]. The classification of posterior and anterior nodes of the anatomical brain network can found in the supporting material of paper. To quantify the posterior-anterior pattern on the simulation data and make comparison with the empirical results, we compute the PAx value in the same way

$$PAx = \left\{\overline{d_{index}(i,h)}\right\}_{i\in posterior} - \left\{\overline{d_{index}(i,h)}\right\}_{i\in anterior}. \tag{6.3}$$

Another choice of delay time $h = 2.9 seconds$ is taken for comparison with the work from [41]. The computation of delay is interpreted in our notation as

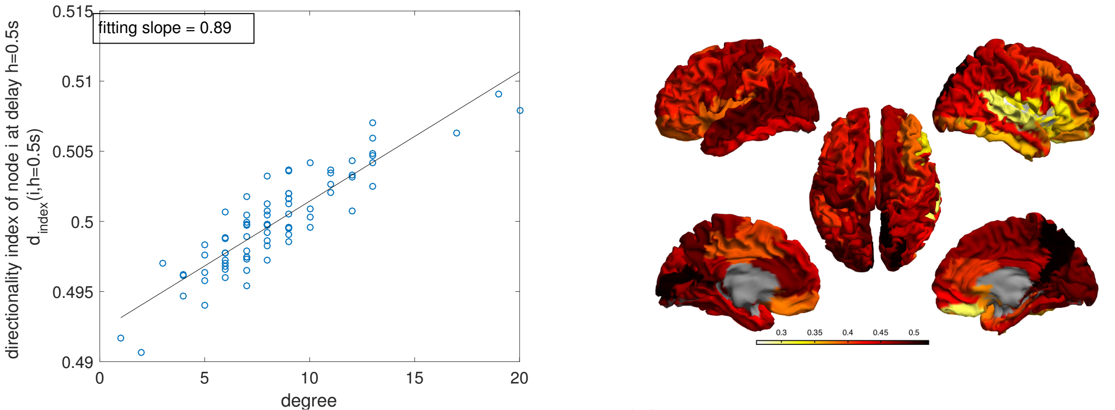
$$delay = \frac{number\ \ of\ \ samples\ \ \times\ \ number\ \ of\ \ ROIs}{total\ \ number\ \ of\ \ status\ \ change\ \ in\ \ the\ \ network},$$

where the number of samples is 40960 (sampling at every 0.1s), and the number of ROIs in our structural brain network is 78 in AAL atlas, and the denominator is counted as the total number



(a) delayed correlation matrix

(b) directed correlation matrix



(c) directionality index, PAx=0.009

(d) directionality index displayed on a brain

Figure 6.1: Directionality measured by normalisation of delayed correlations, h=0.5s. (a) Delayed correlation matrix $corr_{del}(i,j,h)$, computed by Equation 5.3 at $h = 0.5s$. (b) Directed delayed correlation matrix $dcorr(i,j,h)$, computed by Equation 6.1 at $h = 0.5s$. (c) The directionality index $d_{index}(i,h)$ at $h = 0.5seconds$ for all 78 nodes in the structural brain network as computed in Equation 6.2. (d) The directionality index shown on a template brain. The figure displays cerebral cortex of the brain in different views-the left, top, right, right midline and left midline in clockwise order. The brain in the middle (from top view) shows both hemispheres, and the posterior to frontal of the brain is bottom to top. The nodes or cortical regions on the brain with a darker colour in red has higher directionality index and are considered as sending regions, while the light colour in yellow represents receiving regions.

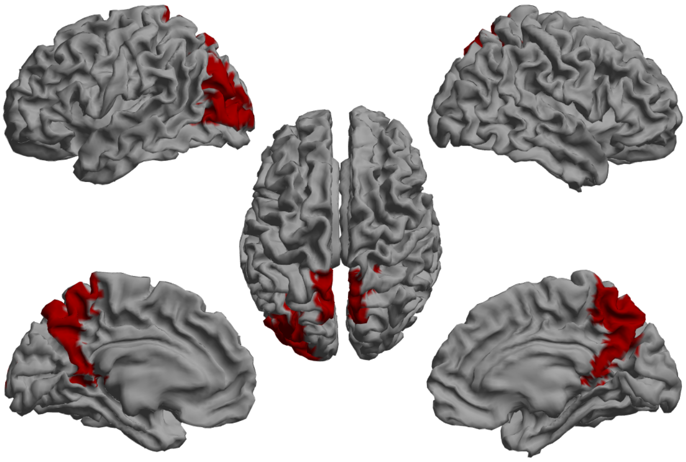


Figure 6.2: Location of the hub nodes on a brain coloured in red. Three nodes, node 21, 23 and 60, with the highest degree are chosen as hubs. The figure displays cerebral cortex of the brain in different views-the left, top, right, right midline and left midline in clockwise order.

of changes between 0s and 1s of all nodes in the whole simulation period of 4096 seconds. And therefore, we computed the $h = 2.9s$. In Figure 6.3 shows the directionality found for time lag $h = 2.9s$, with each subfigure relate to the corresponding subfigure in Figure 6.1 of time lag $h = 0.5s$.

Directionality Measured by Correlation Flux

Another intuitive method to study directionality of information flow is to consider the delayed correlation as a measure of the information flowing at certain delay time and compute the correlation flux. Instead of computing the directed delayed correlation by normalisation in Equation 6.1, we compute the directed flux as

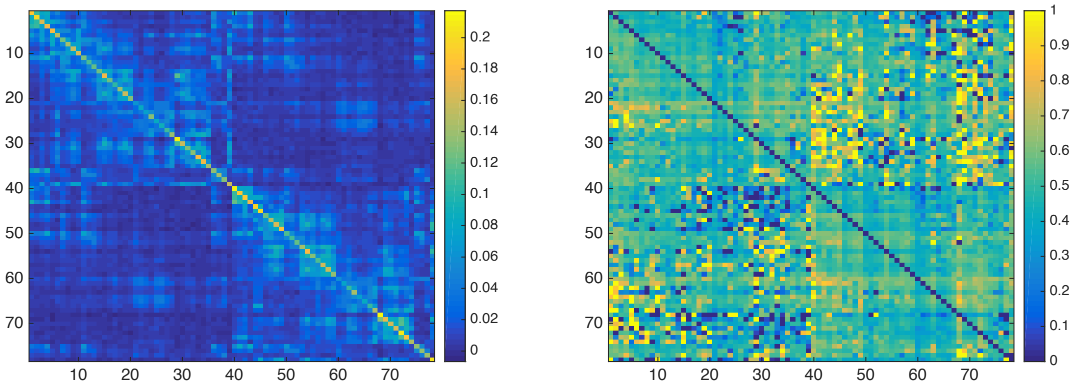
$$dflux(i,j,h) = corr_{del}(i,j,h) - corr_{del}(j,i,h). \tag{6.4}$$

The directed flux is computed similarly as the previous computation for directed delayed correlations in Equation 6.1. Both measures have a one to one relation that a $dcorr(i,j,h) \geq 0.5$ will have $dflux(i,j,h) \geq 0$ and vice versa.

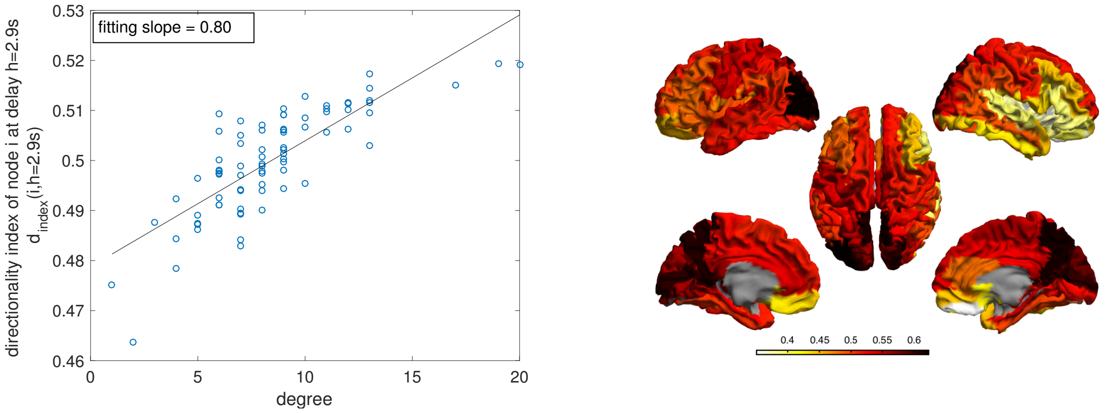
The directionality index of each node is computed as

$$dflux_{index}(i,h) = \frac{1}{N}\sum_{j=1}^{N} dflux(i,j,h), \tag{6.5}$$

which is similar to the computation of directionality index measured by normalised delayed correlations in Equation 6.2.



(a) delayed correlation matrix **(b)** directed correlation matrix



(c) directionality index, PAx=0.019 **(d)** directionality index displayed on a brain

Figure 6.3: Directionality measured by normalisation of delayed correlations, h=2.9s. (a) Delayed correlation matrix $corr_{del}(i, j, h)$. (b) Directed delayed correlation matrix $dcorr(i, j, h)$, (c) The directionality index $d_{index}(i, h)$. (d) The directionality index shown on a template brain. The figure displays cerebral cortex of the brain in different views-the left, top, right, right midline and left midline in clockwise order. The brain in the middle (from top view) shows both hemispheres, and the posterior to frontal of the brain is bottom to top. The nodes or cortical regions on the brain with a darker colour in red has higher directionality index and are considered as sending regions, while the light colour in yellow represents receiving regions.

The directed flux matrix $dflux(i,j,h)$ at $h = 0.5 seconds$ is shown in Figure 6.4(b) and the directionality index measured by flux $dflux_{index}(i,h)$ at the same time delay is shown in Figure 6.4(c). The node with directionality index $dflux_{index}(i,h) \geq 0$ is considered as a more sending node while node with $dflux_{index}(i,h) < 0$ is considered as a more receiving node. The correlation flux directionality index is also displayed on a template brain in Figure 6.4(d). Also, the directionality is examined at another time lag $h = 2.9s$ in Figure 6.5.

6.2 Directionality Measured by Transfer Entropy

Instead of the delayed correlation approach measuring the directionality in brain networks, another method usually used in neuroscience is the transfer entropy (TE) [43, 44]. Data sets from other fields may be much more complicated and thus the transfer entropy method is more useful in analysing output series from oscillators [45]. In MEG measurements, the dynamics time series can often be described as amplitudes and phases (such as a sine wave signal), neuroscientists often apply transfer entropy only to signal phase which is then called phase transfer entropy [46]. However, our SIS time series is binary so the phase transfer entropy would be inapplicable. Following the same concept in finding the directionality by delayed correlation, we will implement the transfer entropy on our SIS dynamics time series to explore the directionality, and compare as a benchmark.

In our notation, the transfer entropy computed on the SIS time series $X_i(t)$ is written as

$$\begin{aligned} TE(i,j,h) = \sum_{k,l,m=\{0,1\}} & Pr[X_j(t+h)=k, X_j(t)=l, X_i(t)=m] \\ & \cdot log_2\left(\frac{Pr[X_j(t+h)=k|X_j(t)=l, X_i(t)=m]}{Pr[X_j(t+h)=k|X_j(t)=l]}\right), \end{aligned} \tag{6.6}$$

and the length of time series $X_i(t)$ in the computation will also be shortened by length h, due to consistency with the series length involving delay. The transfer entropy matrix is shown in Figure 6.6(a) as well as the directed transfer entropy in (b), which is computed as

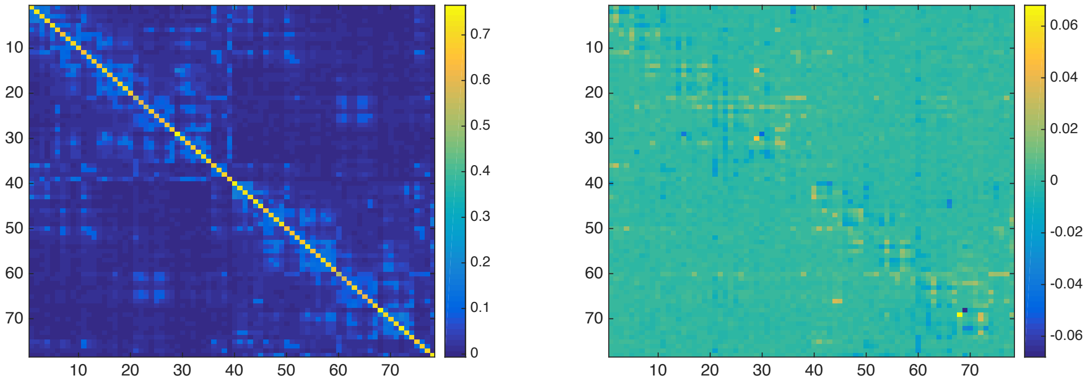
$$dTE(i,j,h) = \frac{TE(i,j,h)}{TE(i,j,h) + TE(j,i,h)}. \tag{6.7}$$

Similarly to Equation 6.2, the transfer entropy directionality index is computed as

$$dTE_{index}(i,h) = \frac{1}{N}\sum_{j=1}^{N} dTE(i,j,h), \tag{6.8}$$

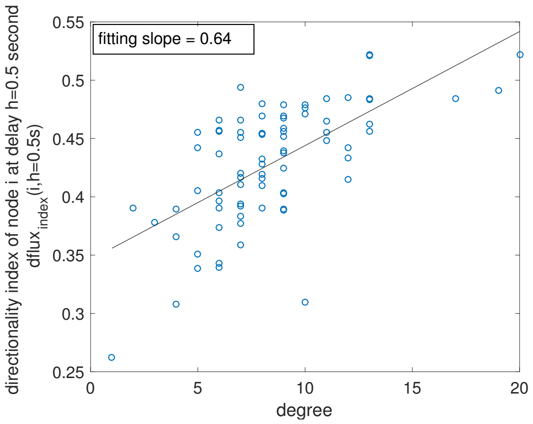
and the result is shown in Figure 6.6(c). The node with directionality index $dTE_{index}(i,h) \geq 0.5$ is considered as a sending node, and node with $dTE_{index}(i,h) < 0.5$ as receiving node. The directionality measured by transfer entropy is shown on a template brain in Figure 6.6(d).

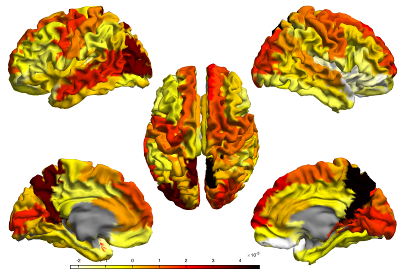
The directionality measured by transfer entropy in another time lag $h = 2.9s$ is also shown in Figure 6.7 as a comparison.



(a) delayed correlation matrix

(b) directed flux matrix

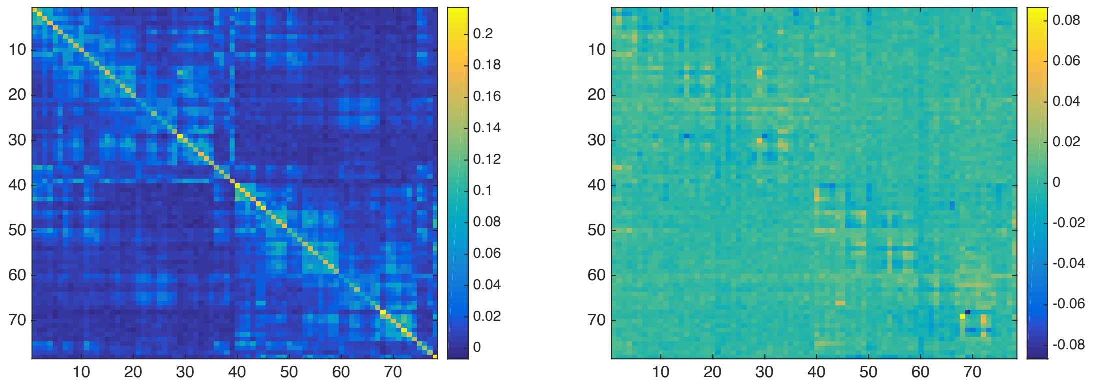




(d) directionality index displayed on a brain

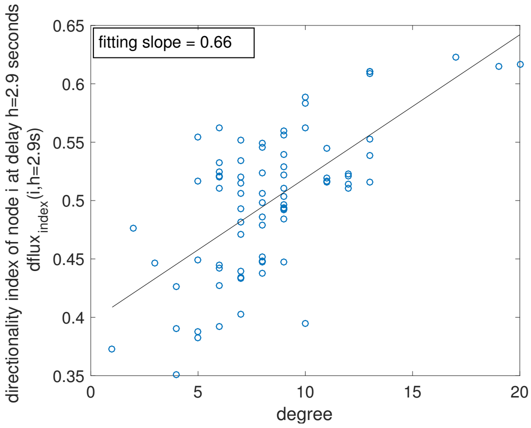
(c) directionality index, PAx = 0.11

Figure 6.4: Directionality measured by correlation flux, h=0.5s. (a) Delayed correlation matrix $corr_{del}(i,j,h)$. (b)Directed flux matrix $dflux(i,j,h)$ at $h = 0.5s$. (c) Directionality index of correlation flux $dflux_{index}(i,h)$ at $h = 0.5s$, of every 78 nodes in the structural brain network. (d) The directionality index shown on a template brain. The figure displays cerebral cortex of the brain in different views-the left, top, right, right midline and left midline in clockwise order. The brain in the middle (from top view) shows both hemispheres, and the posterior to frontal of the brain is bottom to top. The nodes or cortical regions on the brain with a darker colour in red has higher directionality index and are considered as sending regions, while the light colour in yellow represents receiving regions.

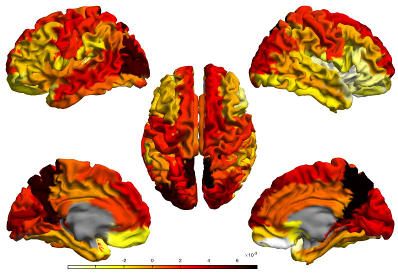


(a) delayed correlation matrix

(b) directed flux matrix

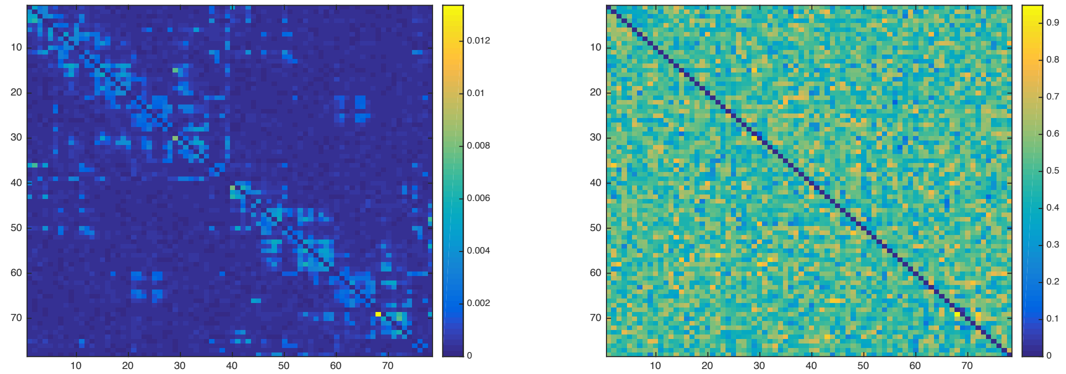


(c) directionality index, PAx = 0.17



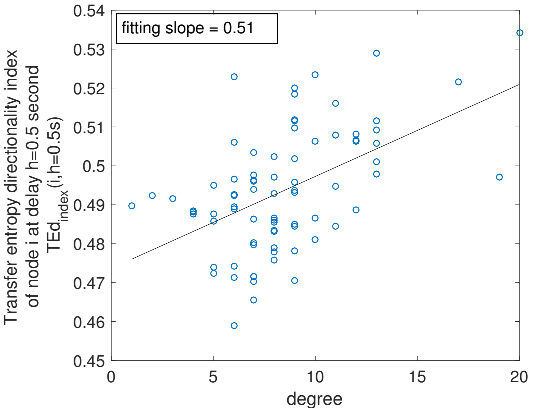
(d) directionality index displayed on a brain

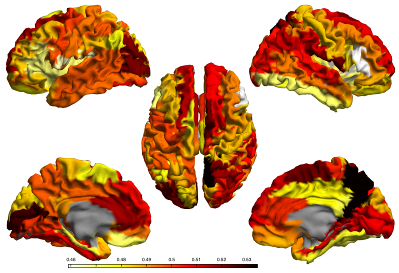
Figure 6.5: Directionality measured by correlation flux, h=2.9s. (a) Delayed correlation matrix $corr_{del}(i,j,h)$. (b)Directed flux matrix $dflux(i,j,h)$ at $h=2.9s$. (c) Directionality index of correlation flux $dflux_{index}(i,h)$ at $h=2.9s$, of every 78 nodes in the structural brain network. (d) The directionality index shown on a template brain. The figure displays cerebral cortex of the brain in different views-the left, top, right, right midline and left midline in clockwise order. The brain in the middle (from top view) shows both hemispheres, and the posterior to frontal of the brain is bottom to top. The nodes or cortical regions on the brain with a darker colour in red has higher directionality index and are considered as sending regions, while the light colour in yellow represents receiving regions.



(a) transfer entropy matrix

(b) directed transfer entropy matrix

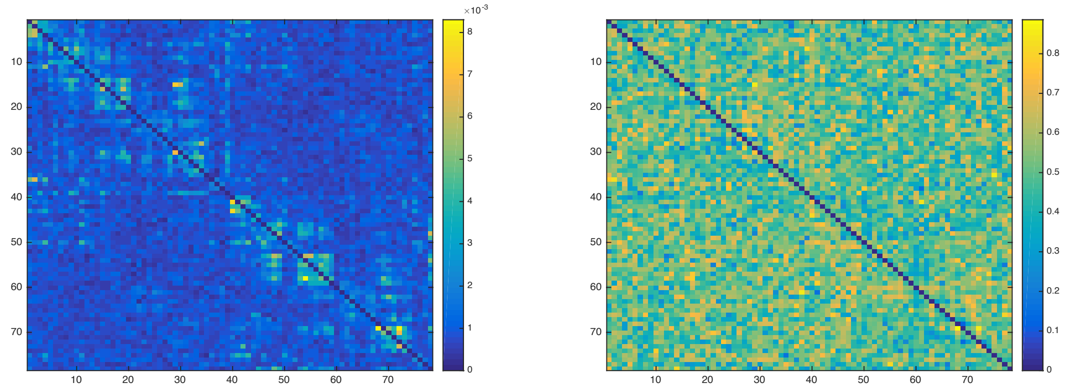




(d) directionality index displayed on a brain

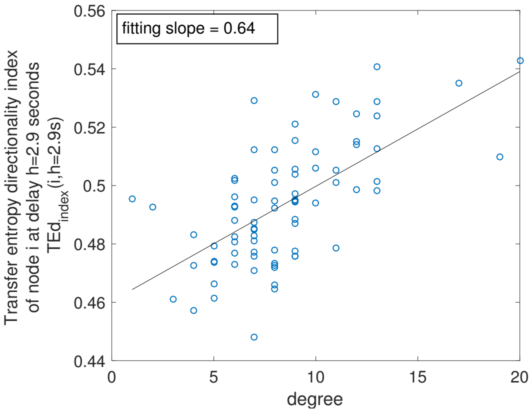
(c) directionality index, PAx=0.0153

Figure 6.6: Directionality measured by transfer entropy, h=0.5s. (a) Transfer entropy matrix $TE(i,j,h)$ at $h = 0.5s$. (b) Directed transfer entropy matrix $dTE(i,j,h)$ at $h = 0.5s$. (c) The transfer entropy directionality index $dTE_{index}(i,h)$ at $h = 0.5s$, for all 78 nodes in the structural brain network. (d) The directionality index displayed on a template brain. The figure displays cerebral cortex of the brain in different views-the left, top, right, right midline and left midline in clockwise order. The brain in the middle (from top view) shows both hemispheres, and the posterior to frontal of the brain is bottom to top. The nodes or cortical regions on the brain with a darker colour in red has higher directionality index and are considered as sending regions, while the light colour in yellow represents receiving regions

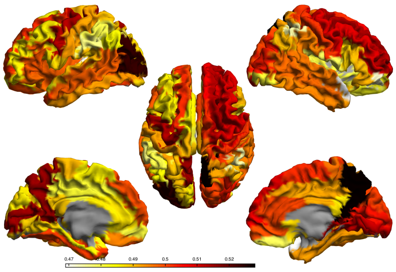


(a) transfer entropy matrix

(b) directed transfer entropy matrix



(c) directionality index, PAx=0.027



(d) directionality index displayed on a brain

Figure 6.7: Directionality measured by transfer entropy, h=2.9s. (a) Transfer entropy matrix $TE(i,j,h)$ at $h=2.9s$. (b) Directed transfer entropy matrix $dTE(i,j,h)$ at $h=2.9s$. (c) The transfer entropy directionality index $dTE_{index}(i,h)$ at $h=2.9s$, for all 78 nodes in the structural brain network. (d) The directionality index displayed on a template brain. The figure displays cerebral cortex of the brain in different views-the left, top, right, right midline and left midline in clockwise order. The brain in the middle (from top view) shows both hemispheres, and the posterior to frontal of the brain is bottom to top. The nodes or cortical regions on the brain with a darker colour in red has higher directionality index and are considered as sending regions, while the light colour in yellow represents receiving regions

6.3 Discussion

Both the delayed correlation matrix (both in Figure 6.1(a) and Figure 6.3(a)) and the transfer entropy matrix (Figure 6.6(a) and Figure 6.7(a)) show the same strucutre of the structural brain network in Figure 4.1: there are strong connection along the matrix diagonal, and two 'blobles' connections areas on the off-diagonal direction. The directed delayed correlation matrix (Figure 6.1(b) and Figure 6.3(b)) computed by normalisation on delayed correlations shows some kind of pattern that some nodes seem to be more sending while other nodes seem to be more receiving. Viewing row-wise, some rows (e.g. rows 21 to 27) are with higher values (in yellow) and is considered as sending more than receiving, while some rows (e.g. rows 41 to 50) are with lower values (in blue) and is considered as receiving more than sending. The sending or receiving feature of a node is represented clearly by the directionality index (Figure 6.1(c) and Figure 6.3(c)), plotted over the degrees.

Both the directed transfer entropy matrix (Figure 6.6(b) and Figure 6.7(b)) and the correlation flux matrix (Figure 6.4(a) and Figure 6.5(a)) does not show visible patterns. The directionality measured by flux is computed based on the value of delayed correlations, and there is a one-to-one mapping that a directed correlation above 0.5 is equivalent to a positive directed flux.

The cortical region communication directionality by three methods: by directed delayed correlation by normalisation, by correlation flux and by transfer entropy, shows a general posterior to anterior (back to front) directionality. The posterior to anterior direction is quantified as the PAx value where a positive PAx value indicates the strength of a back to front global direction. Therefore, the direction of back to front information flow in the brain could be concluded which is also in accordance with empirical results in [41]. The back to front information flow could be generated by the hubs locating at the back of the brain. Intuitively, hub nodes have more neighbours and therefore, they have can largely impact on the network during dynamics [47].

The comparison of different measures also showed that simple analysis can already show global dynamics patterns on the brain which is of highly research interest in characterising the dynamics on a brain network [41].

Chapter 7

Conclusions and Future Work

In this thesis, network properties on the structural brain network have been studied. A simple SIS epidemic model has been applied to the structural network to study the dynamic spreading processes. Previous research results were confirmed with a continuous-time simulator that is a closer imitation of the reality world. The functional connectivity as a measure of the correlation of time series is confirmed to resemble the underlying structural network. In comparison with empirical data, the functional connectivity agrees more with the empirical data measured from MEG than that from fMRI. The Effective connectivity as a measure of the conditional probability on the time series is confirmed to be a rather fixed pattern that diagonal links are missing. The node activation as a function of degree and connection weight versus degree product were also confirmed that high degree nodes are activated more frequently and neighbours have higher connection weights.

The delayed correlations of dynamic time series of the brain are closely analysed to reveal dynamic pattens as a function of different time delays. Pairwise delayed correlations seem to be decreasing in general, while within a small delay interval, the correlation conducts a small increase phase before monotonic decreasing. Pairwise correlations at one delay time seem to be ranking descending according to distance increase that neighbours have higher correlations than that of other nodes at each unit delay time. However, the correlation is negatively related with degrees at each unit delay time that nodes with more neighbours seem to be changing status more often due to a direct spreading between neighbours. In the SIS model used to study the dynamics of the brain, delays within a small time interval are most interesting during which dynamics are rapidly spreading. This increasing interval seems to be corresponding to the average spreading time in the network.

The delayed correlation also showed a direction of information flow from posterior to anterior, which also agrees with results in [41]. The directionality is benchmarked with transfer entropy method in the favour of medical fields. Hubs located at the back of the brain seem to be conducting crucial role in the spreading to the whole network. The directionality of spreading flow can be considered as a type of influence that information given to the other node can ensure the spreading towards the target nodes while the reverse cannot be guaranteed. Researching on the direction of spreading flow reveals general sending and receiving properties of nodes in a network. The sending and receiving property can be useful in conducting effective and efficiency spreading in a network as well as in minimising an undesired spreading process. For example, an important epidemic alert message spreading on a social network could benefit from the direction of spreading flow that imposing the alert message on the sending nodes can result in a fast and effective spreading to the whole network. Since hubs are found to be strong sending nodes, they have more influence on the network. Therefore our results could be applied in the epilepsy treatment that treatment should focus on stopping hubs to send overloaded signals on a disordered brain system.

Applying the simple SIS model enables us to study the general connectivity patterns in the brain. The simple delayed correlation method shows similar results to analysis of transfer entropy method which is much more complex. Conducting dynamic processes on the brain network shows global patterns and the emergence from the underlying topology can be studied via various network metrics. Knowledge of how the topology influences the dynamics can help to explain the underlying dynamical mechanism. Dynamical outcomes can even be predicted if the topology influence is well-studied, which can improve on the brain disorder treatments.

Future Work

The structural brain network used in this work is a rather small network containing only the important regional connections and weaker detailed connections are discarded. A larger brain network considering more regions could be applied to study the general properties of the brain network. Other special topologies should be investigated on the emergence of a global pattern. For example, does a lattice graph with no hubs still form a posterior-anterior pattern? Furthermore, the machine learning approach has been implemented in computational neuroscience for brain imaging on a vast amount of neuronal data, which also could be utilised to analyse the dynamic connectivity pattern over a large amount of time series data.

Interdisciplinary research has provided valuable approaches to study the brain. Network science is one major approach which can provide both local and universal network properties. Both partially neuronal level or general regional level analysis are crucial results in enriching the knowledge about the brain.

Appendix A

AAL list

A list of cortical regions to the corresponding node index of the 78-node structural network [41] is presented on the next page. The regions coloBrickRed in BrickRed are posterior regions and regions coloBrickRed in OliveGreen are anterior regions.

1 Rectus-L
2 Olfactory-L
3 Frontal-Sup-Orb-L
4 Frontal-Med-Orb-L
5 Frontal-Mid-Orb-L
6 Frontal-Inf-Orb-L
7 Frontal-Sup-L
8 Frontal-Mid-L
9 Frontal-Inf-Oper-L
10 Frontal-Inf-Tri-L
11 Frontal-Sup-Medial-L
12 Supp-Motor-Area-L
13 Paracentral-Lobule-L
14 Precentral-L
15 Rolandic-Oper-L
16 Postcentral-L
17 Parietal-Sup-L
18 Parietal-Inf-L
19 SupraMarginal-L
20 Angular-L
21 Precuneus-L
22 Occipital-Sup-L
23 Occipital-Mid-L
24 Occipital-Inf-L
25 Calcarine-L
26 Cuneus-L
27 Lingual-L
28 Fusiform-L
29 Heschl-L
30 Temporal-Sup-L
31 Temporal-Mid-L
32 Temporal-Inf-L
33 Temporal-Pole-Sup-L
34 Temporal-Pole-Mid-L
35 ParaHippocampal-L
36 Cingulum-Ant-L
37 Cingulum-Mid-L
38 Cingulum-Post-L
39 Insula-L

40 Rectus-R
41 Olfactory-R
42 Frontal-Sup-Orb-R
43 Frontal-Med-Orb-R
44 Frontal-Mid-Orb-R
45 Frontal-Inf-Orb-R
46 Frontal-Sup-R
47 Frontal-Mid-R
48 Frontal-Inf-Oper-R
49 Frontal-Inf-Tri-R
50 Frontal-Sup-Medial-R
51 Supp-Motor-Area-R
52 Paracentral-Lobule-R
53 Precentral-R
54 Rolandic-Oper-R
55 Postcentral-R
56 Parietal-Sup-R
57 Parietal-Inf-R
58 SupraMarginal-R
59 Angular-R
60 Precuneus-R
61 Occipital-Sup-R
62 Occipital-Mid-R
63 Occipital-Inf-R
64 Calcarine-R
65 Cuneus-R
66 Lingual-R
67 Fusiform-R
68 Heschl-R
69 Temporal-Sup-R
70 Temporal-Mid-R
71 Temporal-Inf-R
72 Temporal-Pole-Sup-R
73 Temporal-Pole-Mid-R
74 ParaHippocampal-R
75 Cingulum-Ant-R
76 Cingulum-Mid-R
77 Cingulum-Post-R
78 Insula-R

Appendix B

Derivation on autocorrelation

The terms in the nominator of autocorrelation can be written as [48]

$$
\begin{aligned}
E[X_i(t)X_j(t+h)] = (1-\delta h)E[X_i(t)X_j(t)] + \beta h \sum_{k=1}^{N} a_{kj} E[X_i(t)X_k(t)] \\
- \beta h \sum_{k=1}^{N} a_{kj} E[X_i(t)X_j(t)X_k(t)] + o(h),
\end{aligned}
\tag{B.1}
$$

and

$$
E[X_j(t+h)] = (1-\delta h)E[X_j(t)] + \beta h \sum_{k=1}^{N} a_{kj} E[X_k(t)] - \beta h \sum_{k=1}^{N} a_{kj} E[X_j(t)X_k(t)] + o(h). \tag{B.2}
$$

For delayed auto-covariance $j = i$, Equation B.1 becomes Equation B.3, and Equation B.2 becomes Equation B.4.

$$
\begin{aligned}
E[X_i(t)X_i(t+h)] = (1-\delta h)E[X_i(t)X_i(t)] + \beta h \sum_{k=1}^{N} a_{ki} E[X_i(t)X_k(t)] \\
- \beta h \sum_{k=1}^{N} a_{ki} E[X_i(t)X_i(t)X_k(t)] + o(h)
\end{aligned}
\tag{B.3}
$$

$$E[X_i(t+h)] = (1-\delta h)E[X_i(t)] + \beta h \sum_{k=1}^{N} a_{ki} E[X_k(t)] - \beta h \sum_{k=1}^{N} a_{ki} E[X_i(t)X_k(t)] + o(h) \tag{B.4}$$

For $j = i$, the delayed auto-covariance is

$$r_i(t,h) = E[X_i(t)X_i(t+h)] - E[X_i(t)]E[X_i(t+h)] \tag{B.5}$$

Substituting Equation B.3 and Equation B.4 into Equation B.5 we have delayed auto-covariance:

$$\begin{aligned} r_i(t,h) &= (1-\delta h)E[X_i(t)X_i(t)] + \beta h \sum_{k=1}^{N} a_{ki} E[X_i(t)X_k(t)] - \beta h \sum_{k=1}^{N} a_{ki} E[X_i(t)X_i(t)X_k(t)] \\ &- (1-\delta h)E[X_i(t)]E[X_i(t)] - \beta h \sum_{k=1}^{N} a_{ki} E[X_i(t)]E[X_k(t)] + \beta h \sum_{k=1}^{N} a_{ki} E[X_i(t)]E[X_i(t)X_k(t)] + o(h) \\ &= (1-\delta h)\{E[X_i(t)X_i(t)] - E[X_i(t)]E[X_i(t)]\} - \beta h E[X_i(t)] \sum_{k=1}^{N} a_{ki}\{E[X_k(t)] - E[X_i(t)X_k(t)]\} + o(h) \\ &= (1-\delta h)Var[X_i(t)] - \beta h E[X_i(t)] \sum_{k=1}^{N} a_{ki}\{E[X_k(t)] - E[X_i(t)X_k(t)]\} + o(h) \end{aligned}$$

For Bernoulli random variables $X_i \in \{0,1\}$, we have $E[X_i(t)Xj(t)] \le E[X_j(t)]$, so the second term of the above equation is non-positive. Then we have $r_i(t,h) \le Var[X_i(t)] = r_i(t,0)$

Simulation result for delayed auto-covariance

The decreasing feature of delayed auto-covariance $r_i(t,h)$ is proved in the simulation result shown in Figure B.3 (a). However, the decreasing feature is not guaranteed for delayed auto-correlation $\rho(X_i(t), X_i(t+h))$, although the delayed auto-correlation shows a decreasing trend in general, shown in Figure B.3 (b).

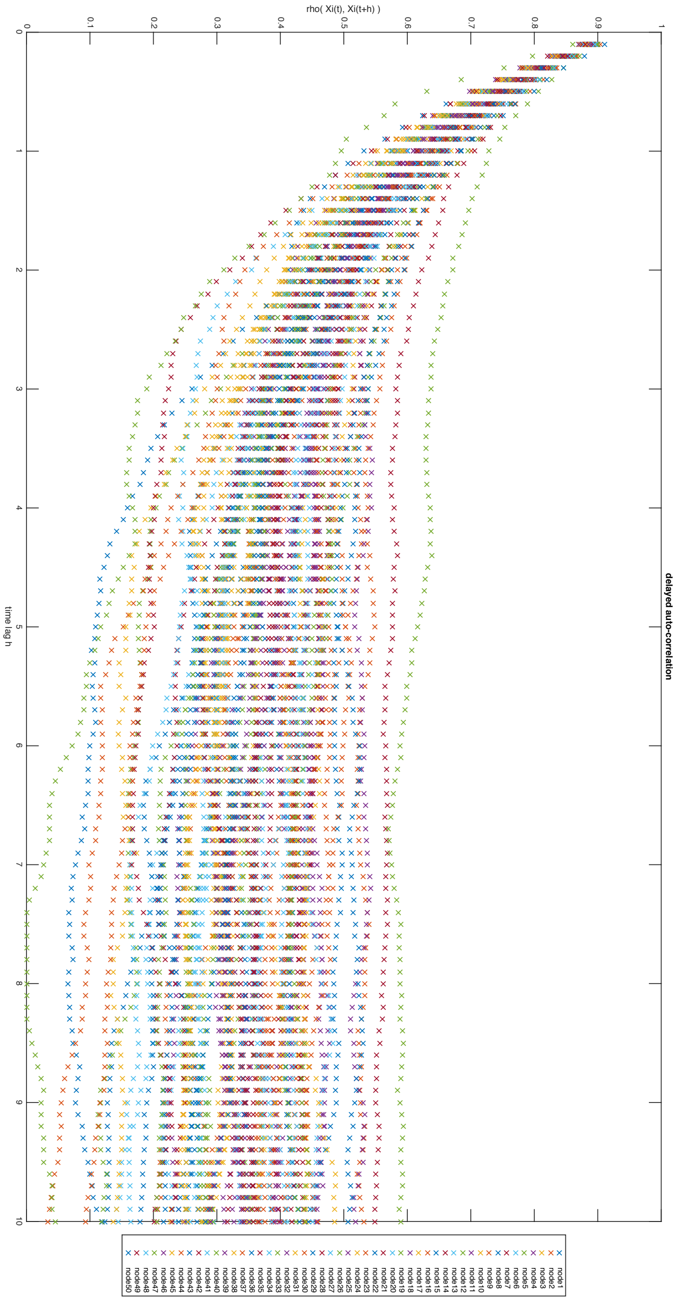


Figure B.1: AutoCorrelation of 1 run

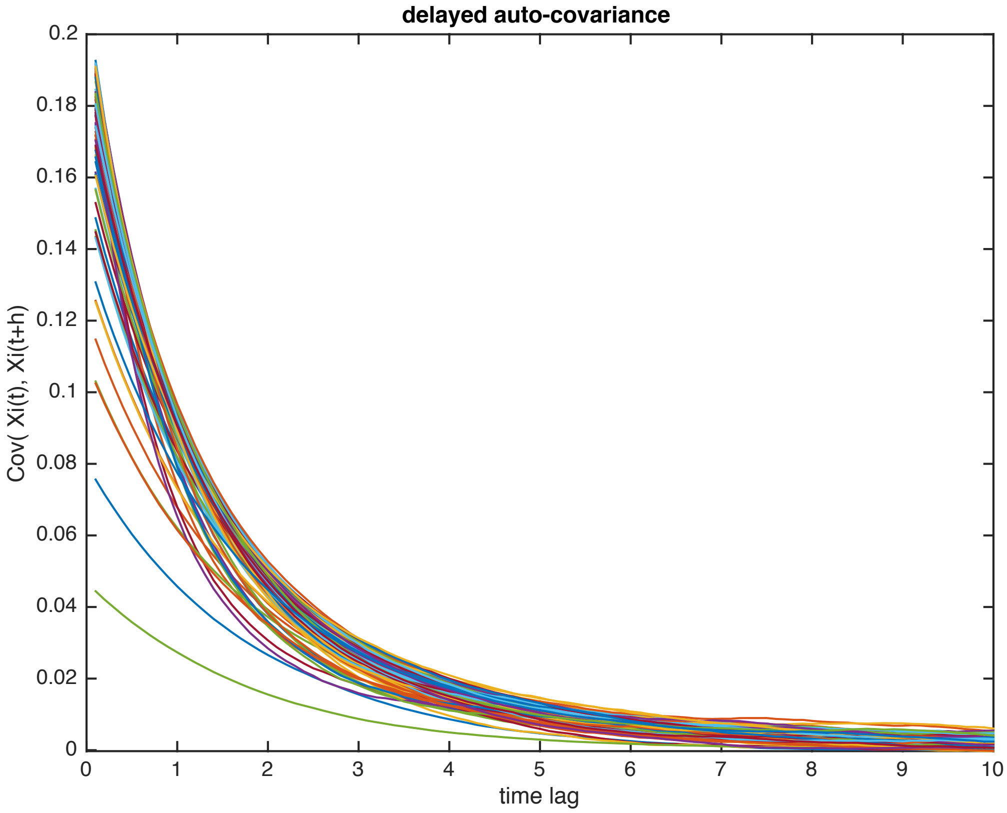


Figure B.2: delayed auto-Covariance

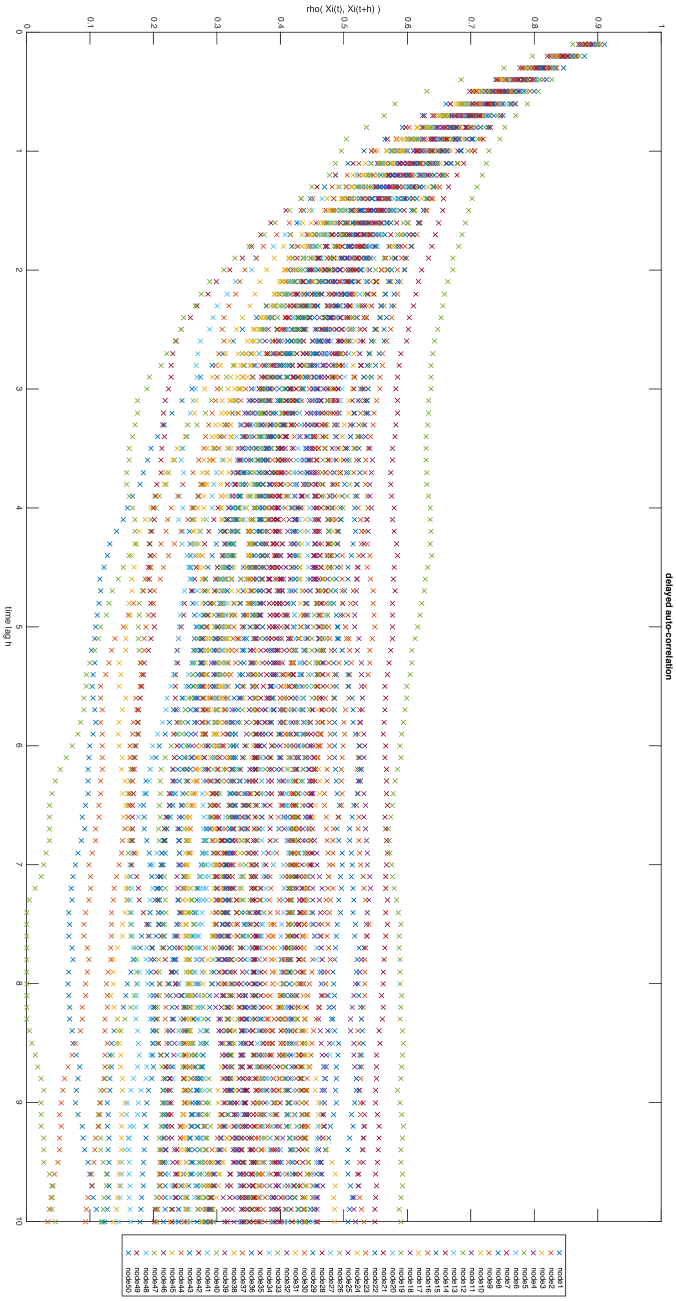


Figure B.3: Auto-Correlation (as a function of delay)

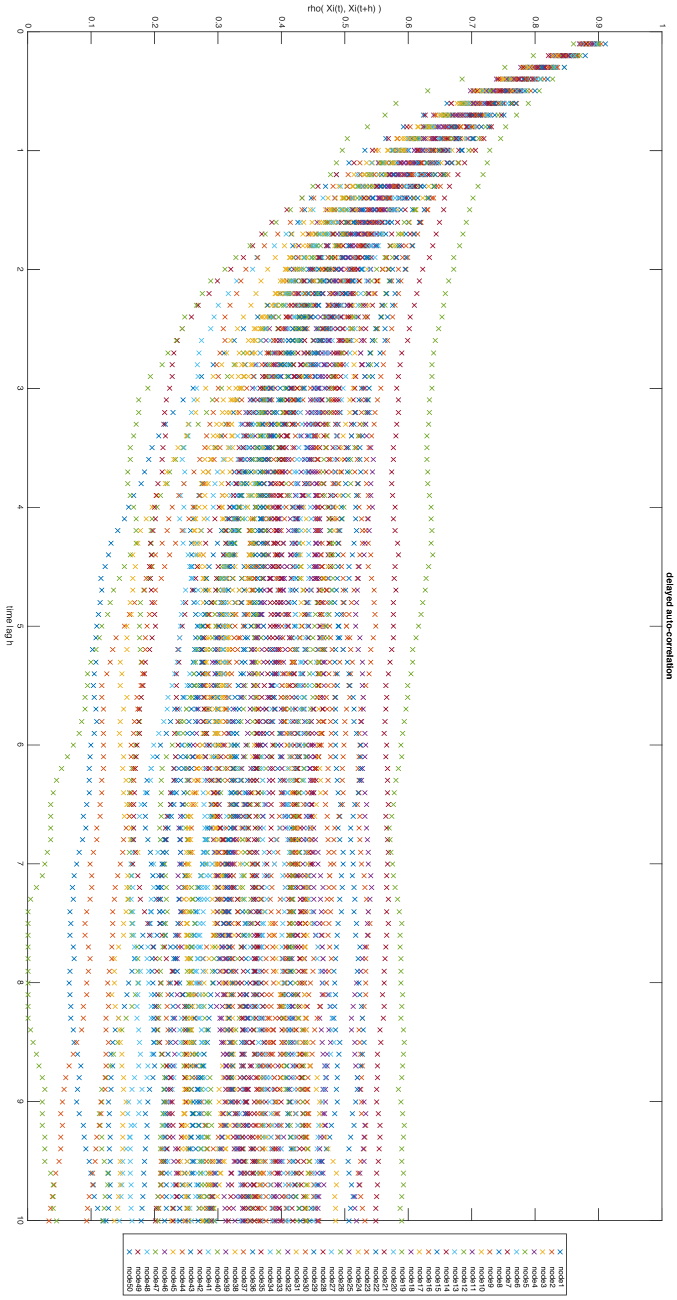


Figure B.4: delayed auto-Cov and delayed auto-Corr

Appendix C

Supporting Figures

C.1 Initial explorations

In Prof. Van Mieghem's Book [8] $equation(2.14)$ gave expected value of a random variable at high moments:

$$E[X^n] = \sum_x x^n Pr[X = x] \tag{C.1}$$

For a Bernoulli random variable, x can only be 0 or 1, thus this equation becomes:

$$\begin{aligned} E[X^n] &= \sum_x x^n Pr[X = x] \\ &= 1^n Pr[X = 1] + 0^n Pr[X = 0] \\ &= 1^n \cdot Pr[X = 1] \\ &= Pr[X = 1] \end{aligned} \tag{C.2}$$

Also :

$$\begin{aligned} E[X] &= \sum_x x Pr[X = x] \\ &= 1 \cdot Pr[X = 1] + 0 \cdot Pr[X = 0] \\ &= Pr[X = 1] \end{aligned} \tag{C.3}$$

Thus, for a Bernoulli random variable we can say $E[X^n] = E[X]$.

The "Probability distribution function" in Prof. Van Mieghem's book section 3.2.3 as well as used in the lecture slides, is actually the cumulative distribution function.

Simulation results explorations

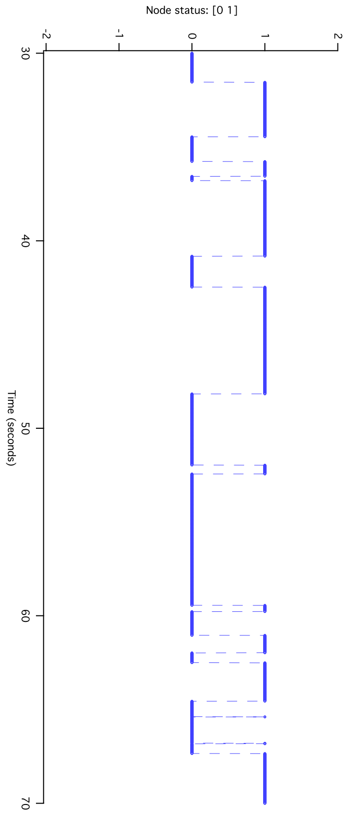


Figure C.1: Time series for node 69 in Run 69, zoomed in from 30s to 70s. Simulator parameter setup: time step is 0.01. Gong's matrix, infection rate is 0.1, curing rate is 0.5, initially infect 15 node. Simulated time is 1000s, time step is 0.01s, 100 runs.

Figure C.1 shows node 69 changes status frequently in this zoomed in time period. From the exact data, we know that this node is infected from 51.97s to 52.43s. The duration of this period is 0.46s. So the 0.01s time step in our Simulator should be good enough (small enough) to (roughly) see the status change, considering both accuracy and programming speed. If we took a larger time step, for example, took data for every 1s, then we wouldn't see the node status changing within 1s, instead we would see the averaged node status in this 1s, accuracy is worse.

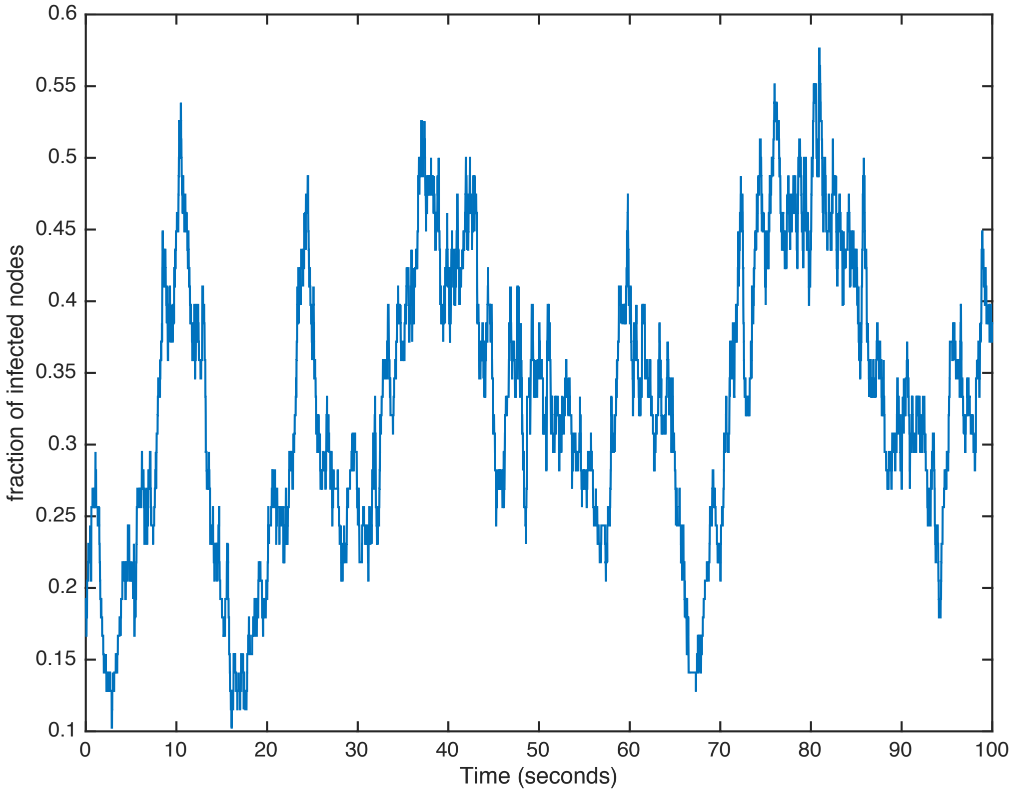


Figure C.2: Fraction of infected node in 1 run (Run 68)

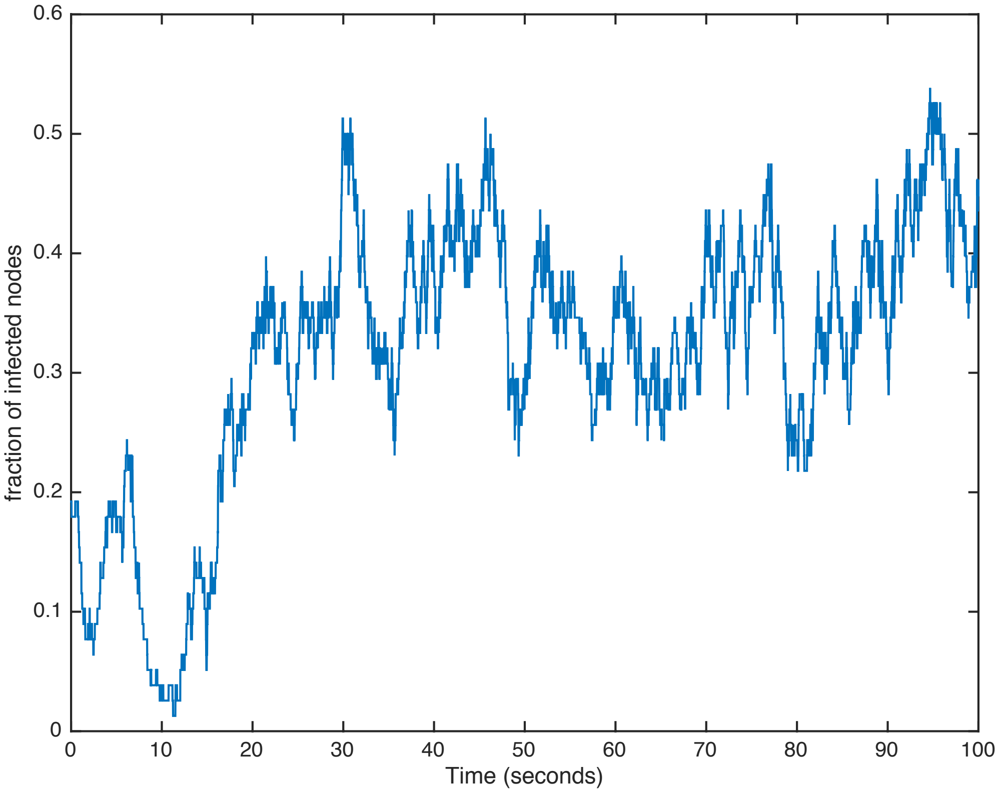


Figure C.3: Fraction of infected node in 1 run (Run 69), this run has a last node which was immuned of die out

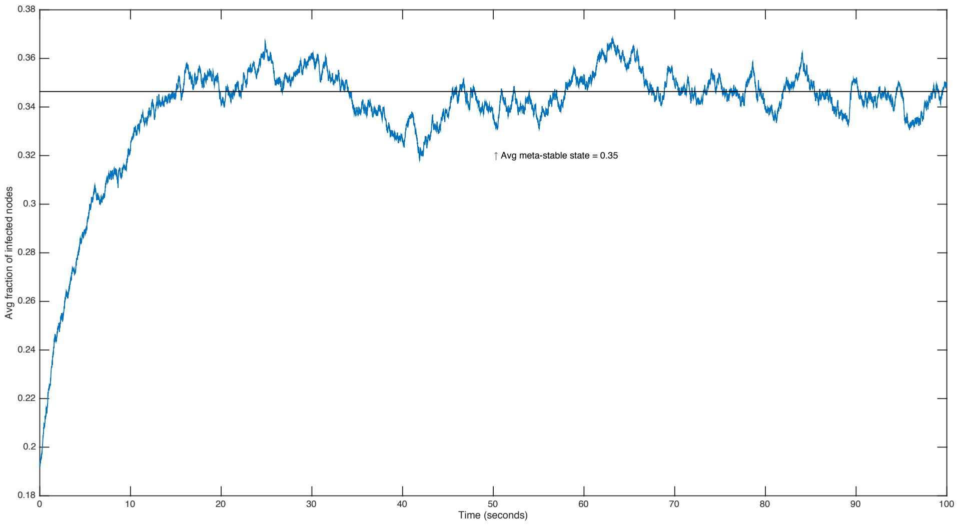


Figure C.4: Fraction of infected nodes averaged over 100 runs

ER network initial explorations

The infection fraction is also checked for ER network, generated with p=0.1096 which is calculated according to gong's matrix. Other settings are the same.

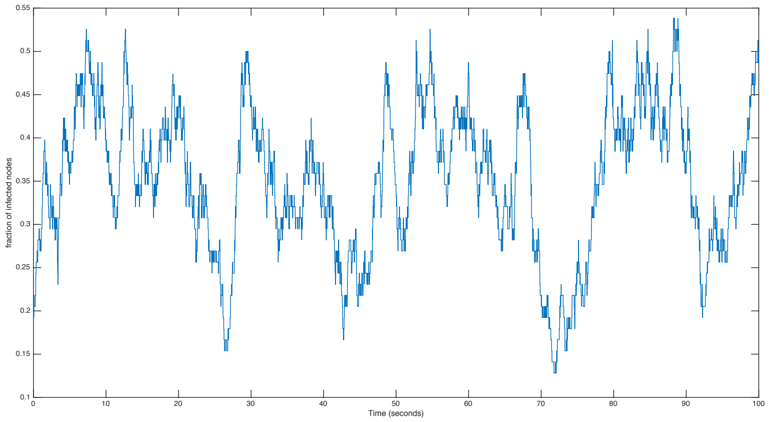


Figure C.5: ER Fraction of infected node in 1 run (Run 6)

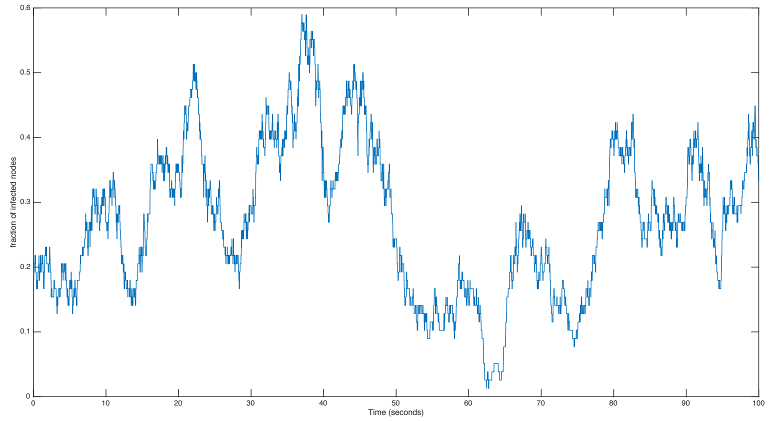


Figure C.6: Fraction of infected node in 1 run (Run 7), this run has a last node which was immuned of die out

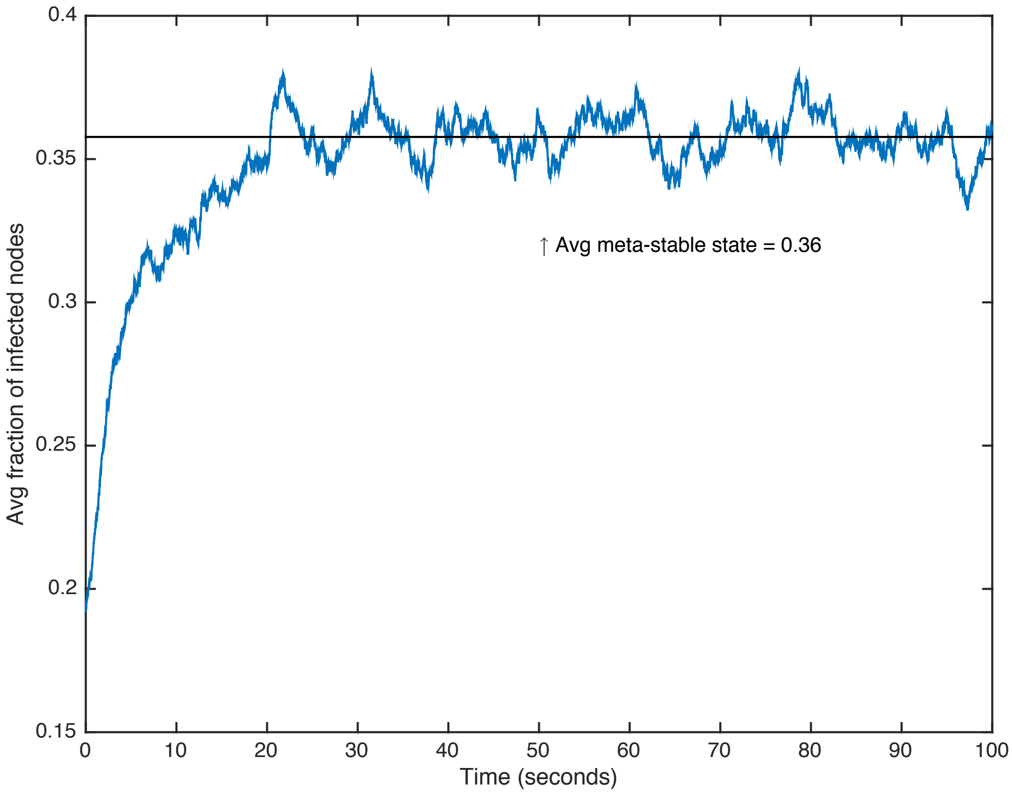


Figure C.7: ER Fraction of infected nodes averaged over 100 runs

C.2 More Confirmed Plots

In this section, the simulation graphs serve the purpose of verifying the results from Prof. Stam's paper in our continuous-time simulator. The equations and methods simulated are exactly the same as stated in Prof. Stam's paper. here we clarify two important concept and equations.

Functional Connectivity

The "functional connectivity" is defined as the Pearson correlation coefficient of the integrated time series.

For a node activation time series $X_i(t)$, the node states are represented in binary format as $X_i(t) = 1$ if node i is activated or infected, and $X_i(t) = 0$ if excitable. Take integration window $w = 10$, the integrated time series becomes

$$X_i(n) = \frac{1}{w}\sum_{k=1}^{w} X_i(n+k)$$

Then calculate the Pearson correlation coefficient for the integrated time series $X_i(n)$ by

$$\rho(X_i(n), X_j(n)) = \frac{E[X_i(n)X_j(n)] - E[X_i(n)]E[X_j(n)]}{\sqrt{Var[X_i(n)]}\sqrt{Var[X_j(n)]}}$$

Effective Connectivity

The "effective connectivity" is used to measure the causal flow between pairs of nodes by computing the probability that node j is activated at time $t+h$ when node i was activated at time t and vice versa (h is the time lag, take $h = 1$ in Stam's.) Since we have undirected graph,

$$C_{eff} = Pr[X_j(t+h) = 1|X_i(t) = 1] + Pr[X_i(t+h) = 1|X_j(t) = 1]$$

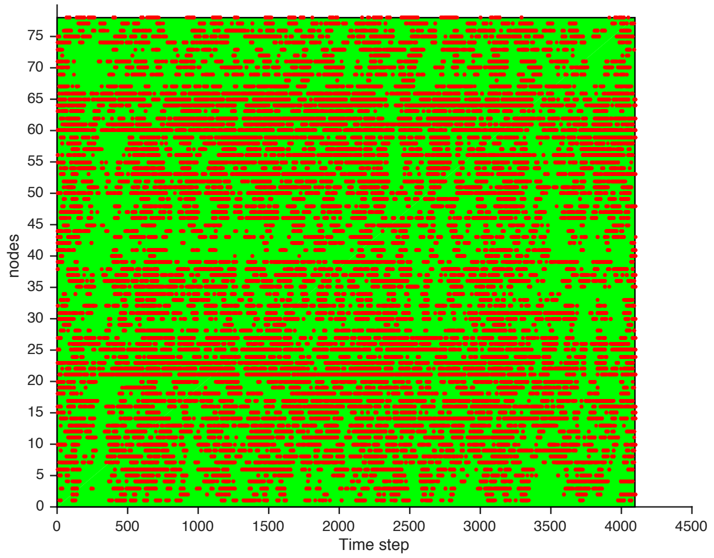


Figure C.8: File 'Gong.txt' . Node Status. N=78 `p=0.1096` I0=15 beta=0.1 delta=0.5 Run=1 Time=420s Frame=10 `E(D)=8.4` `maxEig=10.47`

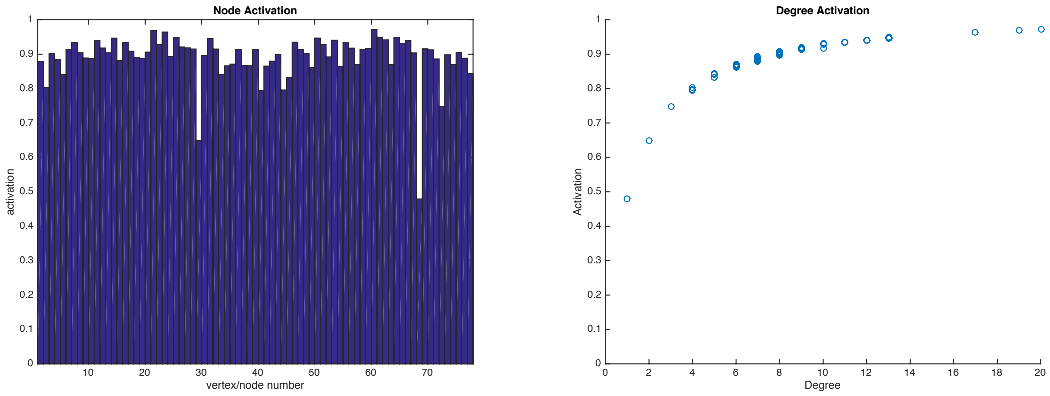
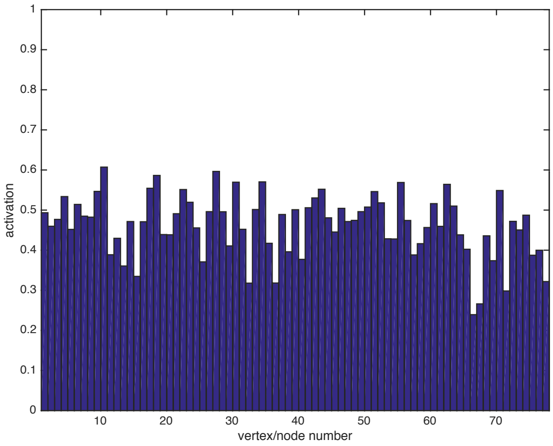
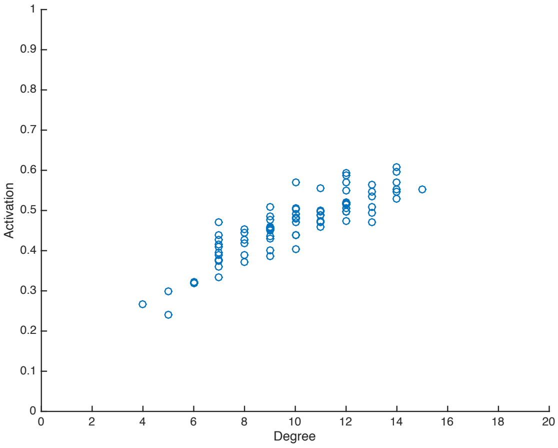


Figure C.9: Activation time as function of nodes an degrees with $\beta = 0.5$ ($\delta = 0.5$), resulting in an effective spreading $\tau = 1.0$. Comparing to Prof. Stam's work in discrete-time simulation, the max value is different. We found the continuous-time simulator often give higher max value

ER



(a) Node vs Activation. Activation time of every node during the simulation period.



(b) Degree vs Activation.Similar to Stam's.

Figure C.10: ER random graph

C.3 Delayed Correlations

Initial Explores

Delayed SIS Correlation - 1-1

N=78 `p=0.1096` I0=15 beta=0.1 delta=0.5
Run=1 Time=420s Frame=10 `E(D)=8.4` `maxEig=10.47`

Source files: Gong's matrix. GR1T420F10B1V1.txt.

Correlation with time lag l for all node pairs

Since Gong's matrix has N=78 nodes, so we have $\binom{78}{2} = 3003$ node pairs. Take time lag l, the correlation of all node pairs is displayed in Figure C.18 . There is 3003 dots on the graph.

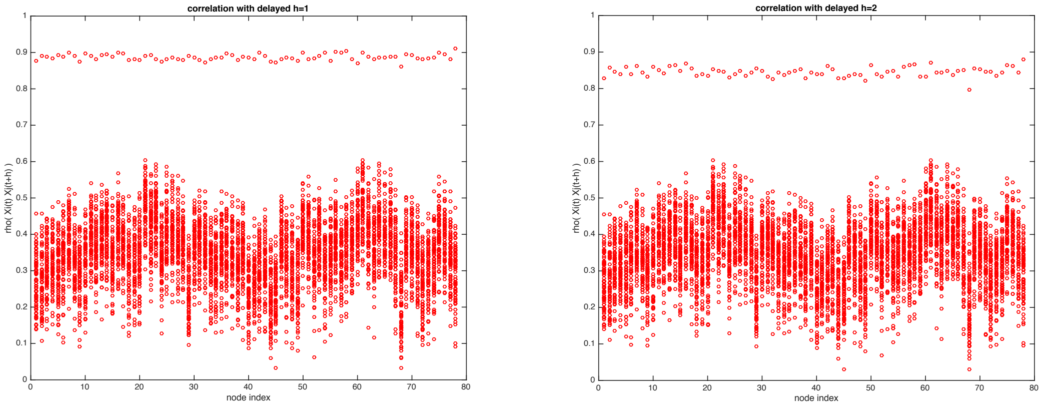


Figure C.11: Correlation of all node pairs at time lag l

Correlation of one node to all other nodes as function of time lag

For each node, it forms 78 node pairs to each of other nodes. Take node 1 for example, we have node pair 1-2, 1-3,...,1-78. Correlation of these node pairs versus different time lags (0.1s to 10s) is shown in Figure C.27. The circled line is the correlation between node 1 and its neighbours.

Since there are altogether 78 lines on Figure C.27, it is not easy to see the details. So we plot correlation with node i's neighbours and without neighbours separately in Figure C.28.

Fitting slope of Correlations versus time lag - node 1

For each correlation line in previous graph, we have a fitting line and we noted down the slope. In our example, the slope of all 78 node pairs of node 1 is then plotted in Figure C.29 on which the neighbours of node 1 is marked by *. Each point is the slope of node pair 1-1, 1-2, ..., 1-78.

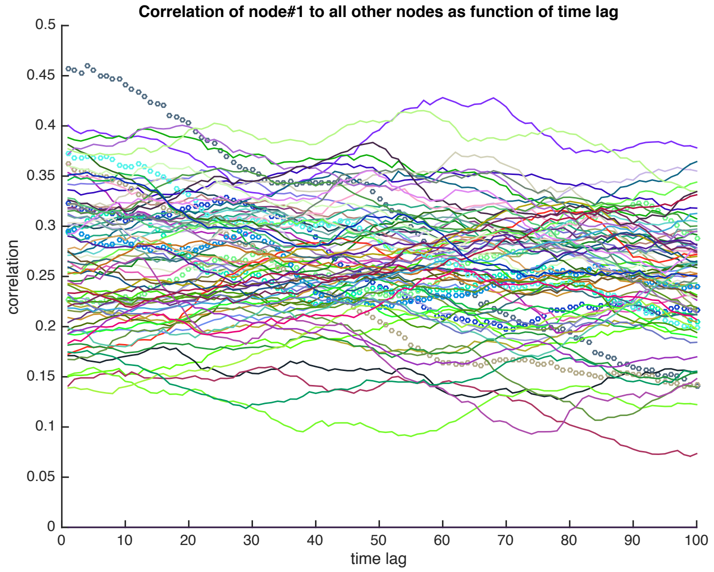


Figure C.12: Correlation of node 1 to all nodes versus time lag

The graph shows slope fluctuates around 0, so we cannot conclude the correlation is enhancing or decreasing as a function of time lag.

Figure C.32 shows all 3003 slope values, they also fluctuates around 0.

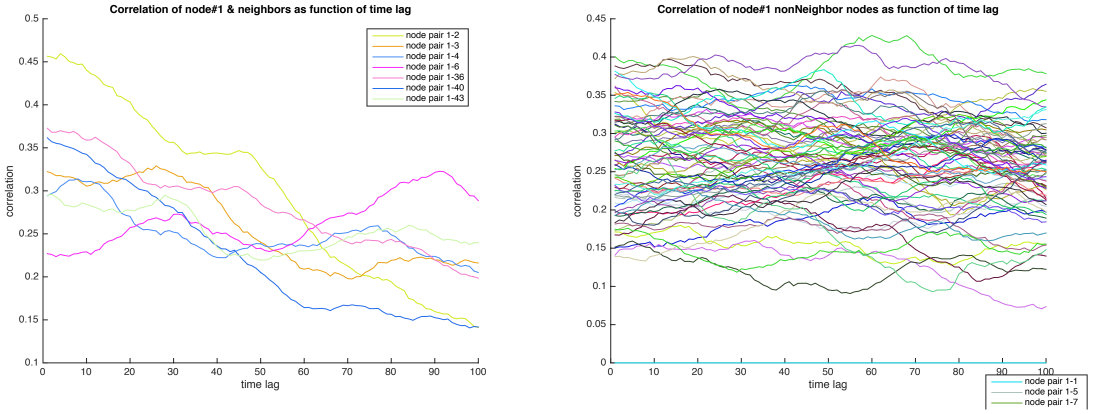


Figure C.13: Correlation of node 1 as function of time lag 0.1 10s

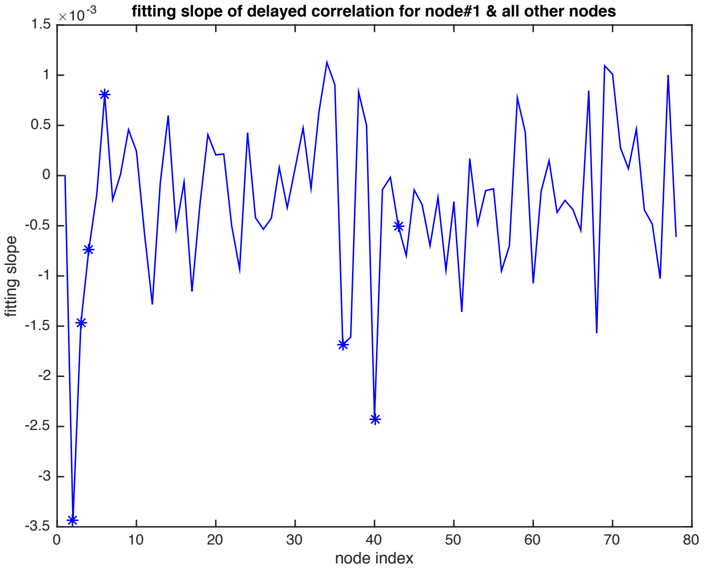


Figure C.14: Slope of node 1 to all nodes

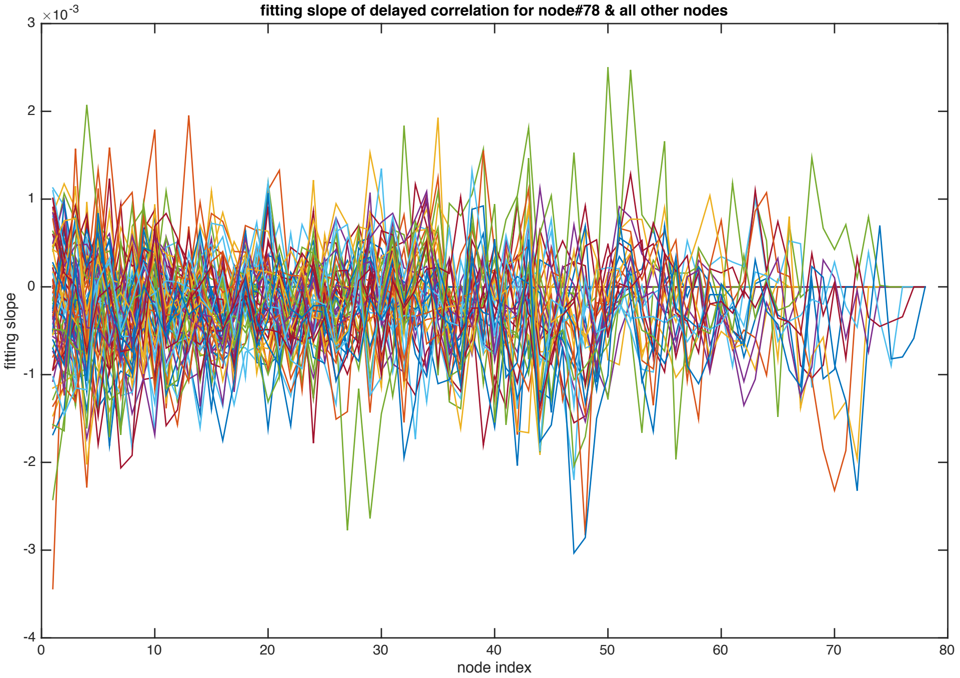


Figure C.15: Slope of all nodes

Fitting slope of avg correlations for each node versus time lag

In previously Figure C.29, we take the average of this line, i.e. the average slope of node 1 to all 78 nodes. For each node, we can plot such a graph as in Figure C.29. Therefore, for each node, we have a averaged slope. This averaged slope is plotted in Figure C.30, and with and without neighbours is plotted separately in Figure C.31.

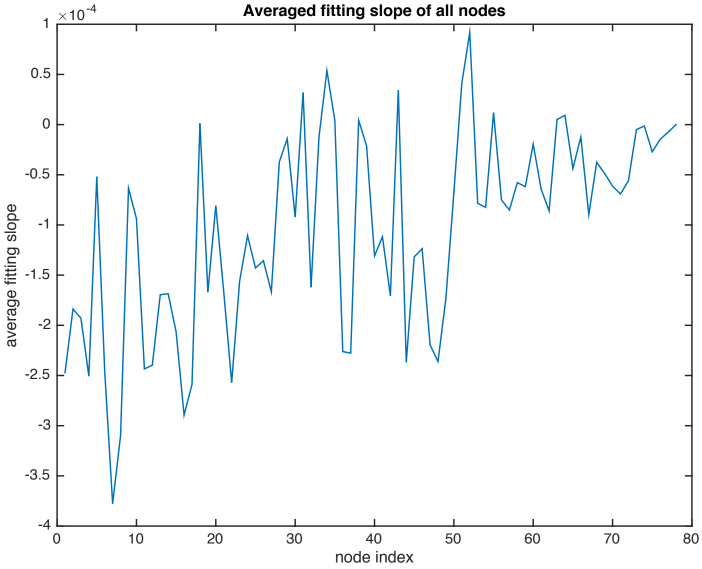


Figure C.16: Avg slope of each node

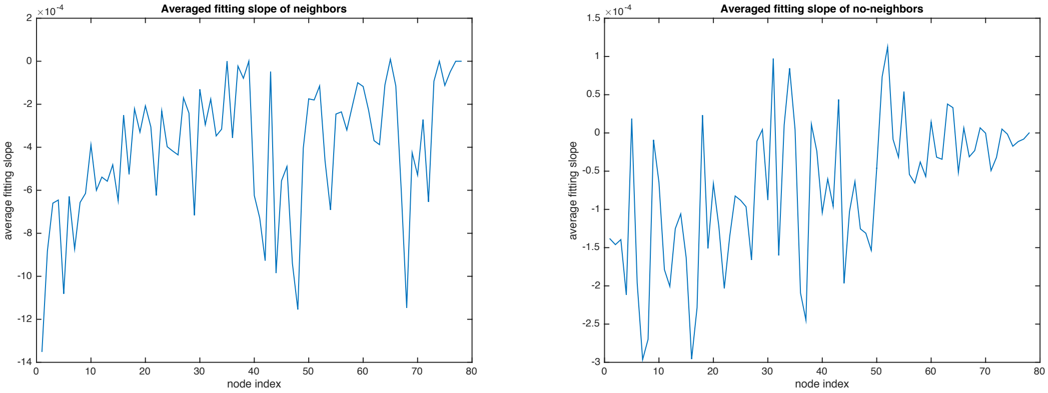


Figure C.17: Avg slope of each node

Delayed SIS Correlation - 1-2

N=78 `p=0.1096` I0=15 beta=0.1 delta=0.5
Run=1 Time=420s Frame=10 `E(D)=8.4` `maxEig=10.47`

Source files: Gong's matrix. GR1T420F10B1V1.txt.

Correlation with time lag l for all node pairs

Since Gong's matrix has N=78 nodes, so we have $\binom{78}{2} = 3003$ node pairs. Take time lag l, the correlation of all node pairs is displayed in Figure C.18 . There is 3003 dots on the graph.

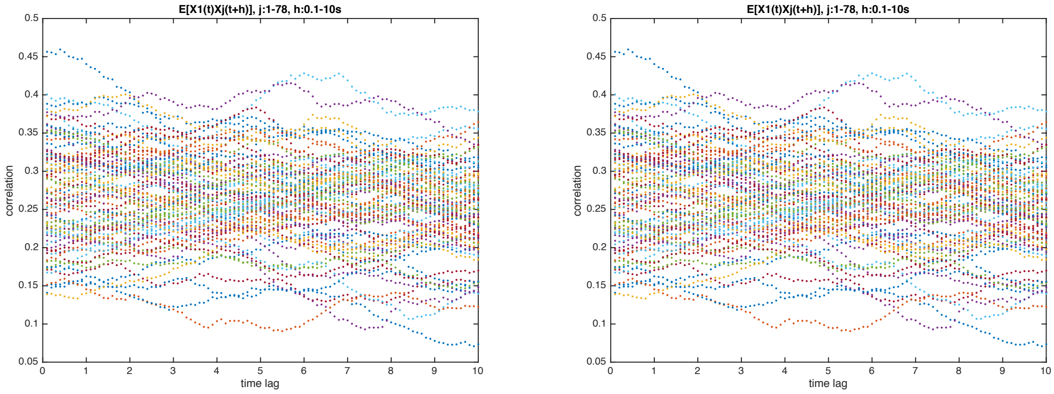


Figure C.18: Correlation of all node pairs at time lag l

Correlation of one node to all other nodes as function of time lag

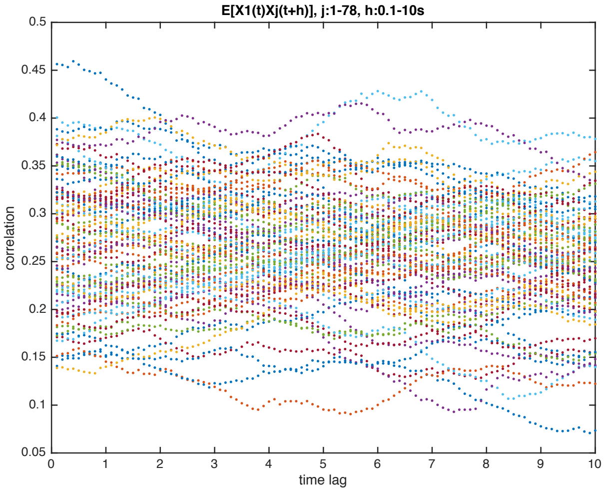


Figure C.19: Correlation of node 1 to all nodes versus time lag

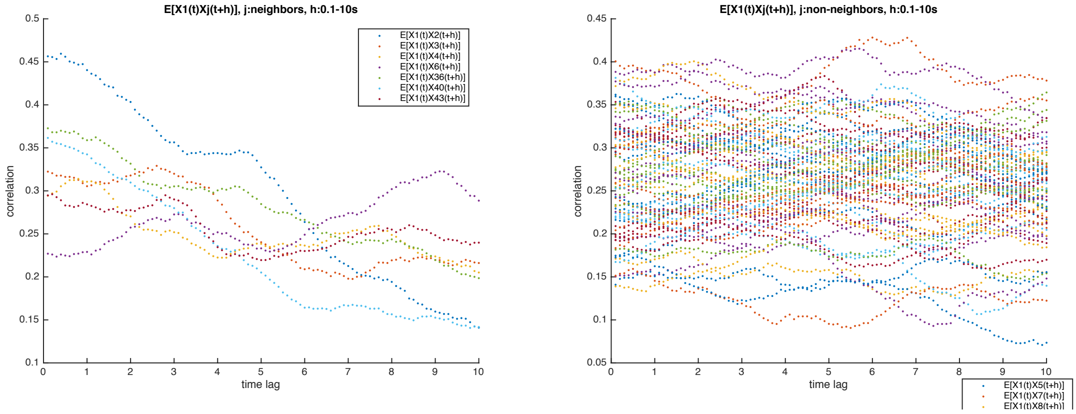


Figure C.20: Correlation of node 1 as function of time lag 0.1 10s

Fitting slope of Correlations versus time lag - node 1

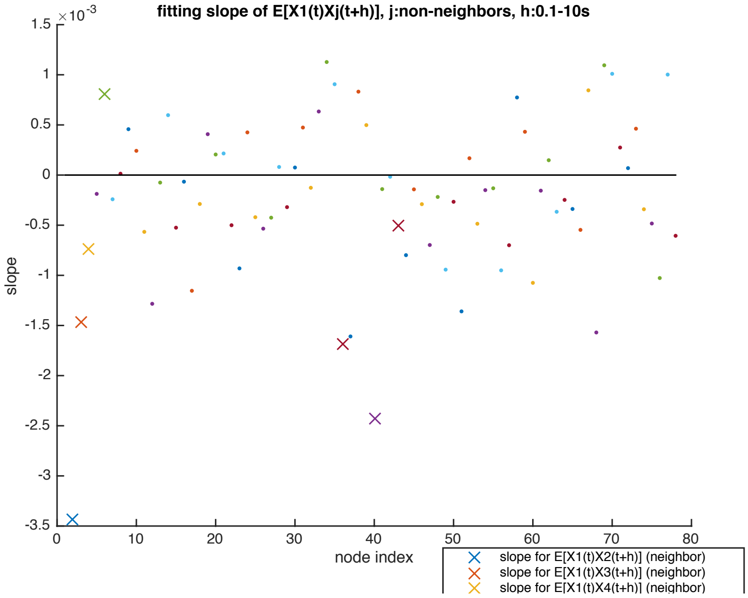


Figure C.21: Slope of node 1 to all nodes

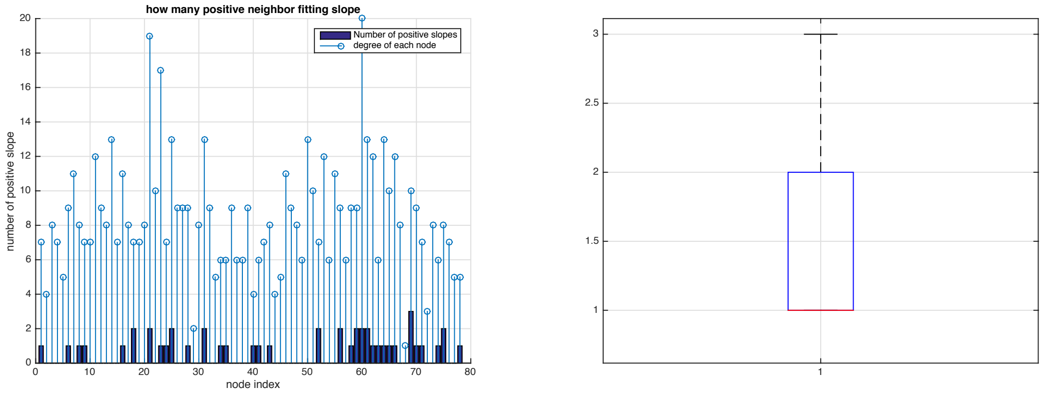


Figure C.22: slope of all nodes over u=0.1-10s

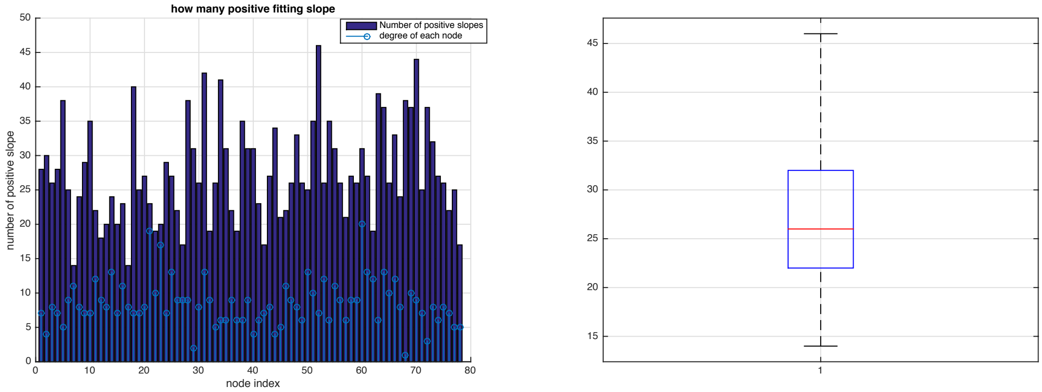


Figure C.23: slope of all nodes over u=0.1-10s

Zoom into u=0.1-1s :

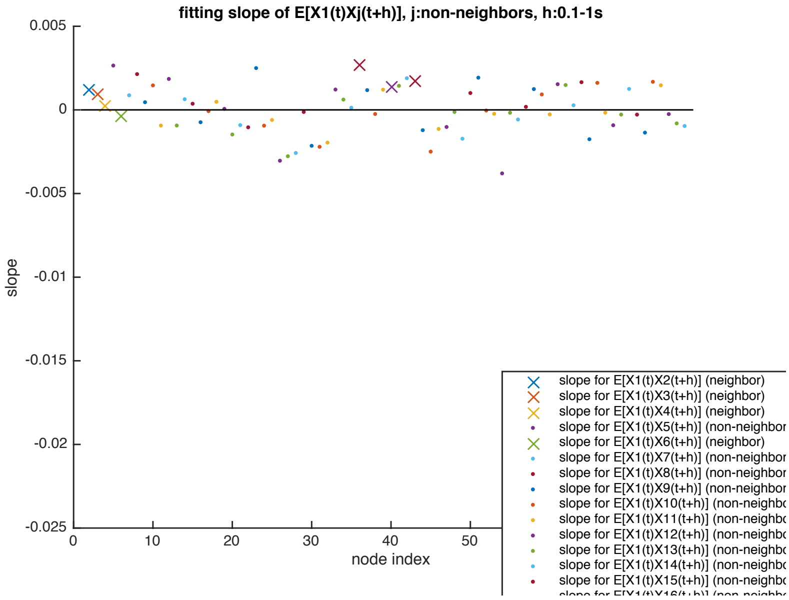


Figure C.24: Slope of node 1 to all nodes

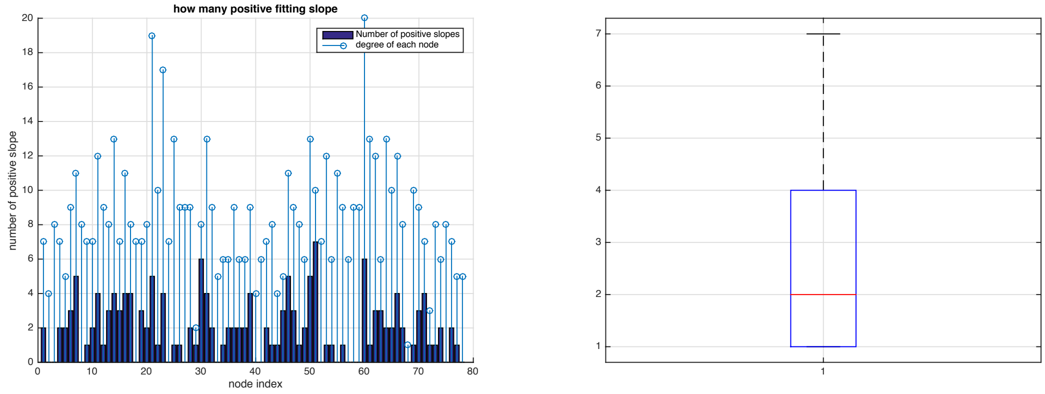


Figure C.25: slope of all nodes Zoom into u=0.1-1s

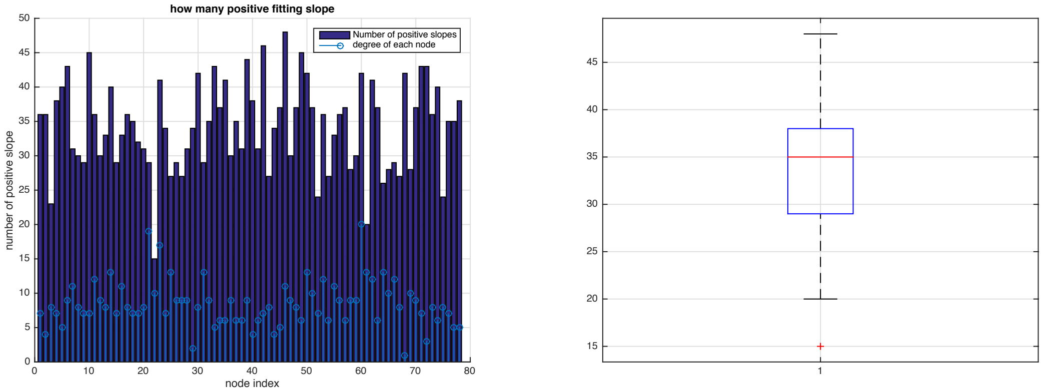


Figure C.26: slope of all nodes Zoom into u=0.1-1s

Delayed SIS Correlations - 1-3

Correlation of one node to all other nodes as function of time lag

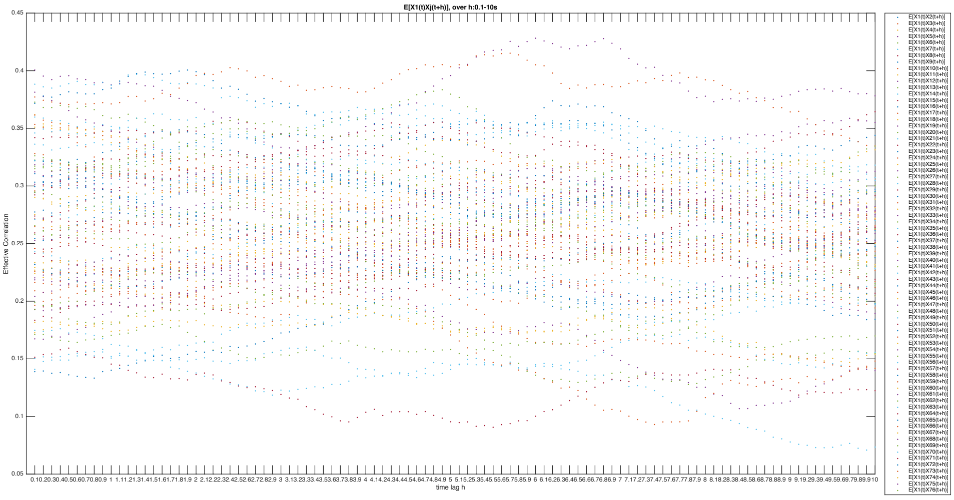


Figure C.27: Correlation of node 1 to all nodes versus time lag

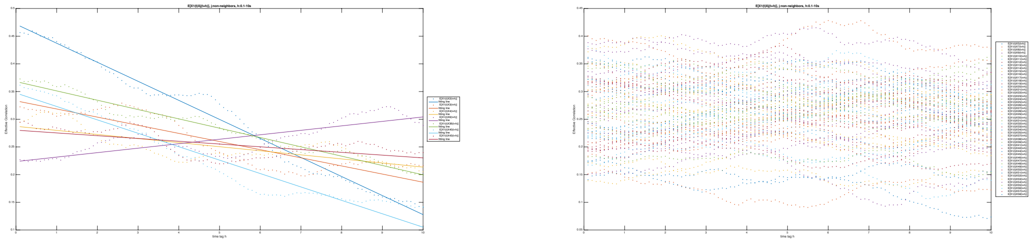


Figure C.28: Correlation of node 1 as function of time lag 0.1 10s

Fitting slope of Correlations versus time lag - node 1

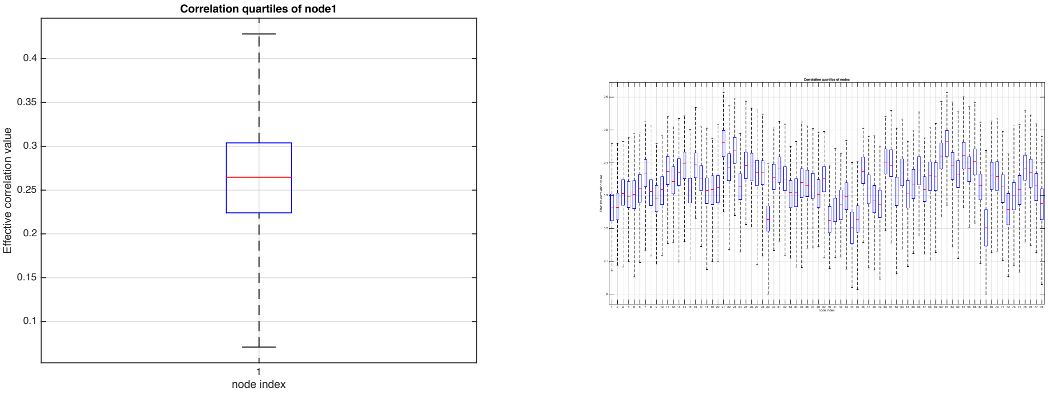


Figure C.29: boxplot

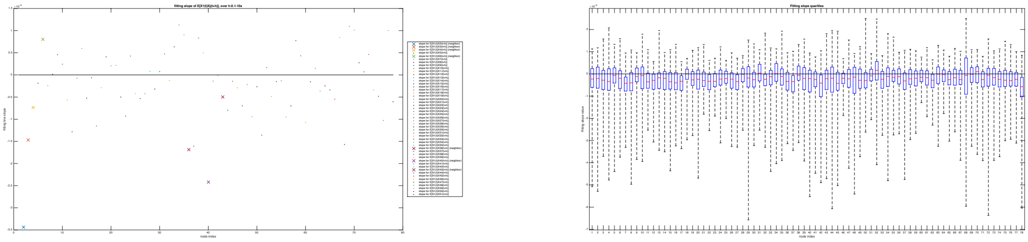


Figure C.30: fitting slope of node 1 to all 78 nodes

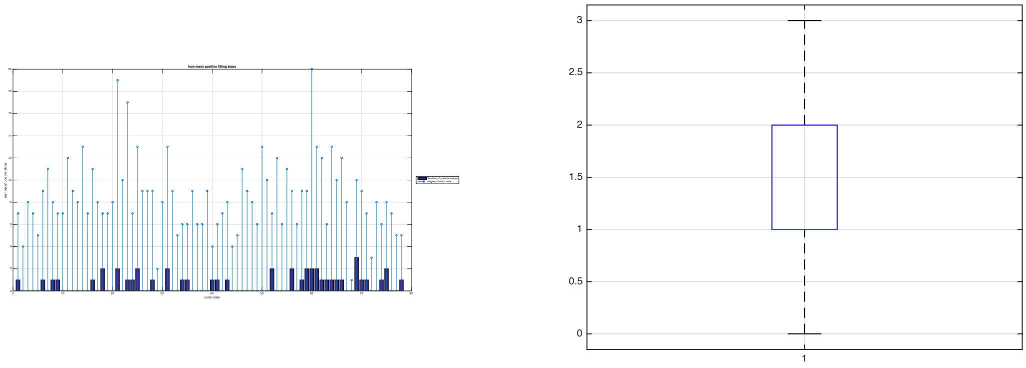


Figure C.31: number of positive slopes

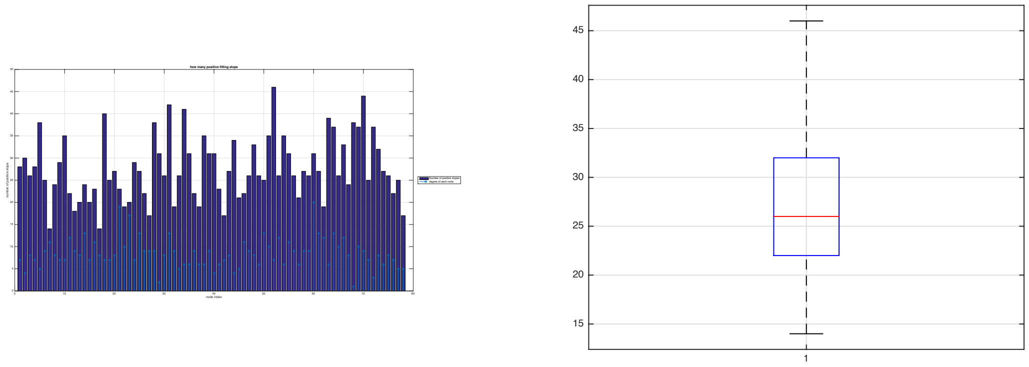


Figure C.32: number of positive slopes

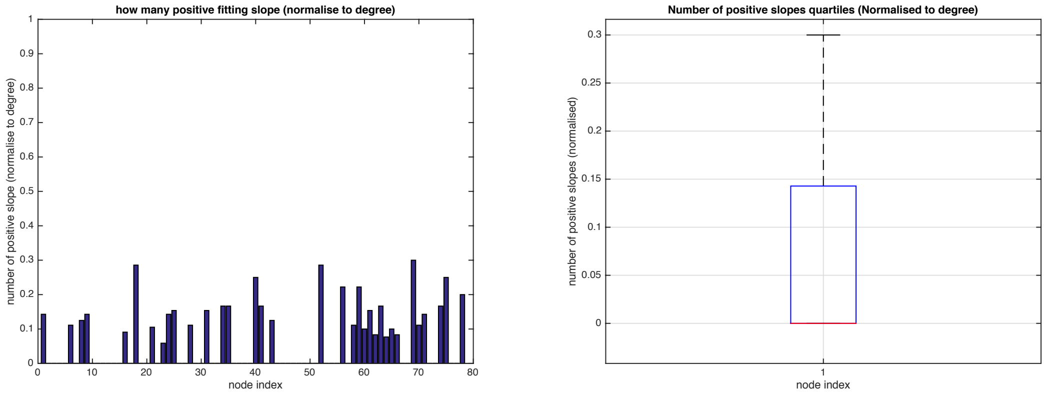


Figure C.33: normalised number of positive slopes

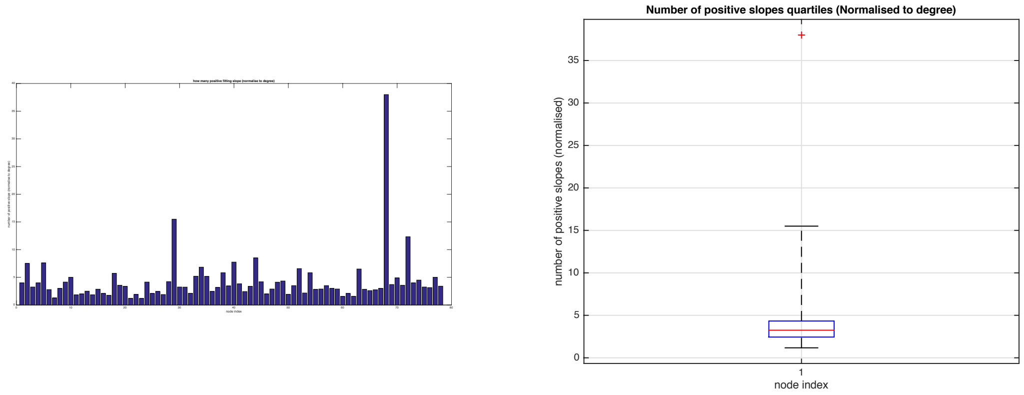


Figure C.34: normalised number of positive slopes

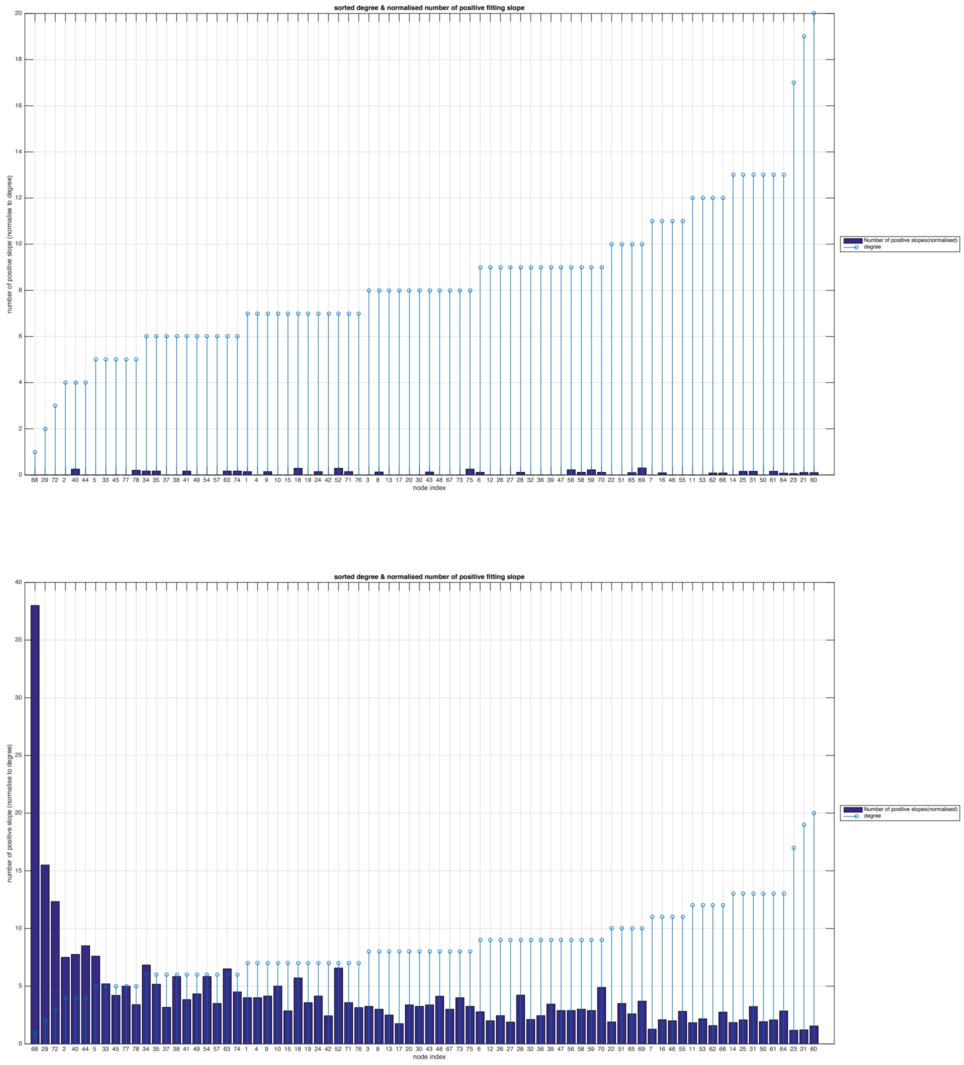


Figure C.35: sorted normalised number of positive slopes

Supplimentary Figures

In this section, the degree of all 78 nodes is presented in a sorted order and delayed correlations of nodes with an average degree or lowest degree are provided.

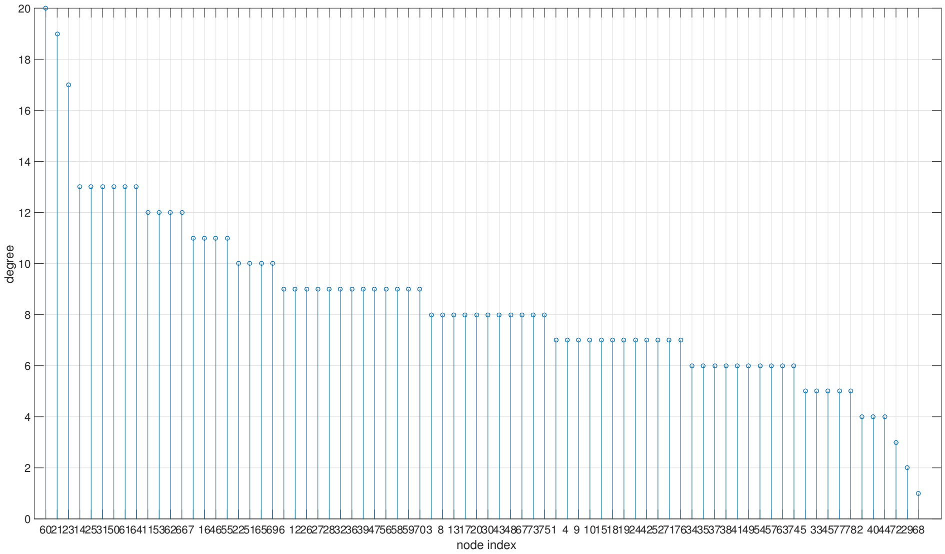


Figure C.36: Sorted degree of all 78 nodes in the structural brain network.

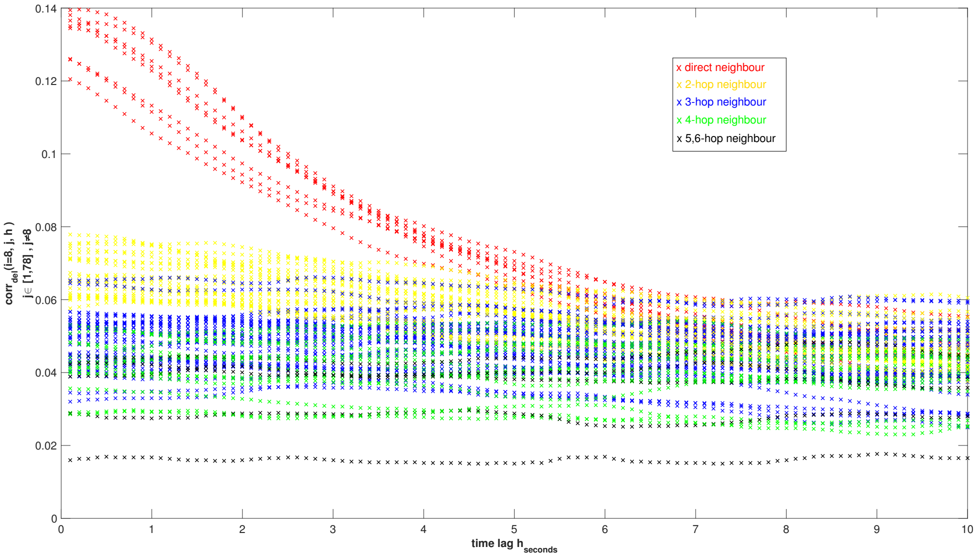


Figure C.37: Example of delayed correlations of node 8 to all other nodes. Node 8 has the average degree.

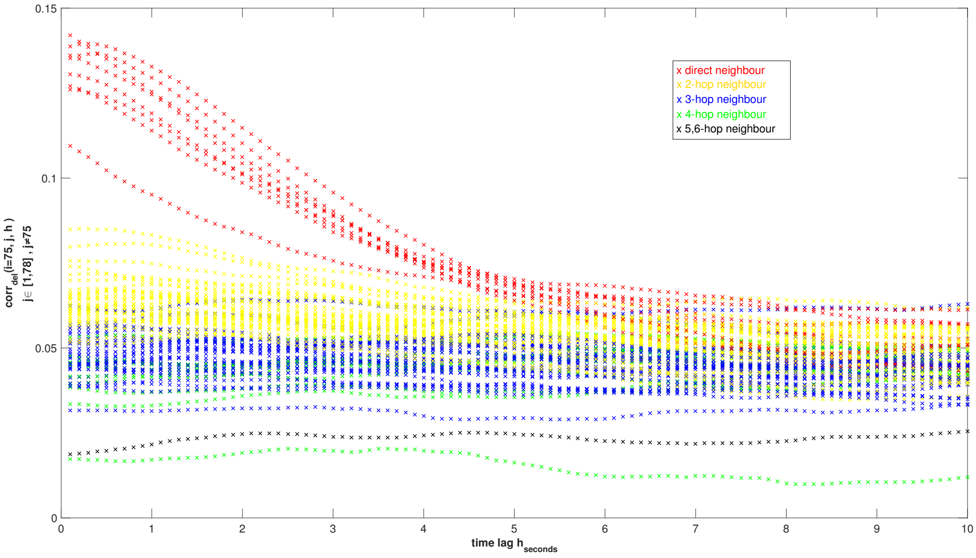


Figure C.38: Example of delayed correlations of node 75 to all other nodes. Node 75 has the average degree.

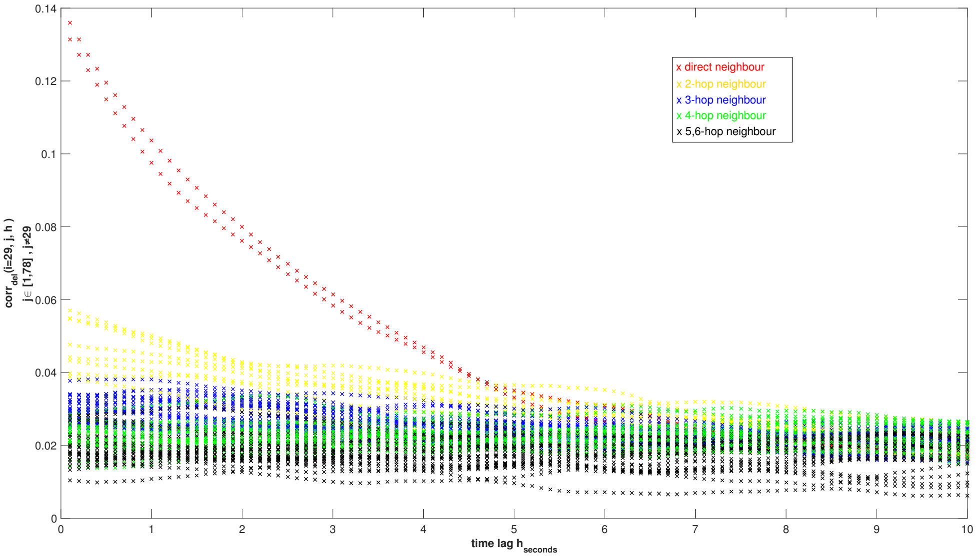


Figure C.39: Example of delayed correlations of node 29 to all other nodes. Node 29 is one of the three nodes of lowest degree.

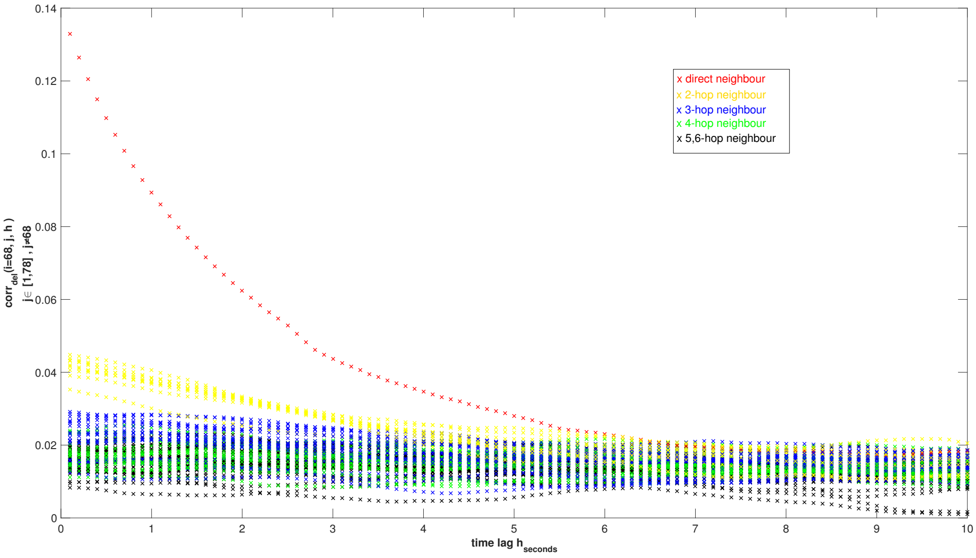


Figure C.40: Example of delayed correlations of node 68 to all other nodes. Node 68 is one of the three nodes of lowest degree.

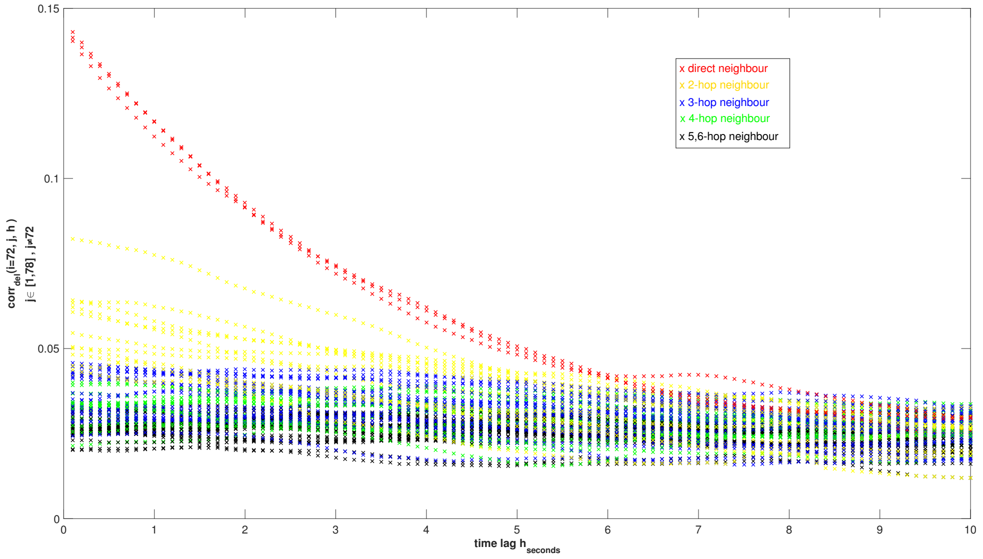


Figure C.41: Example of delayed correlations of node 72 to all other nodes. Node 72 is one of the three nodes of lowest degree.

Appendix D

Scientific manuscript about the work

The major contents of the thesis work is included in the scientific manuscript "Applying the Epidemic Spreading Model to Explain Brain Activity". The manuscript was submitted to Conference on Complex Systems (CCS) on May 13, 2016 and was accepted. The manuscript will be presented at CCS on September 22, 2016 in Amsterdam. The full manuscripted is attached for reference.

Applying the Epidemic Spreading Model to Explain Effective Connectivity in Brain Networks

J. Meier[*], X. Zhou[†], A. Hillebrand[‡], P. Tewarie[§], C.J. Stam[¶] and P. Van Mieghem[‖]

August 24, 2016

Abstract

Explaining how brain regions communicate with each other given the underlying anatomical connections still remains an open research question. A recent study observed a dominant posterior-anterior pattern of information flow in empirical data using the measure of transfer entropy. In this study, we apply a simple SIS epidemic spreading model on the human connectome to analyze the structural topological properties that drive this overall directionality using a continuous-time simulator. In order to analyze the influence of one brain region on the other, we analyze the transfer entropy values for all node pairs under different time delays. We find that just above the critical threshold direct structural connections induce higher transfer entropy between two brain regions and that transfer entropy decreases with increasing distance between nodes (in terms of hops in the structural network). Hubs seem to play a special role in the network dynamics and we show that they send more information to the network than that they receive. Using the measure of transfer entropy, we can confirm the previously empirically observed dominant back-to-front pattern with our SIS model both with the help of analytic derivations and continuous-time simulations. We also show that the posterior hubs seem to be responsible for this observed directionality spreading pattern emerging from an underlying undirected structural network. To sum up, our analysis shows that the posterior-anterior pattern of information flow seems to appear as a result of differences in the spatial distribution of the structural degree.

---

[*] *Email*: J.M.Meier@tudelft.nl, Delft University of Technology

[†] *Email*: x.zhou@student.tudelft.nl, Delft University of Technology

[‡] *Email*: a.hillebrand@vumc.nl, Department of Clinical Neurophysiology and Magnetoencephalography Center, VU University Medical Centre, Amsterdam, The Netherlands

[§] *Email*: Prejaas.Tewarie@nottingham.ac.uk, University of Nottingham, UK

[¶] *Email*: CJ.Stam@vumc.nl, VU University Medical Center, Department of Clinical Neurophysiology

[‖] *Email*: P.F.A.VanMieghem@tudelft.nl, Delft University of Technology

1 Introduction

Many interesting properties of the human brain have been discovered by analyzing the brain as a network (Stam, 2014, Stam and Van Straaten, 2012, Bullmore and Sporns, 2012). However, different measurement techniques capture different aspects of brain networks. Methods such as diffusion tensor imaging (DTI) allow for the reconstruction of the structural brain network, which consists of a map of anatomical connections between brain regions. Functional imaging techniques, such as magnetoencephalograpy (MEG), electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), can measure the activity or communication between brain regions, from which we can extract the functional brain networks. Two different types of connectivities can be computed based on the brain regions' activation series: functional connectivity refers to the correlation of activities between brain regions; effective connectivity tries to capture the causal effect of one region's activity to the other regions' activities (Aertsen et al., 1989, Friston, 1994). Whereas most of the studies so far analyzed functional connectivity, recent approaches have been focusing on effective connectivity to gain knowledge about directionality in functional brain networks (Hillebrand et al., 2016, Moon et al., 2015, Stam and van Straaten, 2012). Patients suffering from brain disorders can have altered anatomical brain connections. In order to anticipate the different brain dynamics resulting from those changes in the structural networks, we need to understand the properties of the underlying structural human connectome that facilitate the communication processes in the functional networks.

To estimate the effective connectivity between two brain regions based on their activation time series, the measure of transfer entropy (TE) is often used for MEG or EEG (Schreiber, 2000). TE from node A to node B measures how much better a prediction of a next value of B gets when we not only include the previous value of B but also the previous value of A. TE can be interpreted as a delayed correlation measure that is corrected for auto-correlation (see Appendix C). In the sense of Wiener's principle (Wiener, 1956), TE can be interpreted as the causal influence from one brain region on the other. Recently, the measure of TE has been expanded to a measure for phase-based connectivity, the so-called Phase Transfer Entropy (PTE) (Paluš and Stefanovska, 2003, Lobier et al., 2014). Using PTE, (Hillebrand et al., 2016) recently found a surprisingly consistent overall spreading pattern from posterior to anterior brain regions in empirical data in higher frequency bands (*alpha1*, *alpha2*, and *beta* band). (Hillebrand et al., 2016) hypothesized that this global direction of information flow could be explained by strong hub connections in the posterior regions and the fact that hubs possess the highest levels of neuronal activity in the network (de Haan et al., 2012, Moon et al., 2015).

Recent studies showed that simple models of activity spread contribute to our understanding of brain dynamics (Abdelnour et al., 2014, Deco et al., 2012). Using a simple deterministic cascade model (Mišić et al.,

2015) revealed that structural hubs and the shortest path structure have a high influence on the efficiency of spreading dynamics of the brain network. Even though those simple models ignore microscopic details of the spreading process, the scarcity of parameters allows for direct analysis of the global spreading patterns. Further, there is evidence that the brain operates near a critical phase transition (Haimovici et al., 2013, Rubinov et al., 2011, Yu et al., 2013). In addition, it is known from statistical physics that the details of the applied model become irrelevant near such a phase transition which could be another explanation why simple models have been successful in capturing more complicated model findings (Honey et al., 2007, Honey et al., 2009).

A dynamic spreading process on networks is often approximated by a simple epidemic process, which has been used for various applications, e.g. information propagation and gossip spreading in networks (Pastor-Satorras et al., 2015). The advantage of the epidemic spreading model is that the only a-priori chosen parameter is the effective spreading rate τ and that we can study the model also analytically. Here, we focus on a Susceptible-Infected-Susceptible epidemic (SIS epidemic) as one of the simplest models of an epidemic. In the case of the functional brain network, brain regions can be activated (infected) and spread this activation to their anatomically neighboring regions. In an SIS epidemic process, a node can be in two states, either activated or excitable (and can be activated by one of its neighbors). The SIS process can be described as a continuous-time Markov chain with 2^N states where N is the number of nodes in the network. The embedded Markov chain can be used to transform the continuous-time SIS process into a discrete-time process. A previous study (Stam et al., 2016) applied a discrete-time epidemic process on the structural brain network and identified the degree product as a driving force for the effective connectivity between two nodes. Applying the well-developed theory of epidemics may lead to a better understanding of the activity spreading in the brain and in particular reveal the structural properties that drive the global spreading dynamics. However, the real underlying SIS process is per definition a continuous-time process, thus any discrete-time process will always be an approximation. The embedded discrete-time Markov chain then contains the transition probabilities but no longer the precise timing of the events (Van Mieghem, 2014b). Thus, by simulating the SIS epidemic as a discrete-time process we miss the smaller-scale dynamics in which we are especially interested in for this study (Van Mieghem, 2014b). In this paper, we simulate an epidemic spreading process on the structural brain network in a continuous-time framework which is closer to reality (see Figure 1). Since the activation time series of our SIS model are binary and thus a phase-based measure is not applicable, we will use the TE for our computations of pairwise interactions.

The aim of this paper is to reveal the topological properties of the structural brain network that drive the empirically observed effective connectivity in the brain network by applying a continuous-time SIS model.

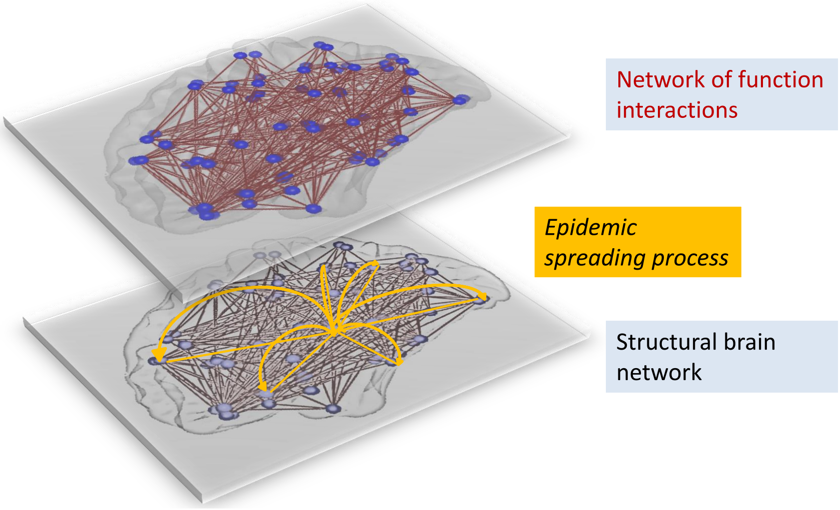


Figure 1: Visualization of the approach taken in this study: we run a dynamic SIS process on the underlying structural network of the brain and then calculate the transfer entropy between the time series of the different nodes to construct a network of the same nodes but with the link weights representing the function interactions between node activities.

For that, we first analyze different effective spreading rates τ as the only model parameter and compare the resulting pairwise dynamics to empirical data. Then, we relate different structural properties of the underlying network like degree, eigenvector and betweenness centrality to the SIS spreading process and the pairwise transfer entropies using simulations and analytic reasoning.

2 Methods

Structural Network

For the structural network, we used a literature-based DTI network from a previous study based on 80 healthy subjects (for details see (Gong et al., 2009)). In short, for every individual two cortical regions from the 78 cortical automated anatomical labeling (AAL) brain areas were considered to be connected if the end points of two white matter tracts were located in these regions(Gong et al., 2009). Via a non-parametric sign test only the significant links were included in the group-averaged structural connectivity matrix. This processing resulted in a binary connectivity matrix for the structural brain network (see Figure 6(a)), which we will further refer to as the structural adjacency matrix A.

Susceptible-Infected-Susceptible Process

The Susceptible-Infected-Susceptible (SIS) epidemic process is one of the simplest dynamic processes on a network. In an SIS epidemic process, on an undirected and unweighted graph G with N nodes and L links, the state of a node i at time t is specified by a Bernoulli random variable $X_i(t) \in \{0, 1\}$: $X_i(t) = 0$ for an excitable node and $X_i(t) = 1$ for an activated node. A node i at time t can be in one of the two states: activated, with probability $w_i(t) = \Pr[X_i(t) = 1]$ or excitable, with probability $1 - w_i(t)$. We replaced here the states healthy and infected from classic epidemic theory with the names excitable and activated which fits better to the status of a brain region as a node. Here, we consider a continuous time model which can formally be defined by the following differential equation

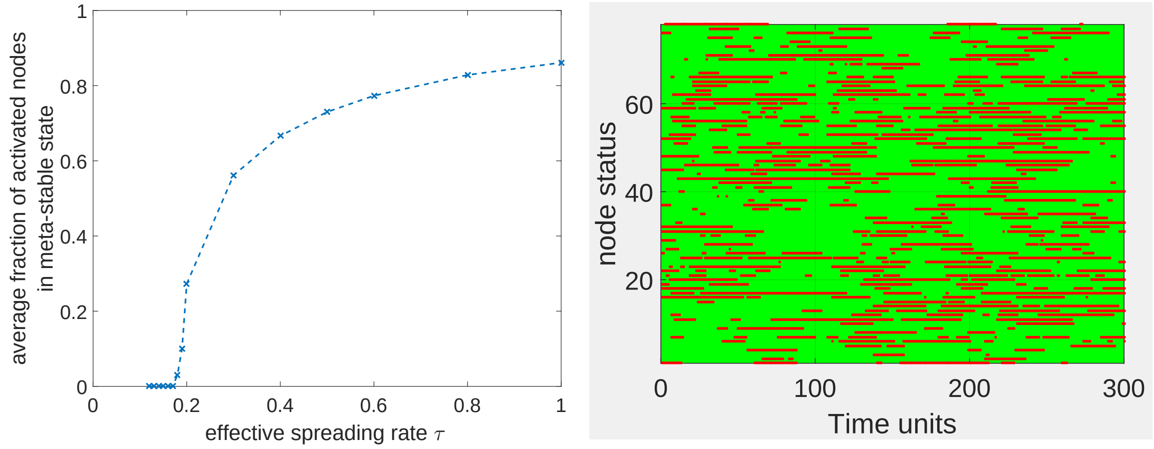
$$\frac{dX_j}{dt} = -\delta X_j + (1 - X_j)\beta \sum_{k=1}^{N} a_{kj} X_k, \tag{1}$$

where we assume that the deactivation process (which means the process of changing from an activated to an excitable status) per node i is a Poisson process with rate δ (Van Mieghem, 2014b). Further, the activation rate per link is a Poisson process with rate β, where $\sum_{k=1}^{N} a_{kj} X_k$ counts the number of infected neighbors of node j. The deactivation and activation process are identical with the curing and infection process from epidemic theory (Van Mieghem, 2014b). Obviously, only when a node is activated, it can activate its direct neighbors, that are still excitable. Both the deactivation and activation Poisson process are independent. The effective infection rate is defined by $\tau = \frac{\beta}{\delta}$. There is evidence that the brain operates with its dynamics near a critical phase transition (Haimovici et al., 2013, Rubinov et al., 2011, Yu et al., 2013). Thus, we follow the previous study and continue to choose our parameters such that the SIS dynamics are near the critical epidemic threshold (Stam et al., 2016).

Details of the simulation

We used the continuous-time simulator SIS simulator (SISS) (van de Bovenkamp, 2015) to simulate an SIS-epidemic on the structural network with $\beta = 0.1$ and $\delta = 0.5$ similar to Stam et al.'s previous simulations(Stam et al., 2016) in order to compare our results with his approach in discrete time, which results in an effective spreading rate $\tau = \beta/\delta = 0.2$. This choice of parameters is slightly above the epidemic threshold τ_c (see Figure 2(a)). For every simulation run, we initially activated 15 random nodes, which is approximately 20% of the whole network (Stam et al., 2016), and all our results are averaged over 100 simulation runs.

We ran a simulation of 4096 time units similar to Stam et al. (Stam et al., 2016). We took 0.1 time units as a sample interval resulting in 40960 time points for one simulation, forming for each node an activation time series of zeros (node not activated at time instance t) and ones (node is activated at time instance t). An example activation series is displayed in Figure 2b (red is the activated state and green means susceptible). To focus on the metastable state, we disregarded the initial phase of the spreading process by calculating all our results based on the second half of the simulation time from 2048 to 4096 time units.


(a) Average fraction of activated nodes versus τ.

(b) Example of activation time series.

Figure 2: (a) Average fraction of activated nodes for different τ. (b) One example of an activation time series for 300 time units of the simulation for all 78 nodes of the underlying structural network of an SIS epidemic, where red means that the node is active and green that it is inactive/excitable (resulting in 3000 time steps, $\beta = 0.1$, $\delta = 0.5$).

Transfer Entropy

In order to capture the delayed influence, we calculate for every node pair i and j the transfer entropy defined as

$$TE_{i\to j}(h) = \sum_{k,l,m=\{0,1\}} \Pr[X_j(t+h) = k, X_j(t) = l, X_i(t) = m] \cdot \log\left(\frac{\Pr[X_j(t+h) = k | X_j(t) = l, X_i(t) = m]}{\Pr[X_j(t+h) = k | X_j(t) = l]}\right) \quad (2)$$

for a certain time delay h.

Motivated by a recent paper of Hillebrand et al. (Hillebrand et al., 2016), we analyze the sending or receiving property of a node. This property means the relation for a node i between $TE_{i\to j}$ and $TE_{j\to i}$ averaged over all other nodes j, and investigates which flow direction outweighs. Similar to (Hillebrand et al., 2016) we define the directed TE (dTE) for node i and j as

$$dTE_{i\to j}(h) = \frac{TE_{i\to j}(h)}{TE_{i\to j}(h) + TE_{j\to i}(h)}. \tag{3}$$

Since the TE can only take positive values, this definition of dTE is well-defined and its value ranges between 0 and 1. If the predominant flow of information is from node i to node j, then $0.5 < dTE < 1$. In the other case, when the information flow from j to i is outweighing the other flow direction, then $0 < dTE < 0.5$. For every node i, we calculate the average value of the directed delayed correlation with all other nodes of the network. For an interpretation of this value, if this overall average of dTE for node i is larger than 0.5, we can conclude that node i seems to be more sending information to the network (similar to the dPTE in (Hillebrand et al., 2016)). If the averaged dTE is smaller than 0.5 for a node, it would mean that this node is more receiving information from the network. In order to verify our modeling results, we will compare for different values of the effective spreading rate τ our results with empirical MEG matrices. As mentioned earlier, TE can be interpreted as a delayed correlation measure that is corrected for auto-correlation (see Appendix C). Since the measure of TE is quite complicated, we will also use analytic reasoning to analyze its more elementary 'building blocks', the (delayed) correlations and auto-correlations.

3 Results

We ran the SIS-epidemic simulations using a continuous-time simulator on the literature-based structural network. In Appendix A, we verified all previous results obtained by discrete-time simulations. In the following, we first analyze the influence of different time delays h from Eq. (2) on the transfer entropy and then relate structural topological properties to the overall global spreading pattern. In previous studies, the degree as the most straightforward centrality metric has been identified as a driver for effective connectivity (Stam et al., 2016, Mišić et al., 2015). Inspired also by analytic reasoning (see Appendix I), we analyze not only for the degree and but also for more advanced centrality metrics like the eigenvector and betweenness centrality their relationship with the spreading dynamics.

We plotted the TE of all node pairs in the network for different time delays h in Figure 3. From Figure 3, we can conclude that a direct structural connection leads to the highest TE value between two nodes. In addition, the further two nodes are away in terms of hopcount (i.e. the number of hops or links in the shortest path connecting these two nodes), the lower is the TE between them.

We found that the degree is a predictor for the overall directed transfer entropy of a node (Figure 4a) as well

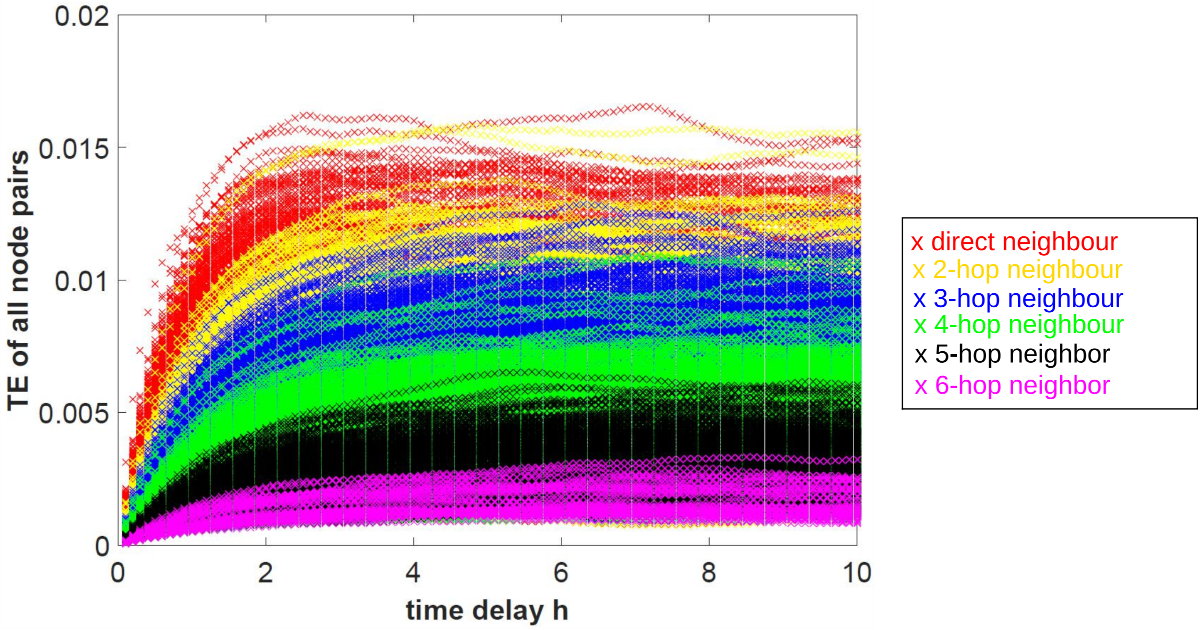
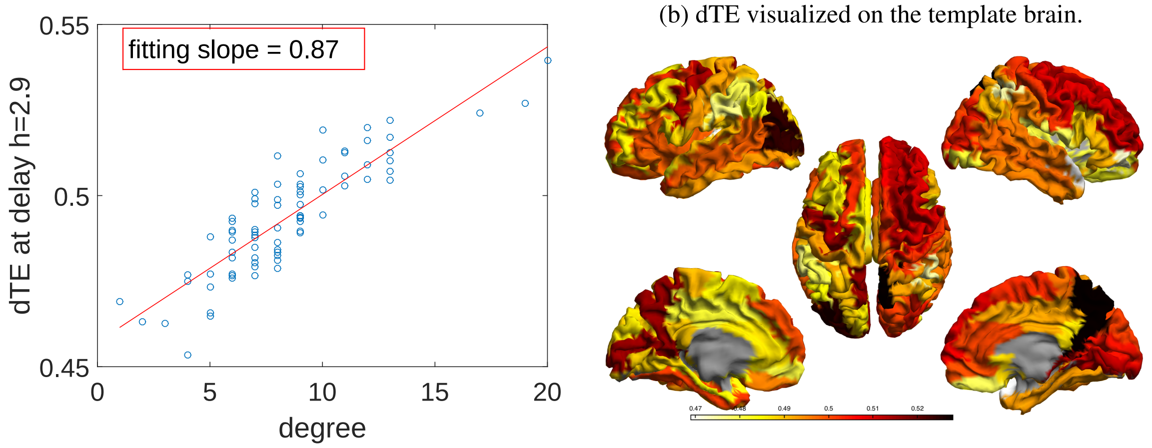


Figure 3: TE of all node pairs for different time delays. The TE values for different time lags *h* are colored in red for all the neighbors, in yellow for all nodes that are 2 hops away, in black for nodes with hopcount 3 and so on. For small time lags it seems that the direct neighbors of a node have the highest TE and that the further away another node is in terms of hopcount the lower is the TE between that node and the regarded node.

as for the overall activation of a node (Figure 5). These results confirm the intuitive assumption that the status of a high degree node has a bigger overall influence on the rest of the network than a node with low degree.

Figure 4b shows that the posterior regions seem to possess more outgoing flow of information (darker colors) whereas anterior regions have a more incoming flow (lighter colors). This finding is consistent with (Hillebrand et al., 2016) where they found in empirical data of healthy controls a predominant information flow in the posterior-anterior direction in most frequency bands. We compared the averaged directed phase transfer entropy values of the empirical matrices with our dTE matrix and obtained a positive correlation ($corr = 0.334$ for the time delay $h = 2.9$ time units, see SI Figure 13). However, the pattern seems to be less clear here than in the empirical global pattern since some rather large part of the frontal right hemisphere seems to be more sending than receiving (see darker colors in Figure 4b).

In addition to computing the measure of transfer entropy popular in the neuroscience community, we also performed similar analysis with the measure of delayed correlation (see Appendix D, E and I). We enlarged our analysis to this 'building block' of the transfer entropy (see Appendix C) since the delayed correlations are a measure that is closer to the theoretical process of SIS epidemics and easier to directly derive analytical results from (see Appendix I). Our analysis of delayed correlations in Appendix D and E using simulations and analytical reasoning shows that the posterior-anterior pattern cannot only be observed when using the measure of transfer entropy but also when using the simple measure of delayed correlations.



(a) dTE versus degree.

(b) dTE visualized on the template brain.

Figure 4: (a) dTE of a node (averaged over its influence on all other nodes) versus its degree. (b) Visualization of the dTE for each brain region on the parcellated template brain, where darker colors represent strong sending brain regions and lighter colors a more receiving property of a region. We show the brain here in clockwise order from the left, top, right, right midline and left midline.

4 Discussion

Using a simple model of activity spread, we were able to reproduce the empirically observed posterior-anterior pattern for effective connectivity. In addition, the structural degree of a node was shown to be a strong indicator for the sending/receiving property of a brain region. Moreover, the further two brain regions are away in terms of hopcount in the structural network, the lower is the TE between them.

It has been shown that hubs play a special role in both the healthy (Gong et al., 2009, Hagmann et al., 2008) and diseased brain (Crossley et al., 2014). Previous studies have also shown that structural brain networks have the strongest hubs in posterior regions (Buckner et al., 2008). These densely connected hubs in the back of the structural brain network have been identified as part of the so-called 'rich-club' of the brain (Senden et al., 2014, van den Heuvel and Sporns, 2011, van den Heuvel et al., 2012, Harriger et al., 2012). A recent study characterized the rich club as a control backbone of the structural brain network (Betzel et al., 2016). Concerning network dynamics, the rich club of the brain has been found to influence the rest of the network in a top-down manner (Harriger et al., 2012, Towlson et al., 2013, van den Heuvel and Sporns, 2013). Our study shows that hubs have a more sending than receiving property. This finding is in line with multiple studies that have shown that hubs seem to be driving the integration of information in the human brain (Sporns et al., 2007). Thus, the global posterior-anterior directionality pattern probably emerges due to the spatial distribution of the hubs accumulating more in the posterior regions and their stronger outgoing property in terms of information flow. More surprisingly, this directed flow direction seems to be emerging on an underlying undirected network using a simple model of activity spread.

The applied model in our study is a simple SIS epidemic spreading model, which ignored microscopic details of the real underlying spreading process in order to analyze global patterns. Even though our model ignored heterogeneity except for the underlying structural network restrictions, we were able to generate the empirically found global, directed spreading pattern (Hillebrand et al., 2016). Our approach is in line with other recent studies about global spreading dynamic principles with the help of simple dynamic models. (Mišić et al., 2015) found that the hubs and a backbone of core pathways facilitate the spreading process and shortest paths accelerate this phenomenon by applying a deterministic cascade model. As another example of a simple model, diffusion models identified the shortest path structure of the structural brain network as a driving force behind the network dynamics (Goñi et al., 2014) and were able to help categorize functional modules of the brain (Betzel et al., 2013,Delvenne et al., 2010). The novelty of our approach is that by using this simple model of activity spread, we analyzed instead of the functional the effective connectivity trying to identify the causal interactions between brain regions. These simple modeling approaches should be considered as complementary to other more traditional modeling studies from computational neuroscience. Moreover, it is known from statistical physics that near a critical phase transition the details of the model become irrelevant (Stam et al., 2016). This observation is a possible explanation why simple models as in (Deco et al., 2012,Haimovici et al., 2013) are able to capture complicated model findings (Honey et al., 2007,Honey et al., 2009). Most importantly, these simpler models allow us to study the basic principles of dynamics on brain networks with a minimum set of a-priori assumptions and parameters. Thus, the emergence of the directionality pattern from posterior to anterior regions could for more complex models be ascribed to any of the complex underlying model properties. In our case, because of the simplicity of the model and the underlying undirected network the emergence of this posterior-anterior pattern can be traced back to the spatially unequal distribution of hubs.

Our study shows that the structural distance between two brain regions seems to have an influence on their transfer entropy. From our simulation results, we can conclude that the further away two nodes are in terms of hopcount in the underlying structural brain network, the lower is their transfer entropy. This result is in line with one of our previous studies that identified the structural hopcount as a driving force for a functional connection (Meier et al., 2016). Moreover, (Goñi et al., 2014) identified that shortest paths of the structural network and detours along these paths are good predictors for the functional connectivity, which also leads to a lower connectivity for node pairs with larger hopcount between them. These results confirm the common assumption that longer paths in the structural brain network only have a small influence on the functional connectivity between two brain regions (Sarkar et al., 2015, Zamora-López et al., 2016, Meier et al., 2016). In our previous study (Stam et al., 2016) the direct structural connection was found to result in the highest effective connectivity and (Honey et al., 2009) stated that indirect connections with the hopcount 2 have a high influence on the functional connectivity strength between brain regions. Our results are in agreement with these earlier studies by identifying the hopcount between two brain regions as an indicator for their functional connectivity and show that these general principles also hold for the effective connectivity.

Methodological limitations

The used structural brain network with 78 nodes obtained from DTI measurements is quite small and newer bigger networks have been published already. By using this rather small but often-used structural network, we are able to compare our results directly to other studies in the field.

The interpretation of the TE from i to j as a causal influence should be taken carefully. Following the definition of causality from the Granger causality, the interpretation is correct but in other fields the term causality can have different meanings (Razak and Jensen, 2014, Pearl, 2010).

The previous study by (Hillebrand et al., 2016) showed that for the *theta* band the directionality pattern was observed in the opposite direction, thus from anterior to posterior regions. The authors hypothesized that those opposite directions of information flow indicate the presence of a loop between the two interacting subsystems of the Default Mode Network, the temporal and the fronto-parietal system representing a mechanism of integration of brain function (Hillebrand et al., 2016). With our current SIS model, we were only able to reproduce the more dominant posterior-anterior pattern from higher frequency bands. Future modelling studies should attempt to solve this question of mirrored information flow for different frequency bands.

5 Conclusion

In this study, we analyzed local and global network dynamics of the brain network by applying an SIS epidemic spreading model on the human connectome. We find that just above the critical threshold direct structural connections induce higher transfer entropy between two brain regions and that transfer entropy decreases with increasing distance between nodes (in terms of hops in the structural network). Hubs seem to play a special role in the network dynamics and we show that they send more information to the network than that they receive. Using the measure of transfer entropy, we can confirm the previously empirically observed dominant back-to-front pattern with our SIS model both with the help of analytic derivations and continuous-time simulations.

We also showed that the posterior hubs seem to be responsible for this observed directionality spreading pattern emerging from an underlying undirected structural network. Future studies should not only investigate these global spreading patterns by using other stochastic models but also analyze how the dynamics are disrupted and what other patterns occur in patients with brain disorders.

Acknowledgments

We thank Gaolang Gong for providing structural network data. Further, we thank Ruud van de Bovenkamp for Figures 22 and 23 in the Appendix and for making his SIS simulator available to us.

References

Abdelnour, F., Voss, H. U., and Raj, A. (2014). Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *NeuroImage*, 90:335–347.

Aertsen, A., Gerstein, G., Habib, M., and Palm, G. (1989). Dynamics of neuronal firing correlation: modulation of" effective connectivity". *Journal of neurophysiology*, 61(5):900–917.

Barrat, A., Barthelemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752.

Betzel, R. F., Griffa, A., Avena-Koenigsberger, A., Goñi, J., THIRAN, J.-P., Hagmann, P., and Sporns, O. (2013). Multi-scale community organization of the human structural connectome and its relationship with resting-state functional connectivity. *Network Science*, 1(03):353–373.

Betzel, R. F., Gu, S., Medaglia, J. D., Pasqualetti, F., and Bassett, D. S. (2016). Optimally controlling the human connectome: the role of network topology. *arXiv preprint arXiv:1603.05261*.

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, 1124(1):1–38.

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

Cator, E., Van de Bovenkamp, R., and Van Mieghem, P. (2013). Susceptible-infected-susceptible epidemics on networks with general infection and cure times. *Physical Review E*, 87(6):062816.

Cator, E. and Van Mieghem, P. (2014). Nodal infection in markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated. *Physical Review E*, 89(5):052802.

Crossley, N. A., Mechelli, A., Scott, J., Carletti, F., Fox, P. T., McGuire, P., and Bullmore, E. T. (2014). The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain*, 137(8):2382–2395.

de Haan, W., Mott, K., van Straaten, E. C., Scheltens, P., and Stam, C. J. (2012). Activity dependent degeneration explains hub vulnerability in alzheimer's disease. *PLoS Comput Biol*, 8(8):e1002582.

Deco, G., Senden, M., and Jirsa, V. (2012). How anatomy shapes dynamics: a semi-analytical study of the brain at rest by a simple spin model. *Frontiers in computational neuroscience*, 6.

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

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping*, 2(1-2):56–78.

Gong, G., He, Y., Concha, L., Lebel, C., Gross, D. W., Evans, A. C., and Beaulieu, C. (2009). Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cerebral cortex*, 19(3):524–536.

Goñi, J., van den Heuvel, M. P., Avena-Koenigsberger, A., de Mendizabal, N. V., Betzel, R. F., Griffa, A., Hagmann, P., Corominas-Murtra, B., Thiran, J.-P., and Sporns, O. (2014). Resting-brain functional connectivity predicted by analytic measures of network communication. *Proceedings of the National Academy of Sciences*, 111(2):833–838.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., and Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biol*, 6(7):e159.

Haimovici, A., Tagliazucchi, E., Balenzuela, P., and Chialvo, D. R. (2013). Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Physical review letters*, 110(17):178101.

Harriger, L., Van Den Heuvel, M. P., and Sporns, O. (2012). Rich club organization of macaque cerebral cortex and its role in network communication. *PloS one*, 7(9):e46497.

Hillebrand, A., Tewarie, P., van Dellen, E., Yu, M., Carbo, E. W., Douw, L., Gouw, A. A., van Straaten, E. C., and Stam, C. J. (2016). Direction of information flow in large-scale resting-state networks is frequency-dependent. *Proceedings of the National Academy of Sciences*, 113(14):3867–3872.

Honey, C., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.-P., Meuli, R., and Hagmann, P. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, 106(6):2035–2040.

Honey, C. J., Kötter, R., Breakspear, M., and 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.

Li, W. (1990). Mutual information functions versus correlation functions. *Journal of statistical physics*, 60(5-6):823–837.

Lobier, M., Siebenhühner, F., Palva, S., and Palva, J. M. (2014). Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. *Neuroimage*, 85:853–872.

Meier, J., Tewarie, P., Hillebrand, A., Douw, L., van Dijk, B. W., Stufflebeam, S. M., and Van Mieghem, P. (2016). A mapping between structural and functional brain networks. *Brain connectivity*, 6(4):298–311.

Mišić, B., Betzel, R. F., Nematzadeh, A., Goñi, J., Griffa, A., Hagmann, P., Flammini, A., Ahn, Y.-Y., and Sporns, O. (2015). Cooperative and competitive spreading dynamics on the human connectome. *Neuron*, 86(6):1518–1529.

Moon, J.-Y., Lee, U., Blain-Moraes, S., and Mashour, G. A. (2015). General relationship of global topology, local dynamics, and directionality in large-scale brain networks. *PLoS Comput Biol*, 11(4):e1004225.

Paluš, M. and Stefanovska, A. (2003). Direction of coupling from phases of interacting oscillators: an information-theoretic approach. *Physical Review E*, 67(5):055201.

Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of modern physics*, 87(3):925.

Pearl, J. (2010). The mathematics of causal relations. *Causality and Psychopathology: Finding the Determinants of Disorders and their Cures. Oxford University Press, Corvallis, OR*, pages 47–65.

Razak, F. A. and Jensen, H. J. (2014). Quantifying 'causality' in complex systems: understanding transfer entropy. *PloS one*, 9(6):e99462.

Rubinov, M., Sporns, O., Thivierge, J.-P., and Breakspear, M. (2011). Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons. *PLoS Comput Biol*, 7(6):e1002038.

Sarkar, S., Chawla, S., and Xu, D. (2015). On inferring structural connectivity from brain functional-mri data. *arXiv preprint arXiv:1502.06659*.

Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, 85(2):461.

Senden, M., Deco, G., de Reus, M. A., Goebel, R., and van den Heuvel, M. P. (2014). Rich club organization supports a diverse set of functional network configurations. *NeuroImage*, 96:174–182.

Sporns, O., Honey, C. J., and Kötter, R. (2007). Identification and classification of hubs in brain networks. *PloS one*, 2(10):e1049.

Stam, C. J. (2014). Modern network science of neurological disorders. *Nature Reviews Neuroscience*, 15(10):683–695.

Stam, C. J., Hillebrand, A., van Dellen, E., Meier, J., Tewarie, P., van Straaten, E., and Van Mieghem, P. (2016). The relation between structural and functional connectivity patterns in complex brain networks. *International Journal of Psychophysiology*, 103:149–60.

Stam, C. J. and Van Straaten, E. (2012). The organization of physiological brain networks. *Clinical Neurophysiology*, 123(6):1067–1087.

Stam, C. J. and van Straaten, E. C. (2012). Go with the flow: use of a directed phase lag index (dpli) to characterize patterns of phase relations in a large-scale model of brain dynamics. *Neuroimage*, 62(3):1415–1428.

Towlson, E. K., Vértes, P. E., Ahnert, S. E., Schafer, W. R., and Bullmore, E. T. (2013). The rich club of the c. elegans neuronal connectome. *The Journal of Neuroscience*, 33(15):6380–6387.

van de Bovenkamp, R. (2015). *Epidemic Processes on Complex Networks: Modelling, Simulation and Algorithms*. PhD thesis, TU Delft.

van den Heuvel, M. P., Kahn, R. S., Goñi, J., and Sporns, O. (2012). High-cost, high-capacity backbone for global brain communication. *Proceedings of the National Academy of Sciences*, 109(28):11372–11377.

van den Heuvel, M. P. and Sporns, O. (2011). Rich-club organization of the human connectome. *The Journal of neuroscience*, 31(44):15775–15786.

van den Heuvel, M. P. and Sporns, O. (2013). An anatomical substrate for integration among functional networks in human cortex. *The Journal of neuroscience*, 33(36):14489–14500.

Van Mieghem, P. (2011). *Graph Spectra for Complex Networks*. Cambridge University Press.

Van Mieghem, P. (2014a). Exact markovian sir and sis epidemics on networks and an upper bound for the epidemic threshold. *arXiv preprint arXiv:1402.1731*.

Van Mieghem, P. (2014b). *Performance Analysis of Complex Networks and Systems*. Cambridge University Press.

Van Mieghem, P., Omic, J., and Kooij, R. (2009). Virus spread in networks. *Networking, IEEE/ACM Transactions on*, 17(1):1–14.

Van Mieghem, P. and Van de Bovenkamp, R. (2013). Non-markovian infection spread dramatically alters the susceptible-infected-susceptible epidemic threshold in networks. *Physical review letters*, 110(10):108701.

Vértes, P. E., Alexander-Bloch, A. F., Gogtay, N., Giedd, J. N., Rapoport, J. L., and Bullmore, E. T. (2012). Simple models of human brain functional networks. *Proceedings of the National Academy of Sciences*, 109(15):5868–5873.

Wiener, N. (1956). The theory of prediction. *Modern mathematics for engineers*, 1:125–139.

Yu, S., Yang, H., Shriki, O., and Plenz, D. (2013). Universal organization of resting brain activity at the thermodynamic critical point. *Frontiers in systems neuroscience*, 7.

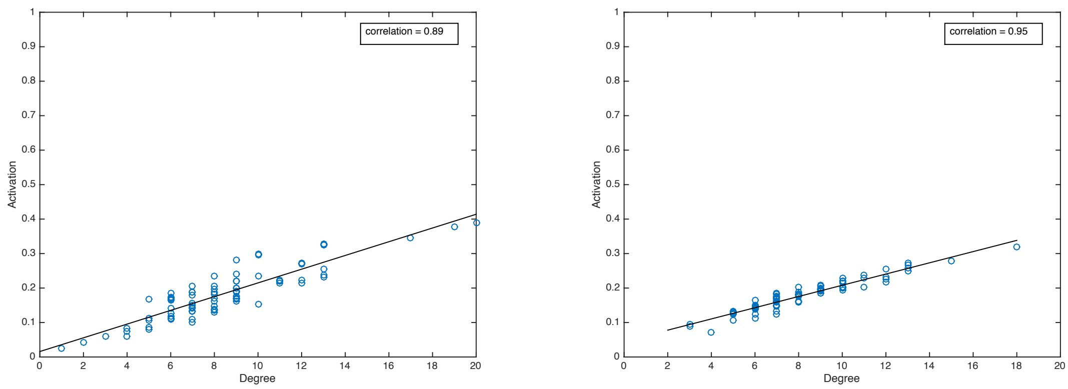
Zamora-López, G., Chen, Y., Deco, G., Kringelbach, M. L., and Zhou, C. (2016). Functional complexity emerging from anatomical constraints in the brain: the significance of network modularity and rich-clubs. *arXiv preprint arXiv:1602.07625*.

Appendix

A Verification of previous results in continuous time

In a previous study, (Stam et al., 2016) applied a discrete-time SIS epidemic model on the same literature-based structural network from (Gong et al., 2009). We verify the achieved results with a continuous-time SIS epidemic simulator (van de Bovenkamp, 2015).

First, we confirm the relationship between the degree of a node and its overall activation which is the percentage of the total time that this node was activated. The conclusion that a higher degree is associated with a higher overall activation is also valid in continuous time (see Figure 5(a)). This seems to be a general result since the two measures have an even higher correlation in an ER graph with similar degree distribution (see Figure 5(b), for details of the used ER graph see Appendix G).



(a) Structural brain network.

(b) Erdös-Renyi.

Figure 5: Scatterplot of a node's degree versus its overall activation for the underlying structural brain network (left panel) and an underlying ER graph (right panel).

To analyze the interactions of the nodes' time series, (Stam, 2014) used two different measures, the functional and the effective connectivity. The functional connectivity between two nodes i and j is defined as the correlation of their activation series

$$\rho(X_i(t), X_j(t)) = \frac{E[X_i(t)X_j(t)] - E[X_i(t)]E[X_j(t)]}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_j(t)]}} \tag{4}$$

over the whole simulation time (Stam et al., 2016). The numerator is also referred to as the covariance between node i and j. Recently, Cator and Van Mieghem (Cator and Van Mieghem, 2014) have proved that

$$E\left[X_i(t)X_j(t)\right] \geq E\left[X_i(t)\right]E\left[X_j(t)\right]$$

is correct for any graph and any time in Markovian SIS epidemics, however, it is false in general for non-Markovian SIS processes on networks as in (Van Mieghem and Van de Bovenkamp, 2013, Cator et al., 2013). Hence, for Markovian SIS epidemics, we know that $\rho\left(X_i(t), X_j(t)\right) \geq 0$ for any node pair (i, j) at time t. The functional connectivity or correlation matrix based on continuous-time simulations is shown in Figure 6(b). We binarized the weighted adjacency matrix here to display the same number of links as the structural network only considering the strongest connections. As for discrete-time simulations, the functional connectivity matrix resembles the underlying structural network (compare Figure 6(a) and (b)).

Stam et al. (Stam et al., 2016) calculated the effective connectivity between node i and node j as

$$C_{eff} = Pr[X_j(t+h) = 1 | X_i(t) = 1] + Pr[X_i(t+h) = 1 | X_j(t) = 1]$$

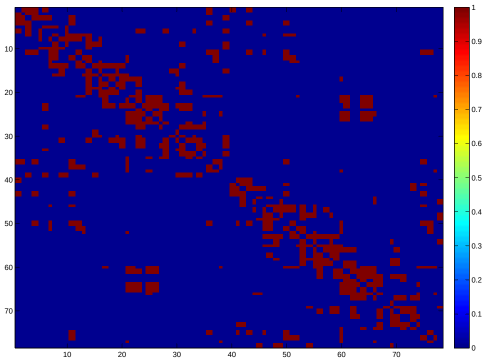
where for discrete-time simulations Stam et al. considered the time lag or time delay $h = 1s$. We calculated the effective connectivity similarly and divided it additionally by 2 to stay between 0 and 1. We plotted the results for the effective connectivity matrix (with the same number of links as the matrix A in a binary form) in Figure 6(c) for $h = 0.1s$, which equals one time step in our simulations, where we can also recognize the pattern of four crosses in the matrix which is common to empirical matrices measured from e.g. MEG or fMRI. To explore the structure of this effective connectivity in more detail, we can analyze different time delays h with our continuous-time SIS simulations: with a larger time lag, the global pattern of four crosses becomes even more apparent (see Appendix A for matrices under different time delays).

Inspired by Barrat *et al.* (Barrat et al., 2004), Stam *et al.* (Stam et al., 2016) have investigated the relation between C_{eff} and the degree product $d_i d_j$, where d_k is the degree of node k. They found a similar kind of relation as Barrat *et al.* (Barrat et al., 2004), namely,

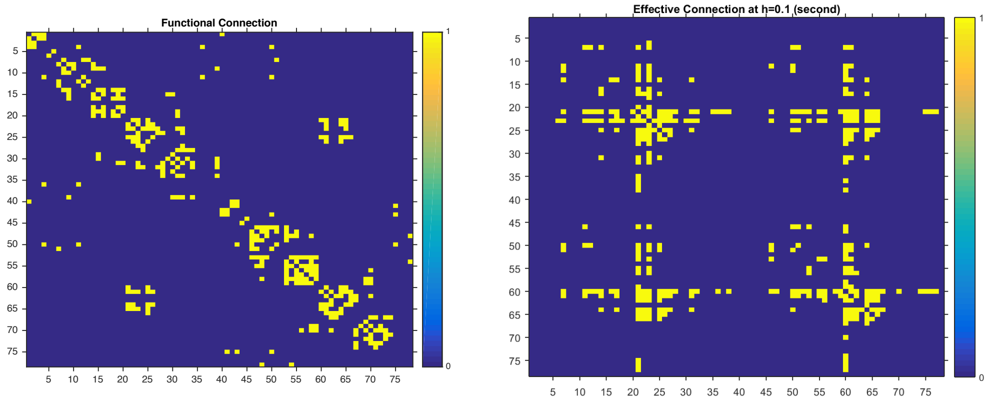
$$C_{eff} = \gamma\left(d_i d_j\right)^{\alpha} \tag{5}$$

where γ is a proportionality constant and α is around $\frac{1}{2}$, but not such a clear relationship for the functional connectivity $\rho\left(X_i(t), X_j(t)\right)$. The law (5) suggests that the degree product of two nodes in G is predictive for the intensity of traffic flowing between them, irrespective of the underlying path structure. We computed the same quantities with our continuous-time simulation and found also that the degree product seems to be a strong predictor for the effective connectivity between two nodes (see Figure 7).

Overall, we obtained similar results as the discrete-time simulations and with the continuous-time simulator we are now able to further analyze the influence of different time delays on the interactions between node



(a) Structural brain network.



(b) Functional Connectivity.

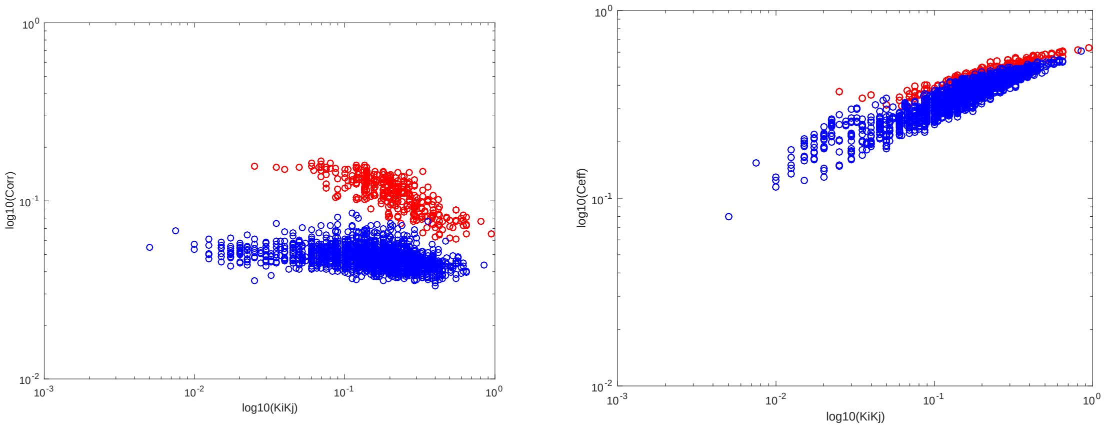
(c) Effective Connectivity for $h = 0.1s$.

Figure 6: (a) Visualization of the structural brain network that we used as an underlying network for our simulations (Gong et al., 2009). (b) Functional Connectivity based on the SIS epidemics time series with $\beta = 0.1$ and $\delta = 0.5$. (c) Effective Connectivity based on the SIS epidemics time series with $\beta = 0.1$ and $\delta = 0.5$ and a delay of $h = 0.1s$, which is one time step in our simulations.

activations.

B Effective connectivity for different time lags

In Figure 8 we show the binarized adjacency matrix for different time lags h. In all of the panels of Figure 8 we can recognize the overall pattern of the four crosses in the colored matrix. This pattern seems to be present for all time lags h and becomes even more clear for increasing h.



(a) Functional Connectivity.

(b) Effective Connectivity.

Figure 7: Functional and effective connectivity of all node pairs against their structural degree product with $\beta = 0.1$ and $\delta = 0.5$ in a log-log plot. Node pairs with a direct structural connection are marked in red and with an indirect connection in blue.

C Correlation versus transfer entropy

The transfer entropy is equal to the conditional mutual information (MI)

$$\begin{aligned} TE_{i\to j}(h) &= MI(X_j(t+h); X_i(t)|X_j(t)) \\ &= \sum_{k,l,m=\{0,1\}} \Pr[X_j(t+h)=k, X_j(t)=l, X_i(t)=m]\cdot \log\left(\frac{\Pr[X_j(t+h)=k|X_j(t)=l, X_i(t)=m]}{\Pr[X_j(t+h)=k|X_j(t)=l]}\right) \\ &= \sum_{k,l,m=\{0,1\}} \Pr[X_j(t+h)=k, X_j(t)=l, X_i(t)=m]\cdot \log\left(\frac{\Pr[X_j(t+h)=k, X_j(t)=l, X_i(t)=m]\Pr[X_j(t)]}{\Pr[X_j(t+h)=k, X_j(t)=l]\Pr[X_i(t), X_j(t+h)]}\right), \end{aligned}$$

where we applied the law of Bayes for the last equality.

The mutual information and the measure of correlation want to measure the same underlying property of two random variables, their 'distance to independence'. The covariance is defined as

$$cov(i, j, h) = E[X_i(t)X_j(t+h)] - E[X_i(t)]E[X_j(t+h)]$$

and measures the distance in terms of expected values of the random variables itself whereas the mutual information can be written as

$$MI(X_j(t+h); X_i(t)) = E[\log(Pr[X_i(t), X_j(t+h)])] - E[\log(Pr[X_i(t)]Pr[X_j(t+h)])]$$

and measures the distance in terms of the expected value of the logarithm of their probabilities.

For the transfer entropy, we can apply the chain rule of the mutual information and obtain

$$\begin{aligned}
TE_{i\to j}(h) &= MI(X_j(t+h); X_i(t)|X_j(t)) \\
&= MI(X_j(t+h); X_i(t), X_j(t)) - MI(X_j(t+h), X_j(t)) \\
&= \sum_{k,l,m=\{0,1\}} \Pr[X_j(t+h)=k, X_j(t)=l, X_i(t)=m] \cdot \log\left(\frac{\Pr[X_j(t+h)=k, X_j(t)=l, X_i(t)=m]}{\Pr[X_j(t+h)]\Pr[X_i(t)=k, X_j(t)=l]}\right) \\
&\quad - \sum_{k,l=\{0,1\}} \Pr[X_j(t+h)=k, X_j(t)=l] \cdot \log\left(\frac{\Pr[X_j(t+h)=k, X_j(t)=l]}{\Pr[X_j(t+h)]\Pr[X_j(t)=l]}\right) \qquad (6)
\end{aligned}$$

For the second term we followed the derivations in (Li, 1990) and used that our activation series are binary resulting in an approximative formula

$$MI(X_j(t+h), X_j(t)) \approx \frac{1}{2}\left(\frac{auto_j}{Pr[X_j(t)=1](1-Pr[X_j(t)=1])}\right)^2,$$

where $auto_j$ denotes the auto-correlation of j. To reach this result, we did assume that $Pr[X_j(t)=1] \approx Pr[X_j(t+h)=1]$ which could be confirmed by our simulations for small values of the time lag h. Thus, the second term of (6) can be interpreted as some correction for the auto-correlation that is included in the transfer entropy.

If we apply the Kirkwood superposition approximation to the first term of (6) which involves all three entities, we can approximate the joint probability of the three terms

$$\begin{aligned}
MI(X_j(t+h); X_i(t), X_j(t)) = &\sum_{k,l,m=\{0,1\}} \Pr[X_j(t+h)=k, X_j(t)=l, X_i(t)=m] \cdot \\
&\log\left(\frac{\Pr[X_j(t+h)=k, X_j(t)=l]\Pr[X_j(t+h)=k, X_i(t)=l]}{\Pr[X_j(t+h)]^2\Pr[X_i(t)=k]\Pr[X_j(t)=k]}\right)
\end{aligned}$$

where e.g. in the case of $k=l=m$ we obtain the element of the sum in the logarithm as

$$\log\left(\frac{\left(auto_j + \Pr[X_j(t)=k]^2\right)}{\Pr[X_j(t)=k]^2} \cdot \frac{\left(corr_{del}(i,j,h) + \Pr[X_j(t)=k]\Pr[X_i(t)=k]\right)}{\Pr[X_j(t)=k]\Pr[X_i(t)=k]}\right)$$

where $corr_{del}(i,j,h)$ is the delayed correlation function between node i and node j which will be further studied in the next section. For the other elements of the sum we can derive similar results in the logarithm reducing the expression to a combination of the auto-correlation of j and the delayed correlation between i and j. For the three-way joint probability in front of the logarithm, we can again use the Kirkwood superposition approximation and obtain e.g. for the element $k = l = m$

$$\Pr[X_j(t+h)=1, X_j(t)=1, X_i(t)=1] \approx \frac{\Pr[X_j(t+h)=1, X_j(t)=1]\Pr[X_j(t+h)=1, X_i(t)=1]\Pr[X_j(t)=1, X_i(t)=1]}{\Pr[X_j(t)=1]\Pr[X_i(t)=1]\Pr[X_j(t)=1]}$$

$$= \frac{(auto_j + \Pr[X_j(t)=1]^2)(corr_{del}(i,j,h) + \Pr[X_j(t)=1]\Pr[X_i(t)=1])(corr(i,j) + \Pr[X_j(t)=1]\Pr[X_i(t)=1])}{\Pr[X_j(t)=1]\Pr[X_i(t)=1]\Pr[X_j(t)=1]}$$

where $corr(i, j)$ is the correlation (or covariance) of the two nodes' binary time series. For the other elements of the sum, the derivation can be conducted similarly. To sum up, we have demonstrated that the transfer entropy from node i to node j can be expressed as a combination of the (delayed) correlation between i and j corrected for the auto-correlation of j. In the following section, we will further examine analytically the (delayed) correlation and auto-correlation as 'building blocks' of the transfer entropy.

D Delayed correlations

In order to better analyze the directionality of the delayed influence, we define the delayed correlation with delay $h \geq 0$ as

$$corr_{del}(i,j,h) = \rho(X_i(t), X_j(t+h)) = \frac{E[X_i(t)X_j(t+h)] - E[X_i(t)]E[X_j(t+h)]}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_j(t+h)]}}$$

(see also Figure9). The correlation or functional connectivity which we defined above is equal to the delayed correlation with delay $h = 0$. The conditional probability $Pr[X_j(t+h) = 1|X_i(t) = 1]$ from the effective connectivity is included in the delayed correlation

$$corr_{del}(i,j,h) = \rho(X_i(t), X_j(t+h)) = \frac{E[X_i(t)] \cdot \left(Pr[X_j(t+h)=1|X_i(t)=1] - E[X_j(t+h)]\right)}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_j(t+h)]}}.$$

where we used Bayes formula and the property of Bernoulli random variables $E[X_i(t)] = Pr[X_i(t) = 1]$. The conditional probability above is the only part which includes the node pair interaction, the other terms are only depending on a single node. Therefore, the conditional probability and the delayed correlation should follow similar behavior with respect to h. Because we are also interested in the direction of spreading, we will not average over both directions (like the effective connectivity) but we will further use the delayed correlation as our measure of influence from one node to the others.

Concerning the relationship between the delayed correlation and the correlation, we found the following relation

$$\frac{\rho\left(X_i(t), X_j(t+h)\right)}{\rho\left(X_i(t), X_j(t)\right)} \approx 1 + \left(\tau d_j - 1\right)\delta h + o(h)$$

(see Appendix I.3.1 for detailed derivations). Since $\tau > \tau_c \geq \frac{1}{\lambda_1}$, we have that $\tau d_j - 1 > \frac{d_j}{\lambda_1} - 1$. Since the spectral radius is bounded (Van Mieghem, 2011) by $\max\left(d_{av}, \sqrt{d_{\max}}\right) \leq \lambda_1 \leq d_{\max}$, where the average degree $d_{av} = \frac{2L}{N}$, the factor $\tau^* d_j - 1$ is positive for a node j with more than average degree, but possibly negative for a node j with low degree. This result means that hubs have a higher delayed influence than direct influence (correlation) than nodes with low degree which have possibly a higher direct influence than delayed influence. Another property of the delayed correlation based on analytical derivations is that it increases stronger for nodes with high degree with regard to increasing h (see Appendix I.1 for analytical derivations).

We plotted the delayed correlation of all node pairs in the network for different time delays h in Figure 10. Whereas (Cator and Van Mieghem, 2014) proved that the correlation is non-negative, the non-negativity of the delayed correlation for $h > 0$ can so far only be shown from the simulations (see Figure 10): if we either simulated long enough ($> 5 \cdot 10^5$ time units) or averaged over a large number of runs ($>$ 200 runs) the delayed correlations were non-negative. From Figure 10, we can conclude that a direct structural connection leads for small h to the highest delayed correlation between two nodes. In addition, the further two nodes are away in terms of hopcount (i.e. the number of hops or links in the shortest path connecting these 2 nodes), the lower is the delayed correlation between them (see Figure 10).

E Global spreading pattern with delayed correlations

Motivated by a recent paper of Hillebrand et al. (Hillebrand et al., 2016), we analyze the sending or receiving property of a node. This property means the relation for a node i between $\rho(X_i(t), X_j(t+h))$ and $\rho(X_j(t), X_i(t+h))$ averaged over all other nodes j, and investigates which flow direction outweighs. Similar to (Hillebrand et al., 2016) we define the directed delayed correlation for node i and j as

$$dcorr_{del}(i, j, h) = \frac{corr_{del}(i, j, h)}{corr_{del}(i, j, h) + corr_{del}(j, i, h)}. \tag{7}$$

If we assume that all delayed correlation values are positive (which can be supported by our simulation results), this definition of $dcorr_{del}$ is well-defined and its value ranges between 0 and 1. If the predominant flow of information is from node i to node j, then $0.5 < dcorr_{del} < 1$. In the other case, when the information flow from j to i is outweighing the other flow direction, then $0 < dcorr_{del} < 0.5$. For every node i, we calculate the average value of the directed delayed correlation with all other nodes of the network. For an interpretation of this value, if this overall average of $dcorr_{del}$ for node i is larger than 0.5, we can conclude that node i seems to be more sending information to the network (similar to the $dPTE$ in (Hillebrand et al., 2016)). If the averaged $dcorr_{del}$ is smaller than 0.5 for a node, it would mean that this node is more receiving information from the network.

Furthermore, we find that the degree is not only a predictor for the overall activation (see Figure 5) but also for the overall delayed correlation of a node (see Figure 11). This result confirms the intuitive assumption that a the status of a high degree node has a bigger overall influence on the rest of the network than a node with low degree. Taking into account the derivations from Appendix I.1.1, we find also analytically that the information flow is more likely to go from a hub node to a lower degree node.

For our data, we plotted for all brain regions this sending/receiving property in Figure 12, where darker colors represent strong sending brain regions and lighter colors a more receiving property of a region. Figure 12 shows that the posterior regions seem to possess more outgoing flow of information whereas anterior regions have a more incoming flow (lighter colors). This finding is consistent with (Hillebrand et al., 2016) where they found in empirical data of healthy controls a predominant information flow in the posterior-anterior direction in most frequency bands.

To compare the results of (Hillebrand et al., 2016) more quantitatively with ours, we plotted their empirical directed Phase Transfer Entropy results in a scatterplot with our averaged dTE values (see Figure 13) and obtained a positive correlation coefficient of 0.334 for the time delay of $h = 2.9$ time units.

F Auto-correlations

We define the auto-correlation as

$$auto_{del}(i,h) = \rho(X_i(t), X_i(t+h)) = \frac{E[X_i(t)X_i(t+h)] - E[X_i(t)]E[X_i(t+h)]}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_i(t+h)]}},$$

which becomes trivial for $h = 0$

$$\begin{aligned} auto_{del}(i,0) = \rho(X_i(t), X_i(t)) &= \frac{E[X_i(t)] - E[X_i(t)]^2}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_i(t)]}} \\ &= \frac{Var[X_i(t)]}{Var[X_i(t)]} = 1. \end{aligned}$$

We show in Appendix I.1 analytically that the auto-correlation function is decreasing for small time delays h. Simulations were able to verify this decreasing behavior for a larger range of h values (see Figure 14). Furthermore, we found from the simulations that for very small time lags h, the auto-correlation function even seems to be exponentially decaying (see Figure 15).

In addition, we find from the simulations that for small time lags h a node with a higher degree has a lower auto-correlation (see Figure 16). This relation could be due to the fact that a node with more neighbors has more external influence on his activation status, changes therefore often its state and thus obtains a lower auto-correlation. Analytically, we were able to show that a higher eigenvector centrality leads to a lower auto-correlation (see Appendix I.3.2).

G Random networks as benchmark

To compare our results to a benchmark, we performed the same analysis on an Erdös Renyi random graph with a similar degree distribution (see Figure 17). The patterns in the obtained matrices for the functional as well as for the effective connectivity seem to be more randomly distributed, not showing any specific recognizable forms (see Figure 18). In addition, the results of the delayed correlation confirm our generally formulated principles (see Figure 19) We also plotted from the structural brain network and from the ER graph only the delayed correlation between the nodes with the highest degree and the rest of the network (see Figures 20 and 21). Furthermore, we were also able to find a distinction in hopcounts for the delayed correlation with a bigger ER graph with 500 nodes (see Figure 22).

H Network Model

So far, the structural network data that we have access to only consists of a network of 78 nodes, where each node represents a region of interest (ROI) in the brain. This network is too small to obtain stable simulation results from the epidemic spreading process. Therefore, we decided to build a bigger structural network using a generative network model. This generative network model should resemble the anatomical connections in the human brain as precisely as possible, but it should be simple at the same time. In consultation with neurophysiologists, we developed a generative network model as an approximation for the structural brain network. In the following, we will present the three steps of the generative model.

- *Step 1:* We randomly place N nodes on a semicircle, which represents the left hemisphere of the human brain. Then, we mirror the node locations to the right semicircle, which represents the right hemisphere of the brain.

- *Step 2:* Within each hemisphere, we connect each node pair with a connection probability $P_{ij} = \exp(-\eta d_{ij})$ that exponentially decays with distance (Vértes et al., 2012), where d_{ij} is the Euclidean distance between node i and j, and η is a parameter that needs to be chosen. This distance penalty for connections is based on the probable aim of the brain to minimize metabolic cost (Vértes et al., 2012). We call the created links the local connections, since they only occur within one hemisphere.
- *Step 3:* The strong long-range connections between the left and the right hemisphere and between the front and the back of the brain are crucial for the dynamics of the human brain. Thus, we first divide the randomly placed nodes into regions by placing 39 equally sized circles onto both semicircles. These circles will later correspond to the 78 brain regions of interest that are measured in our MEG experiments. Then, we pick randomly L circles and connect all nodes in this circle to all nodes in the corresponding circle on the opposite side of the brain. We carry out this procedure once for the connections between the two hemispheres. Then, we divide the circle into a front and a back part by bisecting horizontally and apply the same procedure for generating long-range connections as for the cross-hemisphere connections.

I Analytical Derivations on delayed SIS correlations

I.1 The covariance $\widetilde{\rho}\left(X_i(t), X_j(t+h)\right)$ for a small time lag h

We compute the expectation $E\left[X_i(t) X_j(t+h)\right]$ for a very small time $h > 0$. Although the derivative of a Bernoulli random variable does not exist, we follow the framework in (Van Mieghem, 2014a) and we agree to *formally* define the derivative by the random variable equation

$$\frac{dX_j(t)}{dt} = -\delta X_j(t) + \left(1 - X_j(t)\right)\beta \sum_{k=1}^{N} a_{kj} X_k(t) \tag{8}$$

For small h, the first order expansion of the Taylor series yields

$$X_j(t+h) = X_j(t) + h\frac{dX_j(t)}{dt} + o(h) \tag{9}$$

Then,

$$\begin{aligned} E\left[X_i(t)X_j(t+h)\right] &= E\left[X_i(t)X_j(t) + h\frac{dX_j(t)}{dt}X_i(t) + o(h)\right] \\ &= E\left[X_i(t)X_j(t)\right] + hE\left[\frac{dX_j(t)}{dt}X_i(t)\right] + o(h) \end{aligned}$$

and, after invoking (8), we obtain

$$\begin{aligned} E\left[X_i(t)X_j(t+h)\right] &= E\left[X_i(t)X_j(t)\right] + hE\left[-\delta X_j(t)X_i(t) + \left(1 - X_j(t)\right)X_i(t)\beta\sum_{k=1}^{N} a_{kj}X_k(t)\right] + o(h) \\ &= (1-\delta h)E\left[X_i(t)X_j(t)\right] + \beta hE\left[X_i(t)\sum_{k=1}^{N} a_{kj}X_k(t) - X_j(t)X_i(t)\sum_{k=1}^{N} a_{kj}X_k(t)\right] + o(h) \end{aligned}$$

After some rearrangements, we arrive at

$$\begin{aligned} E\left[X_i(t)X_j(t+h)\right] = (1-\delta h)E\left[X_i(t)X_j(t)\right] + \beta h\sum_{k=1}^{N} a_{kj}E\left[X_i(t)X_k(t)\right] \\ -\beta h\sum_{k=1}^{N} a_{kj}E\left[X_i(t)X_j(t)X_k(t)\right] + o(h) \end{aligned} \tag{10}$$

Similarly,

$$\begin{aligned} E\left[X_j(t+h)\right] &= E\left[X_j(t)\right] + hE\left[\frac{dX_j(t)}{dt}\right] + o(h) \\ &= (1-\delta h)E\left[X_j(t)\right] + \beta h\sum_{k=1}^{N} a_{kj}E\left[X_k(t)\right] - \beta h\sum_{k=1}^{N} a_{kj}E\left[X_j(t)X_k(t)\right] + o(h) \end{aligned} \tag{11}$$

which shows that the probability of infection for node j at time $(t+h)$ depends for small h on three elements: (1) on the state of node j at time t, (2) on the state of the neighbors of node j at time t and (3) on the joint probability with its neighbors.

Thus, the covariance $r = E\left[\left(X_i(t) - E\left[X_i(t)\right]\right)\left(X_j(t+h) - E\left[X_j(t+h)\right]\right)\right]$ for small h is

$$\begin{aligned} r &= E\left[X_i(t)X_j(t+h)\right] - E\left[X_i(t)\right]E\left[X_j(t+h)\right] \\ &= (1-\delta h)E\left[X_i(t)X_j(t)\right] + \beta h\sum_{k=1}^{N} a_{kj}E\left[X_i(t)X_k(t)\right] - \beta h\sum_{k=1}^{N} a_{kj}E\left[X_i(t)X_j(t)X_k(t)\right] + o(h) \\ &\quad - (1-\delta h)E\left[X_j(t)\right]E\left[X_i(t)\right] - \beta h\sum_{k=1}^{N} a_{kj}E\left[X_i(t)\right]E\left[X_k(t)\right] + \beta h\sum_{k=1}^{N} a_{kj}E\left[X_i(t)\right]E\left[X_j(t)X_k(t)\right] \end{aligned}$$

With the definition (4) of the covariance, we find for $i \neq j$ that the j-delayed covariance satisfies

$$\widetilde{\rho}\left(X_i(t), X_j(t+h)\right) = (1-\delta h)\widetilde{\rho}\left(X_i(t), X_j(t)\right) + \beta h \sum_{k=1}^{N} a_{kj}\widetilde{\rho}(X_i(t), X_k(t))$$
$$-\beta h \sum_{k=1}^{N} a_{kj}\left\{E\left[X_i(t)X_j(t)X_k(t)\right] - E\left[X_i(t)\right]E\left[X_j(t)X_k(t)\right]\right\} + o(h) \qquad (12)$$

The derivative of the j-delayed covariance $\widetilde{\rho}\left(X_i(t), X_j(q)\right)$ with respect to the time q and evaluated at time $q = t$ follows as

$$\left.\frac{d\widetilde{\rho}\left(X_i(t), X_j(q)\right)}{dq}\right|_{q=t} = \left(\beta a_{ij} - \delta\right)\widetilde{\rho}\left(X_i(t), X_j(t)\right) + \beta \sum_{k=1;k\neq i}^{N} a_{kj}\widetilde{\rho}(X_i(t), X_k(t))$$
$$-\beta \sum_{k=1}^{N} a_{kj}\left\{E\left[X_i(t)X_j(t)X_k(t)\right] - E\left[X_i(t)\right]E\left[X_j(t)X_k(t)\right]\right\}$$

where

$$\lim_{h\to 0}\frac{\widetilde{\rho}\left(X_i(t), X_j(t+h)\right) - \widetilde{\rho}\left(X_i(t), X_j(t)\right)}{h} = \left.\frac{d\widetilde{\rho}\left(X_i(t), X_j(t+h)\right)}{dh}\right|_{h=0} = \left.\frac{d\widetilde{\rho}\left(X_i(t), X_j(q)\right)}{dq}\right|_{q=t}$$

We observe that the j-delayed covariance $\widetilde{\rho}\left(X_i(t), X_j(t)\right)$ increases stronger when node j has more neighbors.

On the other hand for $i = j$, the autocovariance function for $h > 0$

$$r_i(t,h) = E\left[X_i(t)X_i(t+h)\right] - E\left[X_i(t)\right]E\left[X_i(t+h)\right]$$

and, for $h = 0$,

$$r_i(t,0) = \text{Var}\left[X_i(t)\right] = E\left[X_i(t)\right]\left\{1 - E\left[X_i(t)\right]\right\}$$

becomes for a small time lag h

$$r_i(t,h) = (1-\delta h)\text{Var}\left[X_i(t)\right] - \beta h E\left[X_i(t)\right]\sum_{k=1}^{N} a_{ki}\left\{E\left[X_k(t)\right] - E\left[X_i(t)X_k(t)\right]\right\} + o(h) \qquad (13)$$

In the heavy infection regime ($\tau \to \infty$), we know that $X_i \to 1$ and $\text{Var}[X_i(t)] \to 0$ and the autocovariance function (13) tends to zero, $r_i(t,h) \to 0$. Since $E\left[X_i(t)X_k(t)\right] \leq E\left[X_k(t)\right]$ for any pair of Bernoulli random variables $X_i \in \{0,1\}$ and $X_k \in \{0,1\}$, we observe that the second term is always non-positive so that the autocovariance function $r_i(t,h) \leq \text{Var}[X_i(t)] = r_i(t,0)$. In other words, the autocovariance function $r_i(t,h)$ and thus also its normalized version, the auto-correlation function, decreases with small time lag h for any time t.

I.1.1 Deductions

In general, the j-delayed covariance $\widetilde{\rho}\left(X_i(t), X_j(t+h)\right)$ is different from the i-delayed covariance $\widetilde{\rho}\left(X_j(t), X_i(t+h)\right)$. Indeed, (12) demonstrates that

$$\widetilde{\rho}\left(X_i(t), X_j(t+h)\right) - \widetilde{\rho}\left(X_j(t), X_i(t+h)\right) = \beta h \sum_{k=1}^{N} \left(a_{kj}\widetilde{\rho}(X_i(t), X_k(t)) - a_{ki}\widetilde{\rho}(X_j(t), X_k(t))\right)$$

$$-\beta h \sum_{k=1}^{N} (a_{kj} - a_{ki}) E[X_i(t)X_j(t)X_k(t)] - \beta h \sum_{k=1}^{N} \left(a_{kj} E[X_i(t)] E[X_j(t)X_k(t)] - a_{ki} E[X_j(t)] E[X_i(t)X_k(t)]\right)$$

Starting from the original definition of $\widetilde{\rho}\left(X_i(t), X_j(t+h)\right)$, we can deduce the parts as

$$E\left[X_i(t) X_j(t+h)\right] = E\left[X_i(t) X_j(t)\right] + hE\left[-\delta X_j(t) X_i(t) + \left(1 - X_j(t)\right) X_i(t) \beta \sum_{k=1}^{N} a_{kj} X_k(t)\right] + o(h)$$

$$E\left[X_j(t) X_i(t+h)\right] = E\left[X_j(t) X_i(t)\right] + hE\left[-\delta X_i(t) X_j(t) + (1 - X_i(t)) X_j(t) \beta \sum_{k=1}^{N} a_{ki} X_k(t)\right] + o(h)$$

Subtraction yields

$$\begin{aligned} T &= E\left[X_i(t) X_j(t+h)\right] - E\left[X_j(t) X_i(t+h)\right] \\ &= \beta h E\left[X_i(t)\left(1 - X_j(t)\right) \sum_{k=1}^{N} a_{kj} X_k(t) - X_j(t)(1 - X_i(t)) \sum_{k=1}^{N} a_{ki} X_k(t)\right] \\ &= \beta h E\left[X_i(t) \sum_{k=1}^{N} a_{kj} X_k(t) - X_j(t) \sum_{k=1}^{N} a_{ki} X_k(t) + X_i(t) X_j(t) \sum_{k=1}^{N} \left(a_{ki} - a_{kj}\right) X_k(t)\right] \end{aligned}$$

where $\sum_{k=1}^{N} a_{kj} X_k(t)$ are the infected neighbors of node j. The first equation tells that T is the balance of two cases: either an infection at node i and all infected neighbors of node j try to infect node j or an infection at node j and all infected neighbors of node i try to infect node i.

The other parts of the difference of covariances can be written as

$$E[X_j(t)]E[X_i(t+h)] - E[X_i(t)]E[X_j(t+h)] = \beta h E[X_j(t)] \sum_{k=1}^{N} a_{ki} E[X_k(t)] - \beta h E[X_i(t)] \sum_{k=1}^{N} a_{kj} E[X_k(t)]$$
$$-\beta h \sum_{k=1}^{N} a_{ki} E[X_i(t)X_k(t)]E[X_j(t)] + \beta h \sum_{k=1}^{N} a_{kj} E[X_j(t)X_k(t)]E[X_i(t)]$$

Together we get for the probability flux

$$\widetilde{\rho}\left(X_i(t), X_j(t+h)\right) - \widetilde{\rho}\left(X_j(t), X_i(t+h)\right) = \beta h \sum_{k=1}^{N} a_{kj} \left(E[X_i(t)]E[X_j(t)X_k(t)] - E[X_i(t)X_j(t)X_k(t)]\right)$$
$$-\beta h \sum_{k=1}^{N} a_{ki} \left(E[X_j(t)]E[X_i(t)X_k(t)] - E[X_i(t)X_j(t)X_k(t)]\right)$$

The term $E[X_i(t)]E[X_j(t)X_k(t)] - E[X_i(t)X_j(t)X_k(t)]$ is probably negative for any node triple i, k and j, where k and j are neighbors, thus $E[X_i(t)]E[X_j(t)X_k(t)] \leq E[X_i(t)X_j(t)X_k(t)]$ where it is definitely true for $k = j$ (and i and j neighbors). Then, if $d_i >> d_j$ the flux is likely to be positive and the flow is from node i to node j. This information flow has also been researched by (Hillebrand et al., 2016), where instead of the delayed correlation, (Hillebrand et al., 2016) used the *Phase Transfer Entropy* as a measure of causality. This difference in coavariances is positive if and only if the fraction from (Hillebrand et al., 2016) is larger than 0.5. They found an overall direction flow from the posterior to anterior brain regions (Hillebrand et al., 2016). Since most of the hubs of the structural brain network are known to be located in the back, our analytical derivations can confirm that the information flow is then mostly from the hubs towards the low degree nodes. Thus, overall a posterior-to-anterior flow should emerge, also with our measure of delayed correlations.

Further, we consider the autocovariance function, normalized at $h = 0$,

$$\frac{r_i(t,h)}{r_i(t,0)} = 1 - \delta h - \beta h \sum_{k=1}^{N} a_{ki} \left\{ \frac{E[X_k(t)] - E[X_i(t)X_k(t)]}{1 - E[X_i(t)]} \right\} + o(h)$$

Since $E[X_i(t)X_k(t)] \geq E[X_i(t)]E[X_k(t)]$ is equivalent to

$$\Pr[X_i(t) = 1 | X_k(t) = 1] \geq \Pr[X_i(t) = 1]$$

we bound

$$\frac{E\left[X_k(t)\right]-E\left[X_i(t)X_k(t)\right]}{1-E\left[X_i(t)\right]}=\Pr\left[X_k(t)=1\right]\frac{1-\Pr\left[X_i(t)=1|X_k(t)=1\right]}{1-\Pr\left[X_i(t)=1\right]}$$
$$\leq \Pr\left[X_k(t)=1\right]$$

and we obtain that

$$\frac{r_i(t,h)}{r_i(t,0)}\geq 1-\delta h-\beta h\sum_{k=1}^{N}a_{ki}E\left[X_k(t)\right]+o(h)$$

We define the vector $W(t)$ and $R(t,h)$ with i-th component $w_i(t)=E\left[X_i(t)\right]$ and $r_i(t,h)$, respectively. Further, we define, only for vectors, the function $f(R)=(f(r_1),f(r_2),\ldots,f(r_N))$, then we arrive at the vector inequality

$$\frac{R(t,h)}{R(t,0)}\geq u-h\left(\delta u+\beta AW(t)\right)+o(h)$$

Just above the epidemic threshold (Van Mieghem et al., 2009), the probability of infection is $E\left[X_i(t)\right]=\epsilon\left(x_1\right)_i$, where $\epsilon>0$ is small and where x_1 is the principal eigenvector of the adjacency matrix A belonging to the largest eigenvalue λ_1. With $Ax_1=\lambda_1x_1$, the autocorrelation function just above the epidemic threshold τ_c satisfies

$$\frac{R(t,h)}{R(t,0)}\geq u-\delta h\left(u+\epsilon\tau\lambda_1x_1\right)+o(h)$$

and for the i-th component

$$\frac{r_i(t,h)}{r_i(t,0)}\geq 1-\delta h-\beta\lambda_1\left(x_1\right)_i h\epsilon+o(h)$$

I.2 The variance $\text{Var}\left[X_j(t+h)\right]$

Next, we concentrate on

$$\text{Var}\left[X_j(t+h)\right]=E\left[X_j^2(t+h)\right]-\left(E\left[X_j(t+h)\right]\right)^2=E\left[X_j(t+h)\right]-\left(E\left[X_j(t+h)\right]\right)^2$$

where the nature of a Bernoulli variable, $X_j^m = X_j$ for any positive real number m, has been used. Invoking a Taylor series approach with (9),

$$\begin{aligned}\left(E\left[X_j(t+h)\right]\right)^2 &= \left(E\left[X_j(t)\right]\right)^2 + 2hE\left[X_j(t)\right]E\left[\frac{dX_j(t)}{dt}\right] + o(h) \\ &= (1-2\delta h)\left(E\left[X_j(t)\right]\right)^2 + 2\beta h\sum_{k=1}^{N} a_{kj}E\left[X_j(t)\right]E\left[X_k(t)\right] \\ &\quad - 2\beta h\sum_{k=1}^{N} a_{kj}E\left[X_j(t)\right]E\left[X_k(t)X_j(t)\right] + o(h)\end{aligned}$$

and (11), we obtain

$$\begin{aligned}\operatorname{Var}\left[X_j(t+h)\right] &= \operatorname{Var}\left[X_j(t)\right] + h\delta\left\{\tau\sum_{k=1}^{N} a_{kj}E\left[X_k(t)\right] - E\left[X_j(t)\right]\right\} \\ &\quad - h\delta\left\{\tau\sum_{k=1}^{N} a_{kj}\left\{E\left[X_j(t)X_k(t)\right] + 2E\left[X_j(t)\right]E\left[X_k(t)\right]\right\} - 2\left(E\left[X_j(t)\right]\right)^2\right\} \\ &\quad + 2\beta h\sum_{k=1}^{N} a_{kj}E\left[X_j(t)\right]E\left[X_k(t)X_j(t)\right] + o(h)\end{aligned} \tag{14}$$

Thus, for small values of h, the variance $\operatorname{Var}[X_j(t+h)]$ seems to be approximately equal to $\operatorname{Var}[X_j(t)]$.

I.3 Just above the epidemic threshold

Just above the epidemic threshold (Van Mieghem et al., 2009), the probability of infection is $E[X_i(t)] = \epsilon (x_1)_i$, where $\epsilon > 0$ is small and where x_1 is the principal eigenvector of the adjacency matrix A belonging to the largest eigenvalue λ_1.

I.3.1 The covariance $\widetilde{\rho}\left(X_i(t), X_j(t+h)\right)$

Assuming that the effective infection rate $\tau = \frac{\beta}{\delta} = \tau_c + \epsilon$, then we may discard the last sum with triple expectations (of order $O\left(\epsilon^2\right)$) in (12) so that

$$\widetilde{\rho}\left(X_i(t), X_j(t+h)\right) \approx (1-\delta h)\widetilde{\rho}\left(X_i(t), X_j(t)\right) + \beta h\sum_{k=1}^{N} a_{kj}\widetilde{\rho}\left(X_i(t), X_k(t)\right) + o(h)$$

Since $\widetilde{\rho}\left(X_i(t), X_j(t)\right) \geq 0$ and assuming that $\widetilde{\rho}\left(X_i(t), X_k(t)\right)$ is of about the same magnitude as $\widetilde{\rho}\left(X_i(t), X_j(t)\right)$ for any node k that is a neighbor of j (this is possible because node k as a neighbor of j can only be one hop further away from or nearer to i than j. If i and j are directly connected and k is a common neighbor, then this approximation is even more accurate), then

$$\widetilde{\rho}\left(X_i(t), X_j(t+h)\right) \approx \left(1+\left(\tau d_j-1\right)\delta h\right)\widetilde{\rho}\left(X_i(t), X_j(t)\right)+o(h)$$

and

$$\frac{\rho\left(X_i(t), X_j(t+h)\right)}{\rho\left(X_i(t), X_j(t)\right)} \approx \left(1+\left(\tau d_j-1\right)\delta h\right)\frac{\sqrt{\operatorname{Var}\left[X_j(t)\right]}}{\sqrt{\operatorname{Var}\left[X_j(t+h)\right]}}+o(h)$$

Assuming $\operatorname{Var}\left[X_j(t+h)\right] \approx \operatorname{Var}\left[X_j(t)\right]$ (which is reasonable for small h) the fraction on the right hand side can be approximated by 1 and we obtain

$$\frac{\rho\left(X_i(t), X_j(t+h)\right)}{\rho\left(X_i(t), X_j(t)\right)} \approx 1+\delta h\left(\tau d_j-1\right)+o(h)$$

Finally, since $\tau > \tau_c \geq \frac{1}{\lambda_1}$, we have that $\tau d_j - 1 > \frac{d_j}{\lambda_1} - 1$. Since the spectral radius is bounded (Van Mieghem, 2011) by $\max\left(d_{av}, \sqrt{d_{\max}}\right) \leq \lambda_1 \leq d_{\max}$, where the average degree $d_{av} = \frac{2L}{N}$, the factor $\tau^* d_j - 1$ is positive for a node j with more than average degree, but possibly negative for a node j with low degree.

I.3.2 The autocorrelation function $r_i(t,h)$

Introducing $E\left[X_i(t)\right] = \epsilon\left(x_1\right)_i$ and using the eigenvalue equation $\sum_{k=1}^{N} a_{ki}\left(x_1\right)_k = \lambda_1\left(x_1\right)_i$, the autocorrelation function in (13) for small time lags h simplifies to

$$\begin{aligned} r_i(t,h) &= (1-\delta h)\left(\epsilon\left(x_1\right)_i - \epsilon^2\left(x_1\right)_i^2\right) - \beta h \epsilon^2\left(x_1\right)_i \sum_{k=1}^{N} a_{ki}\left(x_1\right)_k + O\left(\epsilon^3\right) + o(h) \\ &= (1-\delta h)\left(\epsilon\left(x_1\right)_i - \epsilon^2\left(x_1\right)_i^2\right) - \beta h \epsilon^2 \lambda_1\left(x_1\right)_i^2 + O\left(\epsilon^3\right) + o(h) \end{aligned}$$

The normalized autocorrelation function

$$\frac{r_i(t,h)}{r_i(t,0)} = 1 - \delta h - \beta h \frac{E\left[X_i(t)\right]}{\operatorname{Var}\left[X_i(t)\right]} \sum_{k=1}^{N} a_{ki}\left\{E\left[X_k(t)\right] - E\left[X_i(t) X_k(t)\right]\right\} + o(h)$$

is simplified to

$$\begin{aligned}\frac{r_i(t,h)}{r_i(t,0)} &= 1-\delta h-\frac{\beta h}{(x_1)_i}\left(\frac{\epsilon\lambda_1(x_1)_i^2+O\left(\epsilon^2\right)+o(h)}{1-\epsilon(x_1)_i}\right)\\ &= 1-\delta h-\frac{\beta h}{(x_1)_i}\left(\left(\epsilon\lambda_1(x_1)_i^2+O\left(\epsilon^2\right)+o(h)\right)\left(1+\epsilon(x_1)_i+O\left(\epsilon^2\right)\right)\right)\\ &= 1-\delta h-\frac{\beta h}{(x_1)_i}\left(\epsilon\lambda_1(x_1)_i^2+O\left(\epsilon^2\right)+o(h)\right)\end{aligned}$$

and, in terms of the effective infection rate $\tau=\frac{\beta}{\delta}$,

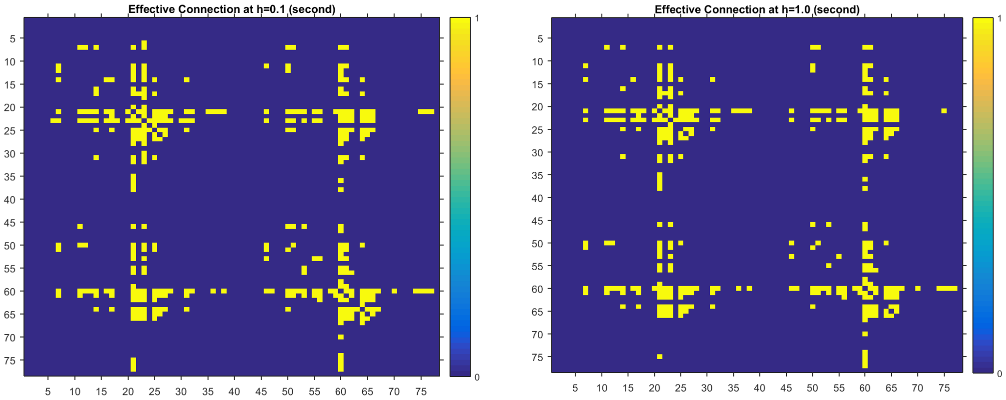
$$\begin{aligned}\frac{r_i(t,h)}{r_i(t,0)} &= 1-\delta h(1+\tau\lambda_1\epsilon(x_1)_i)+O\left(\epsilon^2\right)+o(h)\\ &= 1-\delta h+o(h)\end{aligned}$$

because, just above the epidemic threshold, $\tau\lambda_1\approx 1$. Hence, just above the epidemic threshold, there is essentially only curing as the infectious power is too weak. In addition, the normalized autocovariance can be regarded as an approximation for the auto-correlation function. Then, we have shown that the auto-correlation seems to be smaller for nodes with a higher eigenvector-centrality.

Let us regard the joint delayed probability (10) for one node i and use the property of Bernoulli random variables $E[X_i(t)X_i(t)]=E[X_i(t)]$

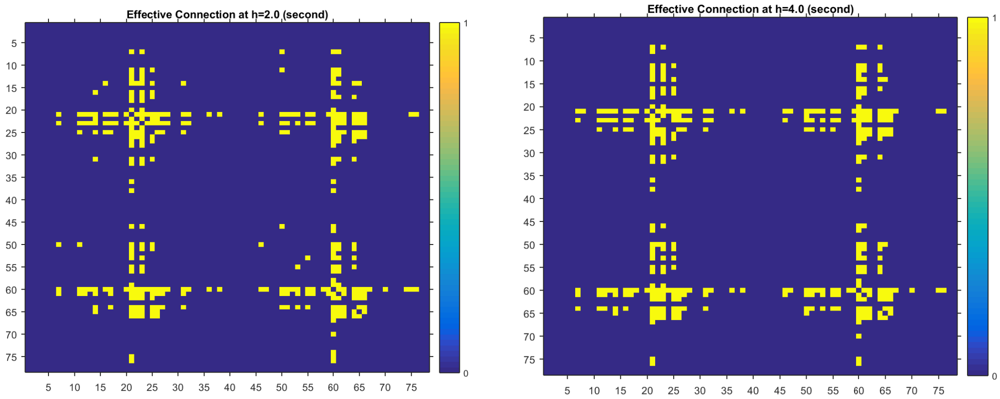
$$E\left[X_i(t)X_i(t+h)\right]=(1-\delta h)E\left[X_i(t)\right]+o(h)$$

such that just above the epidemic threshold the derivative of $E\left[X_i(t)X_i(t+h)\right]$ is $-\delta\epsilon(x_1)_i$.



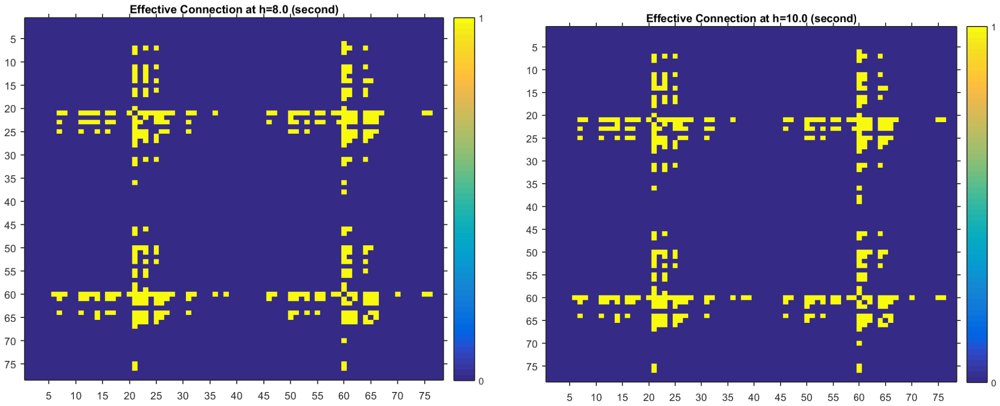
(a) Effective connectivity for $h = 0.1s$.

(b) Effective connectivity for $h = 1s$.



(c) Effective connectivity for $h = 2s$.

(d) Effective connectivity for $h = 4s$.



(e) Effective Connectivity for $h = 8s$.

(f) Effective Connectivity for $h = 10s$.

Figure 8: Effective connectivity for different time lags h.

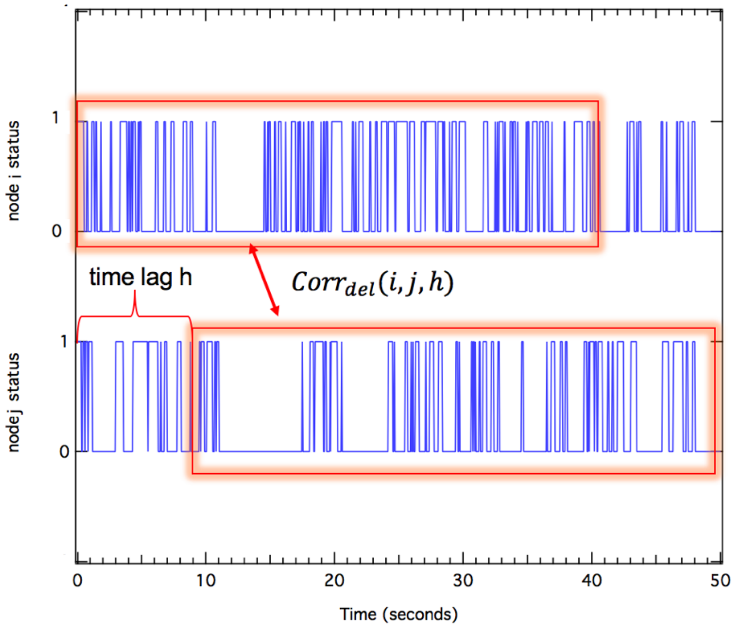


Figure 9: Schematic explanation of how the delayed correlation is calculated in our analysis.

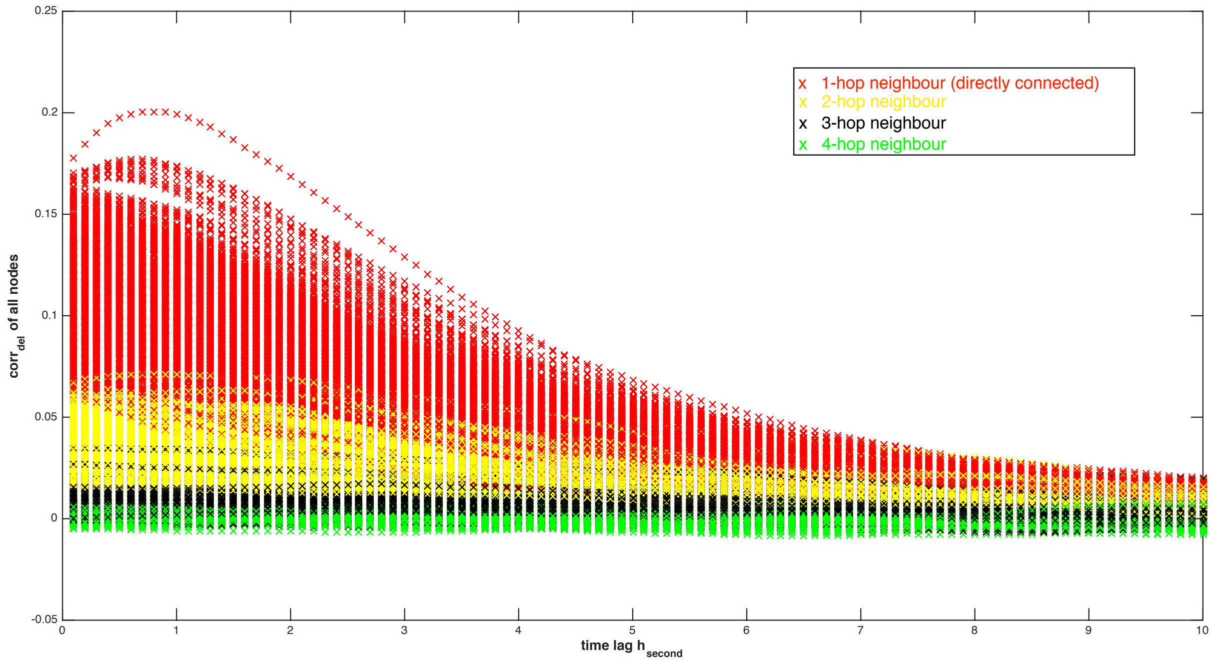


Figure 10: Delayed correlation of all node pairs. The delayed correlations for different time lags h are colored in red for all the neighbors, in yellow for all nodes that are 2 hops away, in black for nodes with hopcount 3 and green for hopcount 4. For small time lags it seems that the direct neighbors of a node have the highest delayed correlation and that the further away another node is in terms of hopcount the lower is the delayed correlation between that node and the regarded node.

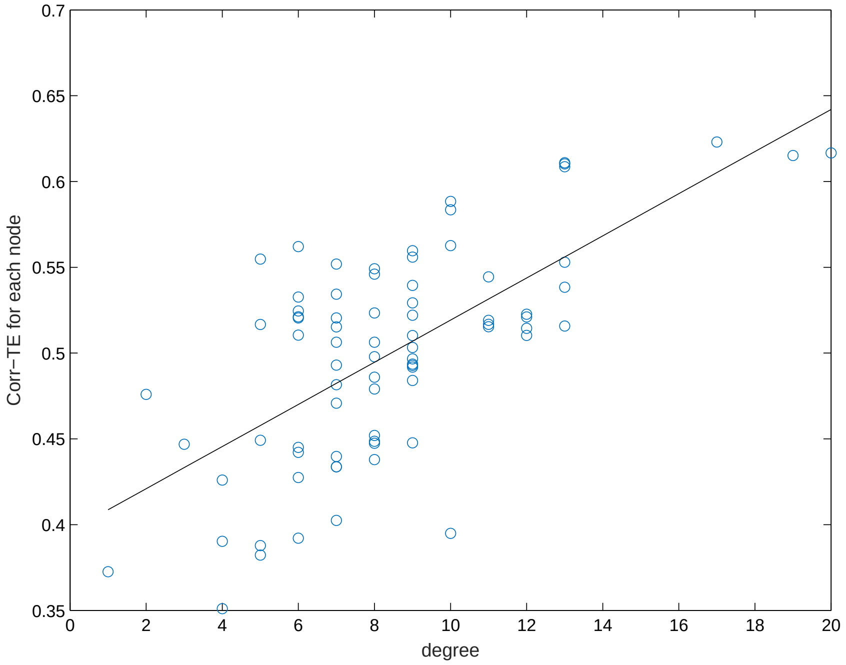


Figure 11: Delayed correlation of a node (averaged over its influence on all other nodes) versus its degree and normalized as in Equation (7).

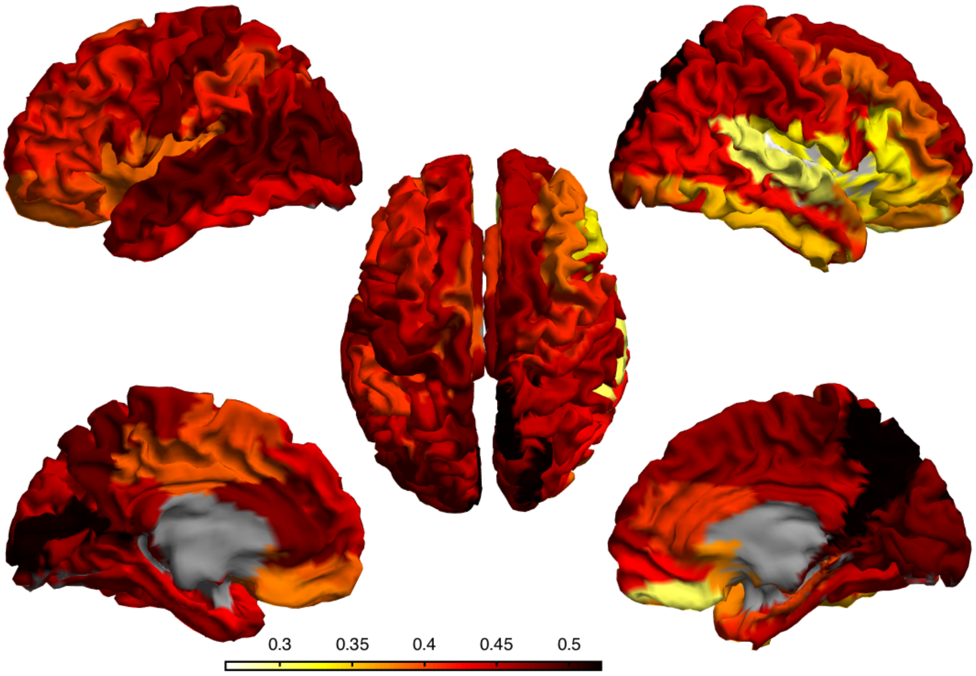


Figure 12: Plot of the averaged normalized delayed correlation (see Equation (7)) for all nodes of the underlying structural network. Visualization of the directed delayed correlation for each brain region on the parcellated template brain, where darker colors represent strong sending brain regions and lighter colors a more receiving property of a region. We show the brain here in clockwise order from the left, top, right, right midline and left midline.

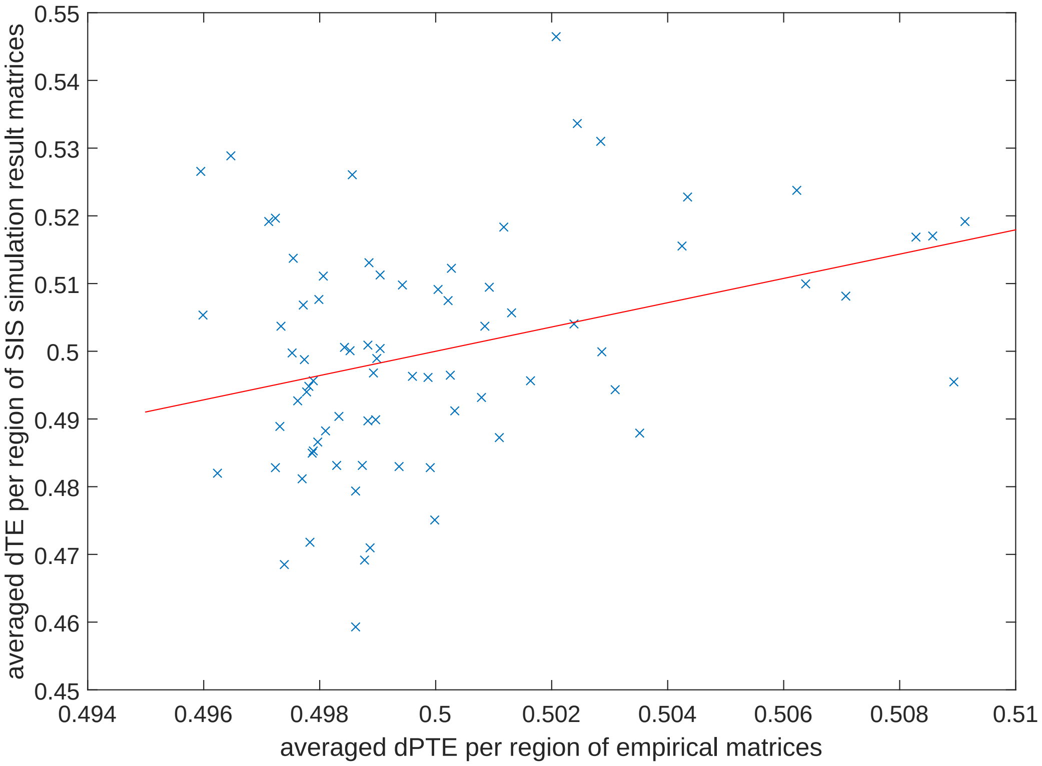


Figure 13: Scatterplot of the empirically observed directed Phase Transfer Entropy results from Hillebrand et al. (2016) and our averaged dTE values. We obtain a positive correlation coefficient of 0.334 for the time delay of $h = 2.9$ time units.

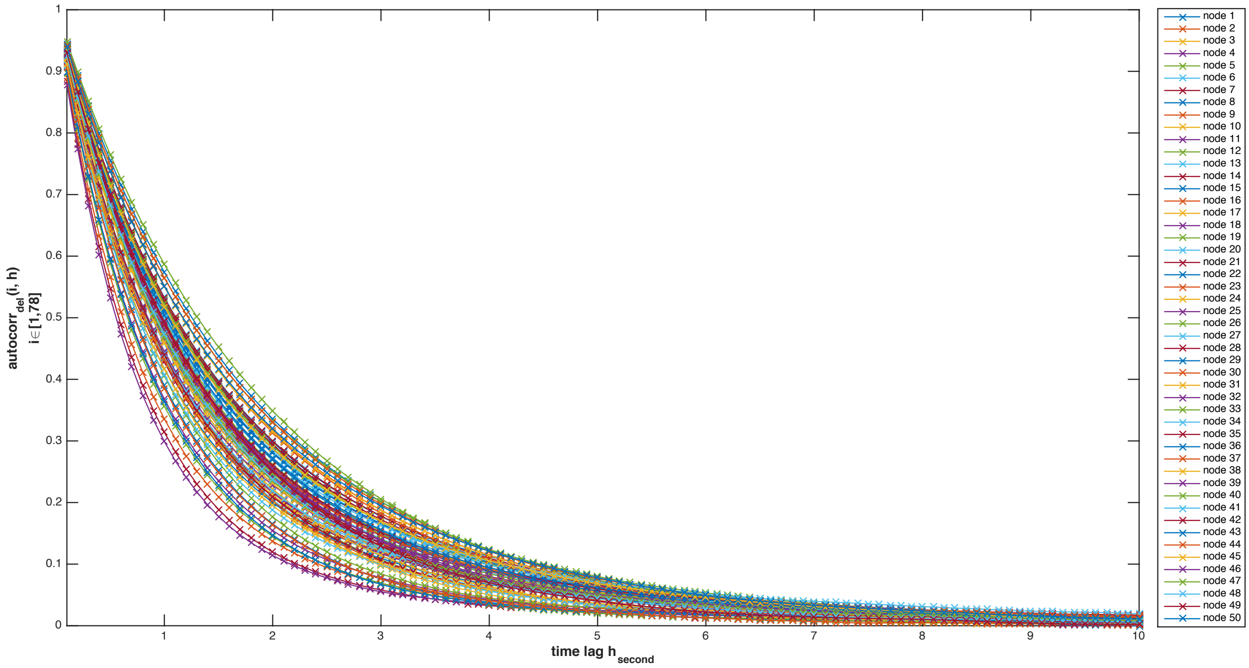


Figure 14: Delayed auto-correlation of all nodes.

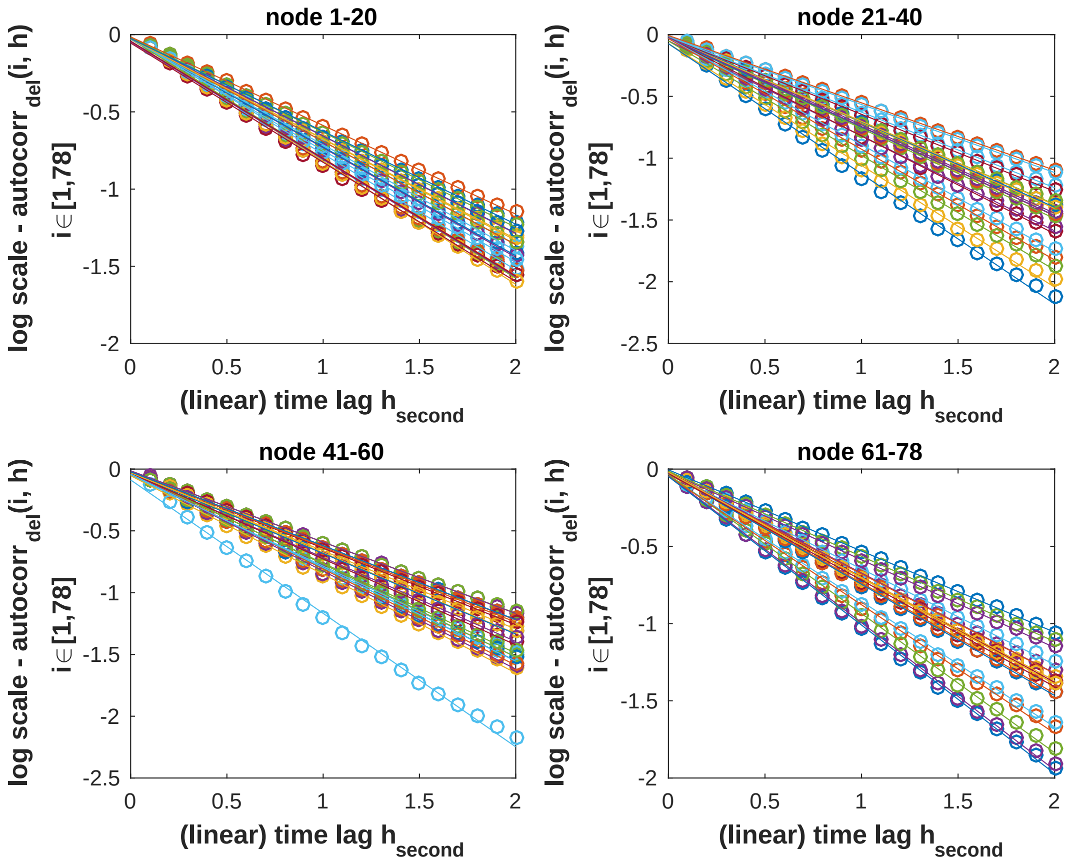


Figure 15: Delayed auto-correlation of all nodes for very small time lags h on a log-linear scale (zoomed-in plot of Figure 14). For small time lags, the auto-correlation seems to decay exponentially fast with h.

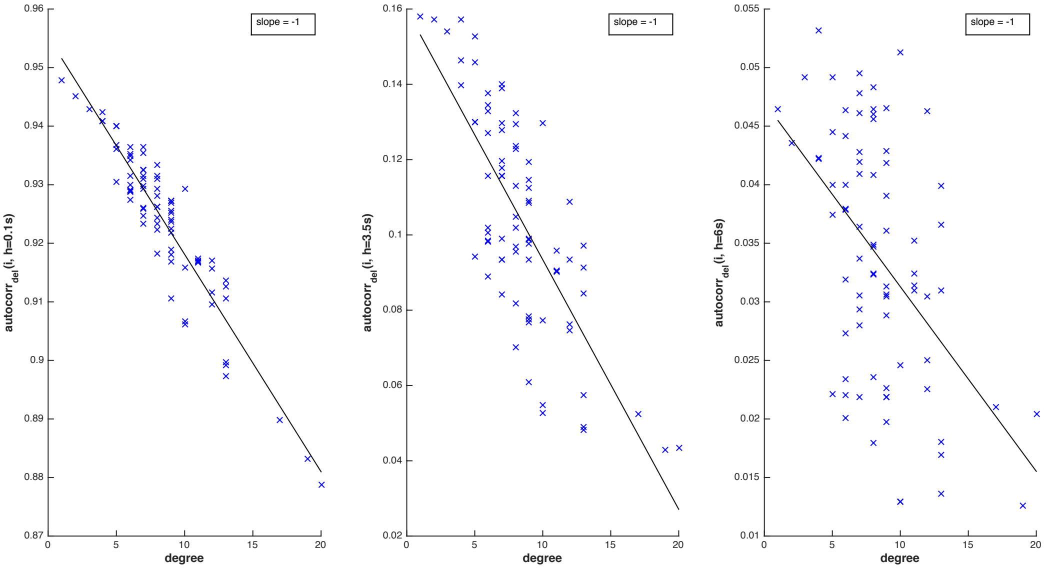


Figure 16: Scatterplots for different delays h of the delayed auto-correlation of all nodes. The delayed auto-correlation for all time lags h seems to be negatively correlated with the degree (more visible for small time lags h, left panel), meaning that hubs have a lower auto-correlation than low-degree nodes.

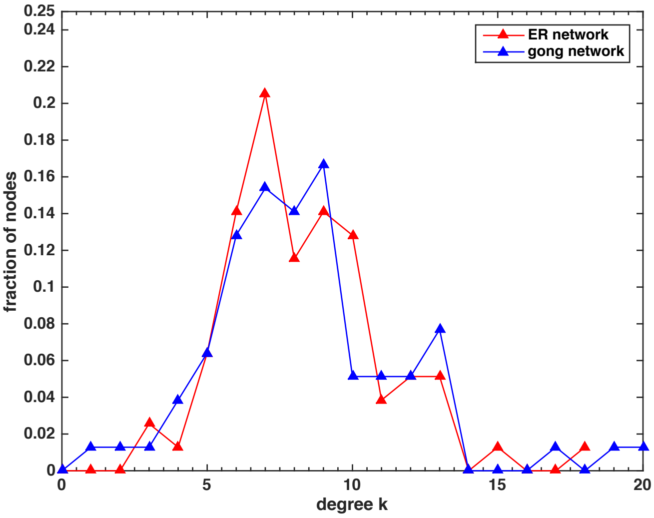
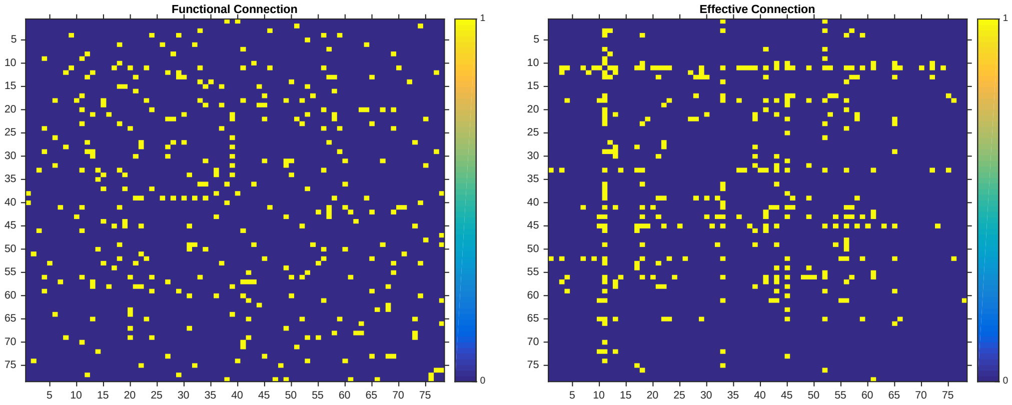


Figure 17: Degree distribution of the structural brain network and an ER random graph with similar properties.


(a) Functional Connectivity.

(b) Effective Connectivity.

Figure 18: (a) Functional Connectivity based on the SIS epidemics time series with $\beta = 0.1$ and $\delta = 1$ on an ER graph. (b) Effective Connectivity based on the SIS epidemics time series with $\beta = 0.1$ and $\delta = 1$ on an ER graph with delay $h = 0.1s$, which is one time step in our simulations.

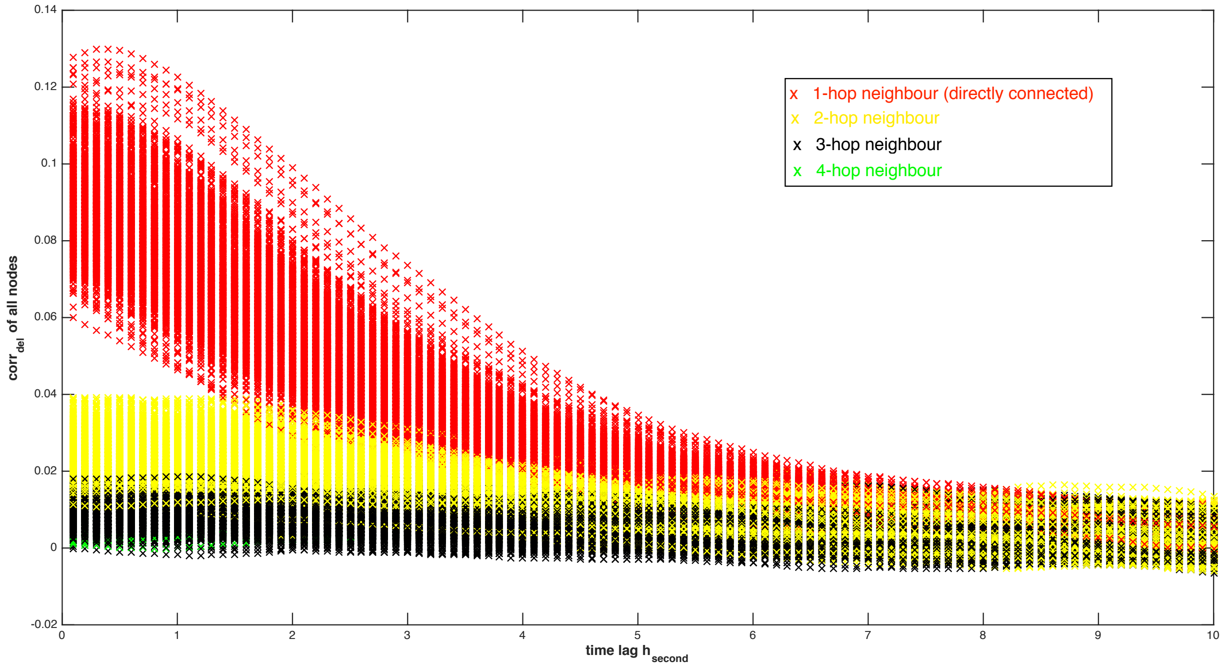


Figure 19: Delayed correlation of all node pairs for an ER graph. The delayed correlations for different time lags h are colored in red for all the neighbors, in yellow for all nodes that are 2 hops away, in black for nodes with hopcount 3 and green for hopcount 4. For small time lags it seems that the direct neighbors of a node have the highest delayed correlation and that the further away another node is in terms of hopcount the lower is the delayed correlation between that node and the regarded node.

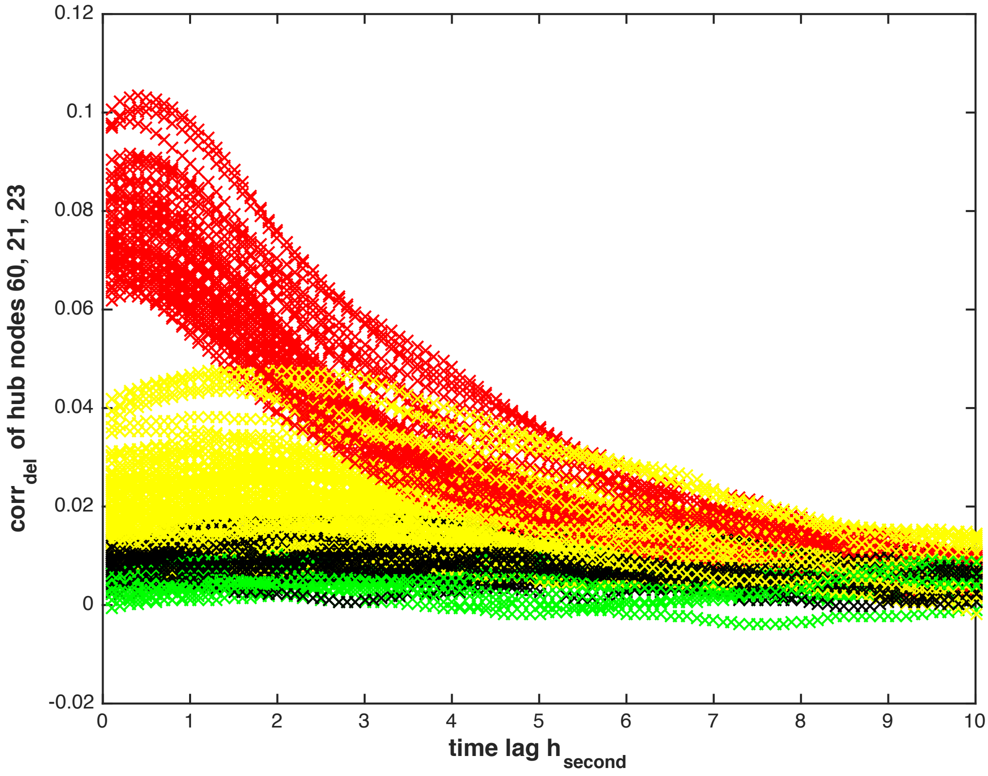


Figure 20: Delayed correlation of the 3 nodes with the highest degrees with the rest of the network. The delayed correlations for different time lags *h* are colored in red for all the neighbors, in yellow for all nodes that are 2 hops away, in black for nodes with hopcount 3 and green for hopcount 4. For small time lags it seems that the direct neighbors of a hub have a much higher delayed correlation with that node than the nodes that are further away.

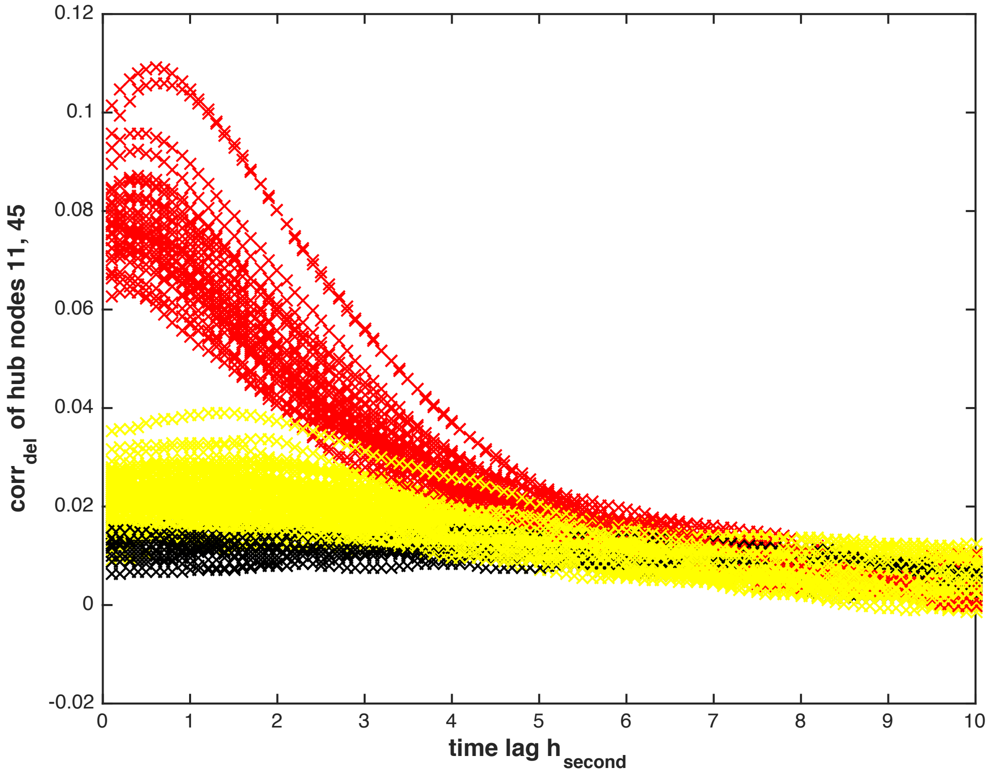


Figure 21: Delayed correlation of the 2 nodes with the highest degrees with the rest of the network for an ER graph. The delayed correlations for different time lags *h* are colored in red for all the neighbors, in yellow for all nodes that are 2 hops away and in black for nodes with hopcount 3. For small time lags it seems that the direct neighbors of a hub have a much higher delayed correlation with that node than the nodes that are further away.

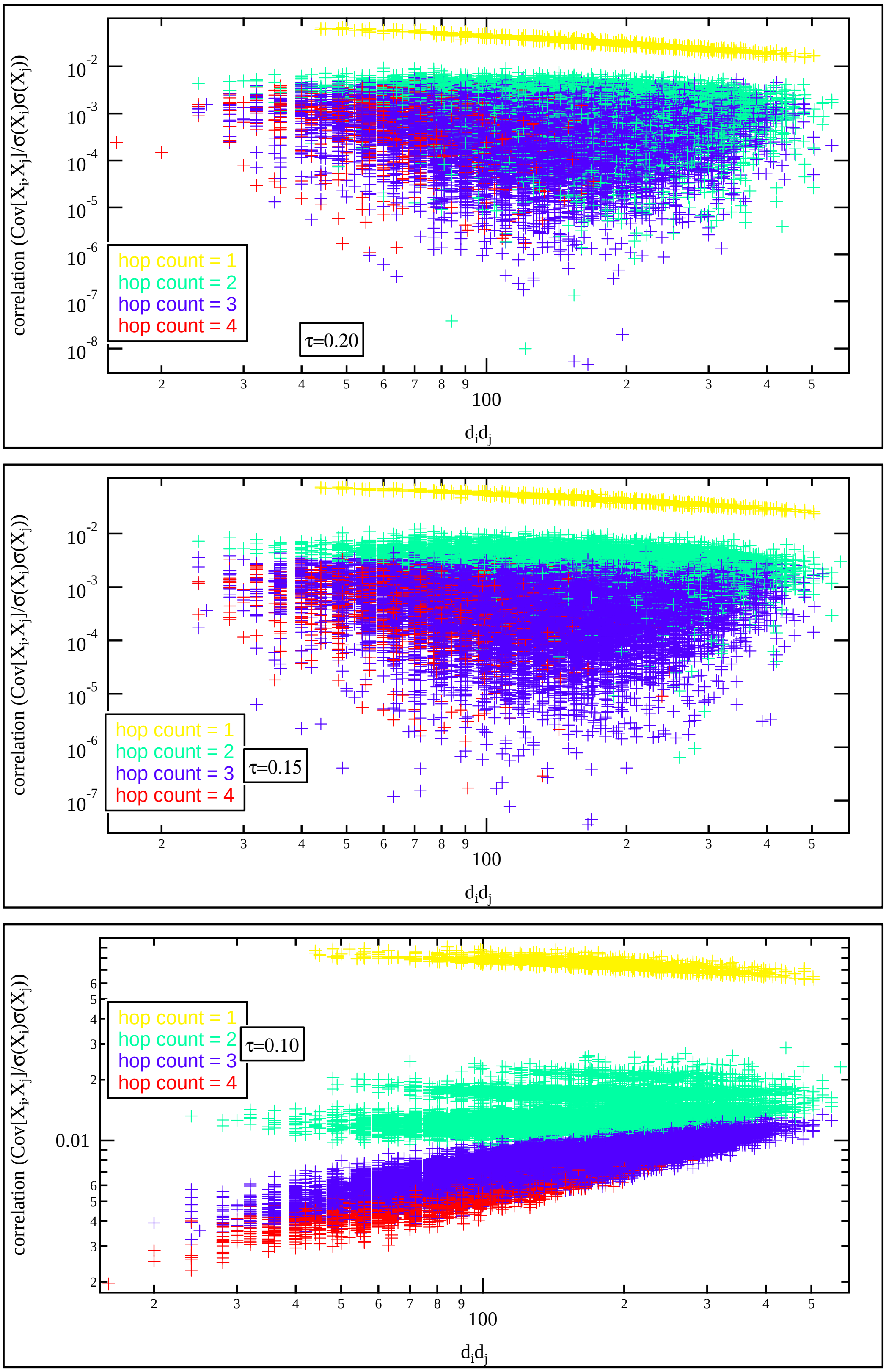


Figure 22: Cross-correlation versus the degree product for different hopcounts for an ER graph with 500 nodes.

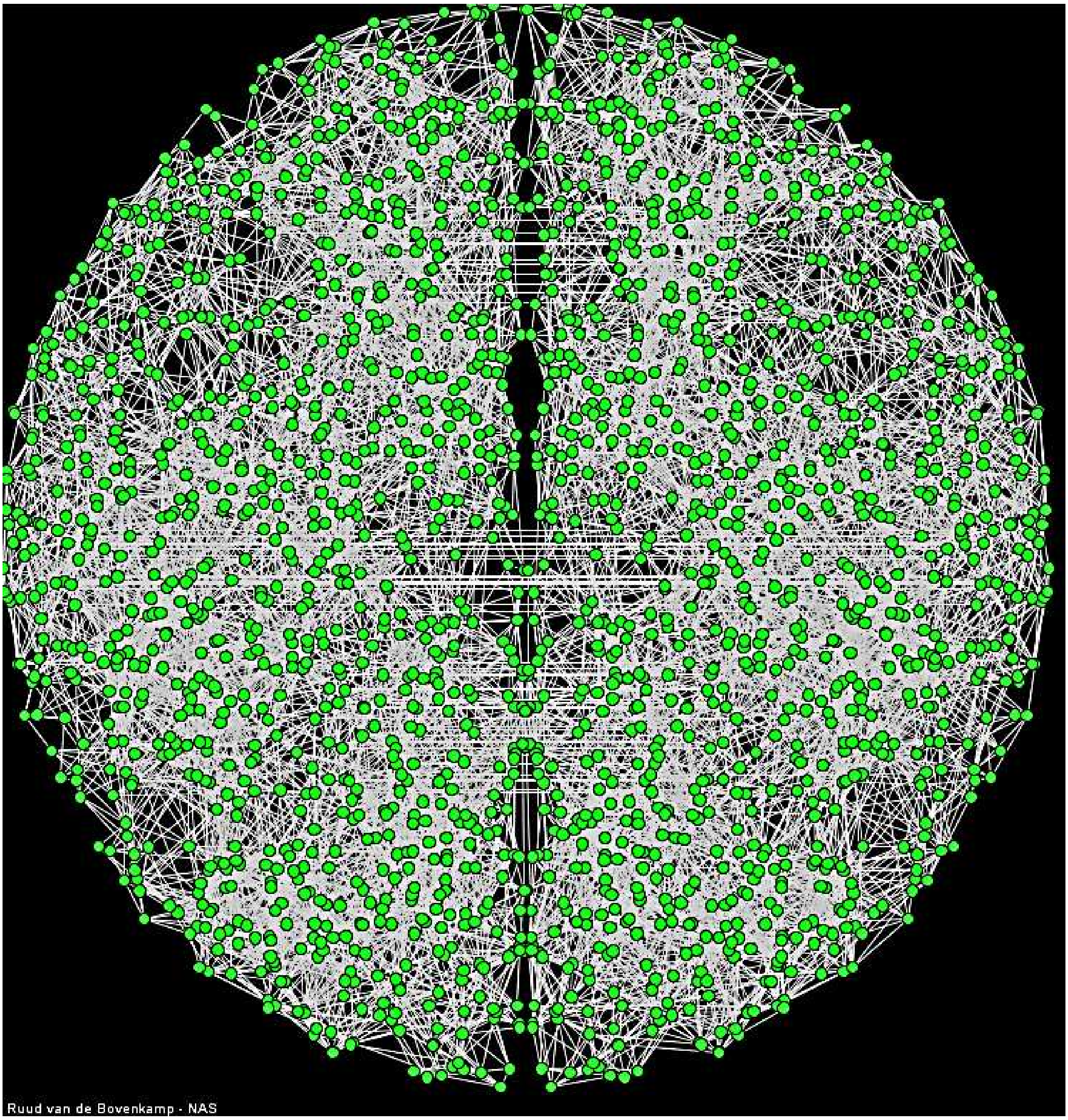


Figure 23: Generated structural brain network from model described earlier.

Bibliography

[1] M. Newman, *Networks: an introduction.* Oxford university press, 2010.

[2] R. Cohen and S. Havlin, *Complex networks: structure, robustness and function.* Cambridge University Press, 2010.

[3] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[4] C. J. Stam, "Modern network science of neurological disorders," *Nature Reviews Neuroscience*, vol. 15, no. 10, pp. 683–695, 2014.

[5] P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1–14, 2009.

[6] L. J. Allen, F. Brauer, P. Van den Driessche, and J. Wu, *Mathematical epidemiology.* Springer, 2008.

[7] C. Li, R. van de Bovenkamp, and P. Van Mieghem, "Susceptible-infected-susceptible model: A comparison of n-intertwined and heterogeneous mean-field approximations," *Physical Review E*, vol. 86, no. 2, p. 026116, 2012.

[8] P. Van Mieghem, *Performance analysis of complex networks and systems.* Cambridge University Press, 2014.

[9] ——, *Graph spectra for complex networks.* Cambridge University Press, 2010.

[10] R. J. Wilson, "An eulerian trail through königsberg," *Journal of graph theory*, vol. 10, no. 3, pp. 265–275, 1986.

[11] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[12] D. J. Watts and S. H. Strogatz, “Collective dynamics of 'small-world' networks,” *nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[13] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *science*, vol. 286, no. 5439, pp. 509–512, 1999.

[14] I. de Sola Pool and M. Kochen, “Contacts and influence,” *Social networks*, vol. 1, no. 1, pp. 5–51, 1978.

[15] M. Faloutsos, P. Faloutsos, and C. Faloutsos, “On power-law relationships of the internet topology,” in *ACM SIGCOMM computer communication review*, vol. 29. ACM, 1999, pp. 251–262.

[16] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, “Epidemic processes in complex networks,” *Reviews of modern physics*, vol. 87, no. 3, p. 925, 2015.

[17] R. van de Bovenkamp and P. Van Mieghem, “Survival time of the susceptible-infected-susceptible infection process on a graph,” *Physical Review E*, vol. 92, no. 3, p. 032806, 2015.

[18] N. Kaur, M. Ghosh, and S. Bhatia, “Modeling and analysis of an sirs epidemic model with effect of awareness programs by media,” *Int J Math Comput Nat Phys Eng*, vol. 8, no. 1, pp. 233–239, 2014.

[19] J. C. Miller and E. M. Volz, “Incorporating disease and population structure into models of sir disease in contact networks,” *PloS One*, vol. 8, no. 8, p. e69162, 2013.

[20] R. Van de Bovenkamp, *Epidemic Processes on Complex Networks: Modelling, Simulation and Algorithms*. TU Delft, Delft University of Technology, 2015.

[21] C. Stam, E. van Straaten, E. Van Dellen, P. Tewarie, G. Gong, A. Hillebrand, J. Meier, and P. Van Mieghem, “The relation between structural and functional connectivity patterns in complex brain networks,” *International Journal of Psychophysiology*, 2015.

[22] E. Bullmore and O. Sporns, “The economy of brain network organization,” *Nature Reviews Neuroscience*, vol. 13, no. 5, pp. 336–349, 2012.

[23] (2014) Human connectome project. [Online]. Available: http://www.humanconnectomeproject.org

[24] (2016) Organization for human brain mapping. [Online]. Available: http://www.humanbrainmapping.org

[25] P. Hagmann, M. Kurant, X. Gigandet, P. Thiran, V. J. Wedeen, R. Meuli, and J.-P. Thiran, “Mapping human whole-brain structural networks with diffusion mri,” *PloS one*, vol. 2, no. 7, p. e597, 2007.

[26] G. Gong, Y. He, L. Concha, C. Lebel, D. W. Gross, A. C. Evans, and C. Beaulieu, “Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography,” *Cerebral cortex*, vol. 19, no. 3, pp. 524–536, 2009.

[27] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, “Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain,” *Neuroimage*, vol. 15, no. 1, pp. 273–289, 2002.

[28] B. Mišić, R. F. Betzel, A. Nematzadeh, J. Goñi, A. Griffa, P. Hagmann, A. Flammini, Y.-Y. Ahn, and O. Sporns, “Cooperative and competitive spreading dynamics on the human connectome,” *Neuron*, vol. 86, no. 6, pp. 1518–1529, 2015.

[29] W. L. Shew and D. Plenz, “The functional benefits of criticality in the cortex,” *The neuroscientist*, vol. 19, no. 1, pp. 88–100, 2013.

[30] D. S. Bassett and E. Bullmore, “Small-world brain networks,” *The neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006.

[31] M. Kaiser, “A tutorial in connectome analysis: topological and spatial features of brain networks,” *Neuroimage*, vol. 57, no. 3, pp. 892–907, 2011.

[32] C. Honey, O. Sporns, L. Cammoun, X. Gigandet, J.-P. Thiran, R. Meuli, and P. Hagmann, “Predicting human resting-state functional connectivity from structural connectivity,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 6, pp. 2035–2040, 2009.

[33] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, “Network structure of cerebral cortex shapes functional connectivity on multiple time scales,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 10 240–10 245, 2007.

[34] J. Meier, P. Tewarie, A. Hillebrand, L. Douw, B. W. van Dijk, S. M. Stufflebeam, and P. Van Mieghem, “A mapping between structural and functional brain networks,” *Brain connectivity*, vol. 6, no. 4, pp. 298–311, 2016.

[35] E. C. van Straaten and C. J. Stam, “Structure out of chaos: functional brain network analysis with eeg, meg, and functional mri,” *European Neuropsychopharmacology*, vol. 23, no. 1, pp. 7–18, 2013.

[36] J. Lee Rodgers and W. A. Nicewander, “Thirteen ways to look at the correlation coefficient,” *The American Statistician*, vol. 42, no. 1, pp. 59–66, 1988.

[37] E. Pereda, R. Q. Quiroga, and J. Bhattacharya, “Nonlinear multivariate analysis of neurophysiological signals,” *Progress in neurobiology*, vol. 77, no. 1, pp. 1–37, 2005.

[38] H. Hinrichs, H. Heinze, and M. Schoenfeld, “Causal visual interactions as revealed by an information theoretic measure and fmri,” *NeuroImage*, vol. 31, no. 3, pp. 1051–1060, 2006.

[39] B. Chai, D. Walther, D. Beck, and L. Fei-Fei, “Exploring functional connectivities of the human brain using multivariate information analysis,” in *Advances in neural information processing systems*, 2009, pp. 270–278.

[40] M. Chávez, J. Martinerie, and M. Le Van Quyen, “Statistical assessment of nonlinear causality: application to epileptic eeg signals,” *Journal of neuroscience methods*, vol. 124, no. 2, pp. 113–128, 2003.

[41] A. Hillebrand, P. Tewarie, E. van Dellen, M. Yu, E. W. Carbo, L. Douw, A. A. Gouw, E. C. van Straaten, and C. J. Stam, “Direction of information flow in large-scale resting-state networks is frequency-dependent,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 14, pp. 3867–3872, 2016.

[42] E. Cator and P. Van Mieghem, “Nodal infection in markovian susceptible-nfected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated,” *Physical Review E*, vol. 89, no. 5, p. 052802, 2014.

[43] R. Vicente, M. Wibral, M. Lindner, and G. Pipa, “Transfer entropyâĂŤa model-free measure of effective connectivity for the neurosciences,” *Journal of computational neuroscience*, vol. 30, no. 1, pp. 45–67, 2011.

[44] J. T. Lizier, J. Heinzle, A. Horstmann, J.-D. Haynes, and M. Prokopenko, “Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fmri connectivity,” *Journal of Computational Neuroscience*, vol. 30, no. 1, pp. 85–107, 2011.

[45] T. Schreiber, “Measuring information transfer,” *Physical review letters*, vol. 85, no. 2, p. 461, 2000.

[46] M. Lobier, F. Siebenhühner, S. Palva, and J. M. Palva, “Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions,” *Neuroimage*, vol. 85, pp. 853–872, 2014.

[47] M. P. Van Den Heuvel and O. Sporns, “Rich-club organization of the human connectome,” *The Journal of neuroscience*, vol. 31, no. 44, pp. 15 775–15 786, 2011.

[48] J. Meier, X. Zhou, A. Hillebrand, C. Stam, and P. Van Mieghem, “Applying the epidemic spreading model to explain effective connectivity in brain networks,” accepted in Conference on Complex Systems (CCS) and will be presented at CCS on September 22, 2016 in Amsterdam.