Generative Models of Mental Disorders

Jessica Toudic

Trinity 2021

1   Introduction

In 2020, the National Institute of Mental Health (NIMH) published a 5-year strategic plan for research, the first goal of which is to 'define the brain mechanisms underlying complex behaviours' relating to mental health disorders [1]. This goal includes an objective to 'identify and characterise the neural circuit mechanisms contributing to human behaviour and their disruption in mental illness' [1]. Therefore, understanding the mechanisms leading to the development of mental health disorders is a research priority.

Mental illnesses are considered to be a result of dysfunctional connections in the brain, established as a result of abnormal brain development [2, 3, 4, 5, 6]. This view of mental disorders was first proposed by Wernicke [7], who suggested that the core of pathophysiology in the brain is the alteration of complex interactions between brain regions [2], and has since been further elucidated by more recent neuroimaging studies [2, 3]. Overall, there has been a marked paradigm shift for mental disorders away from the disease model medicine, which sought to identify a single focal 'lesion' as the cause of mental illness [4], towards a network perspective.

The emerging field of network neuroscience (NN) is ideally suited to studying such system-level dysfunction: brain regions are represented by nodes, and advanced neuroimaging techniques, such as functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI), are used to define connections between brain regions. These connections are represented as edges between the nodes, forming a 'brain network' [4]. Tools from network science can then be employed to analyse these networks, enabling the description of brain networks in terms of a vast set of network metrics [3].

However, despite its success in characterising differences in connectivity between healthy individuals and those with psychiatric disorders [8, 9, 10], NN does not contribute to the NIMH objective to identify the (patho-)physiological mechanisms underlying mental illness: network neuroscience focuses on the descriptive level, rather than the mechanistic level [2]. That is, NN describes features of a network, but doesn't explain the cause of such features [2, 4]. Therefore, to achieve the NIMH goal, a shift of focus from describing how one network is different to another (descriptive) to why one network is different from another (mechanistic) is required [2].

Such a shift in focus can be addressed by investigating generative models of network development [2]. In the context of brain networks, generative modelling involves computationally implementing biologically-motivated parameterised wiring rules for connections (edges) between brain regions (nodes), creating synthetic networks that hope to emulate network properties of empirical brain networks [4, 11, 12]. In this way, potential mechanisms of brain network development can be identified. In particular, possible mechanisms resulting in dysfunctional connections reflective of psychiatric disorders can be found, allowing progress towards the NIMH goal to be made.

In this report, we explore generative models of mental illness at the system-level of the brain. This first requires an understanding of generative models of normal brain networks, which in itself necessitates knowledge of what constitutes a 'normal' brain. Therefore, we begin by providing a short introduction to network neuroscience and characteristics of brain networks in Section 2. This allows us to introduce generative modelling in the context of neuroscience in Section 3. There, we also explain how generative modelling can be used to understand psychiatric disorders. With this knowledge in hand, we then explain current understanding of generative models of brain networks in Section 4, by focusing on the details of particular models. This brings us to a focus on generative models in the context of psychiatric disorders in Section 5. We conclude the report by discussing limitations of generative modelling and suggest possible future directions, both in terms of methods and progress towards the NIMH goal.

2 Network Neuroscience

We begin by introducing the necessary background concerning network neuroscience, as it allows us to identify characteristics of normal brain networks.[1]

2.1 Introducing Network Neuroscience

Network neuroscience (NN) originates from the mathematical field of graph theory, which represents interconnected systems in terms of nodes and edges [2, 4]. Overall, NN involves applying network-based methods to neuroimaging data in order to study the brain: regions of the brain are represented by nodes,[2] and relationships between them (determined by neuroimaging) are represented by edges [4]. An example of such a network is shown in Figure 1.

The relationships between brain regions can be structural (represent white-matter fibre tracts between neuronal populations [13]) or functional (represent statistical relationships between activation time-series of populations of neurons making up a brain region [2, 13]).[3] Functional connectivity measures statistical correlations without an explicit causal affect, and so is referred to as 'model-free' [13].

The pattern of connections is known as the network's 'topology', and is amenable analysis by using tools from network science: the network can be quantified in terms of its topological features. This provides information about the network not readily apparent from a simple list of nodes and edges, or by looking at a visual depiction of the network [2]. Furthermore, these features directly relate to the dynamics that the network is able to support and hence also reflect the brain network's biological function [2, 4]. In particular, the topological network features provide information about the biological function of the brain that can't be gained from *in vitro*[4] studies or by looking at regions of the brain in isolation; without NN, our understanding of how regions of the brain relate to one another at both functional and structural levels would be limited, and our understanding of psychiatric disorders as disorders of large-scale brain connectivity would still be very minimal. For a detailed introduction to networks and their topological features, we refer the reader to [15].
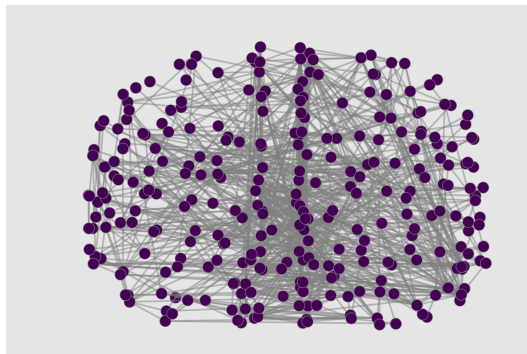

Figure 1: Network representation of a brain. Nodes (purple dots) represent brain regions, and edges (grey lines) represent functional relationships between brain regions. We constructed this network using fMRI data from [6].

---

[1] In this report, whenever we say 'brain network', we mean a system-level network representation of the brain. That is, nodes represent large-scale brain regions.

[2] Nodes can also be defined at many other scales, including as small-scale as individual neurons. In this report, we will be focusing on system-level dysconnectivity in mental disorders, and hence will only consider nodes that represent large-scale brain regions.

[3] Other forms of connectivity exist too, such as effective connectivity [14], but we will not consider such connectivity in this report.

[4] Studies where the brain is outside of a living organism.

2.2 Main Results of Network Neuroscience

During the last two decades, efforts of network neuroscience to characterise structural and functional brain networks using network metrics has culminated in a small set of attributes that collectively characterise most brain networks, since they are consistently observed across neuroimaging studies [4, 11, 16]. These attributes are: (1) small-world architecture, (2) modular community structure, and (3) existence of hubs in a rich-club structure. These features will become important when we consider both motivation for generative mechanisms and how to compare synthetic networks to empirical brain networks, so we now explain each of these attributes mathematically, and relate them to their neurobiological meaning. We first define a network. These definitions can be found in [17].

Definition 1 (Network). A network $\mathcal{G} = (V, E)$ is defined by a set of n nodes $V = \{v_1, \ldots, v_n\}$ and a set of m edges $E = \{e_1, \ldots, e_m\}$, where each edge is a pair of nodes $(v_i, v_j) \in E$. To each edge $e_i \in E$ we can assign a weight $w_i \in \mathbb{R}_{>0}$.

2.2.1 Small-world Architecture

A network with 'small-world architecture' is characterised by: (1) a short average path length, and (2) a high clustering coefficient.

Definition 2 (Average Path Length). The path length $d(v_i, v_j)$ between two nodes $v_i, v_j \in E$ is the length of the shortest sequence of edges between v_i and v_j. If the network is weighted, the distance is the sum of weights of those edges. The average path length L is then defined by

$$L = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{i-1} d(v_i, v_j). \tag{1}$$

Therefore, the average path length is the sum of path lengths between all pairs of nodes in the network, normalised by the total number of pairs of nodes: $\binom{n}{2} = n(n-1)/2$. In practice, a measure of global network efficiency is calculated instead of average path length, as it is numerically easier to use [16].

Definition 3 (Efficiency). The efficiency E of a network is defined by

$$E = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{i-1} \frac{1}{d(v_i, v_j)}. \tag{2}$$

That is, the efficiency is inversely related to path length; a short average path length is equivalent to a high efficiency.

Definition 4 (Clustering Coefficient). The local clustering coefficient C_i of node $v_i \in V$ is defined by

$$C_i = \frac{\text{number of triangles including the } i\text{-th node}}{k_i(k_i - 1)/2}. \tag{3}$$

The (average) clustering coefficient of the network is then

$$C = \frac{1}{n} \sum_{i=1}^{n} C_i. \tag{4}$$

That is, the local clustering coefficient measures the proportion of connected triples (a, b, c) that have all three edges $(a, b), (b, c), (a, c)$ present, and is hence a measure of local 'density'. The normalisation factor $k_i(k_i - 1)/2$ is the total number of pairs of neighbours of node v_i, and hence the number of triples v_i belongs to. This can be seen since k_i is the *degree* of node v_i (the number of edges incident to it, which is equal to the number of neighbours it has), and so the number of pairs of neighbours is $\binom{k_i}{2}$. The clustering coefficient is an average over all n nodes in the network. Figure 2 shows a visual representation of nodes with low and high local clustering coefficients.

Given these two definitions, it becomes clear that a short average path length means that on average two nodes are only a short distance apart, and a high clustering coefficient means that the network is locally dense. Neurobiologically, these features can be attributed to a balance between the rapid transmission of information on the network and efficient local processing [2, 4, 16]. In particular, efficiency measures how efficiently the brain network exchanges information [18]. These were some of the earliest features of brain networks to be observed [4].
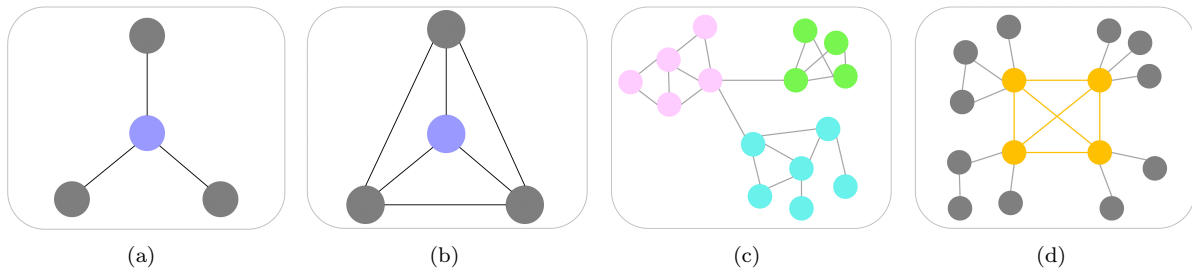
Figure 2: Diagrams explaining network features. (a) Purple node has local clustering coefficient 0 as no triples containing it have all three edges. (b) Purple node has local clustering coefficient 1 because all triples containing it have three edges. (c) Network displaying modular structure; each colour represents a community. (d) The orange nodes form a rich-club.

2.2.2 Modular Structure

A network exhibiting modular structure, otherwise known as community structure, is separated into distinct 'communities' of high local connectivity, and there exists only sparse connectivity between communities. Figure 2 (c) gives a visual representation of a network displaying community structure.

In the context of brain networks, modular structure 'facilitates optimal segregation of information' [2, 4]. In particular, it allows functional specialisation within communities and coordination between communities; it has been proposed that this may facilitate cognition [4].

2.2.3 Hubs and Rich-club Structure

A hub is a node with a comparatively higher degree than the rest of the nodes in the network. Brain networks have been shown to contain hubs that form multiple long-range connections between distant brain regions, are distributed across communities, and are densely connected to one another, overall forming a 'rich-club' [4]. Figure 2 (d) gives a visual representation of such a rich-club structure. The presence of hubs is reflected in a heavy-tailed degree distribution, where the majority of nodes have few connections (low degree), but the number of nodes with high degree is higher than one would expect from the exponential distribution, therefore making the tail 'heavy' [4]. Figure 3 demonstrates this visually.
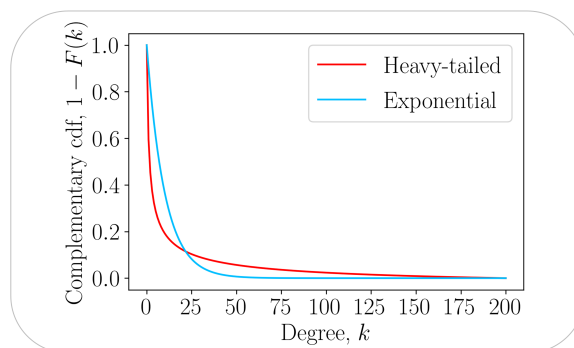

Figure 3: Illustration of a heavy-tailed degree distribution. The complementary cumulative degree distribution $1 - F(k)$, where $F(k)$ is the cumulative degree distribution, is plotted against degree for an exponential distribution $P(k) \propto \exp(-k)$ and a power-law distribution $P(k) \propto k^{-2}$. The higher density of high degrees in the power-law distribution than the exponential makes it 'heavy-tailed'.

Biologically, rich-club nodes allow rapid transmission of information across the brain, and act as a 'backbone' for integrating information between communities [4].
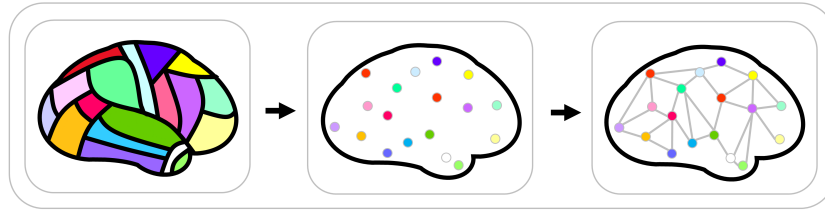
Figure 4: Constructing a brain network: a parcellation of the brain is defined, which defines the nodes of the network. Neuroimaging is then conducted, and analysis of this data is used to define the edges. Diagram inspired from [21].

2.3 Creating Brain Networks

In order to create a brain network, one must define the nodes (brain regions) and the edges (relationships between brain regions). Figure 4 demonstrates the process visually.

Nodes are defined by a parcellation of the brain, that is, a division of the brain into N regions. Typically, the number of different regions N is between 100 and 1000 [4, 13]. Parcellation schemes vary from study to study, but the more nodes, the more computationally demanding the analysis of the brain network [4].

Edges are defined by neuroimaging data, and hence depend on the neuroimaging method employed. Functional brain networks can be imaged using a variety of neuroimaging techniques, each of which give different measures of functional connectivity [13]. One of the most popular methods is functional magnetic resonance imaging (fMRI) due to its high spatial resolution (of the order of millimetres cubed) [3]. fMRI captures blood-oxygen-level-dependent (BOLD) signals as a measure of the underlying functional brain activity [13, 16, 19]: in an MRI scanner, hydrogen nuclei align with a strong magnetic field, and a radio wave is emitted and absorbed by the hydrogen nuclei. This knocks the hydrogen nuclei out of alignment, and when they return to their original position, they emit radio waves (magentic resonance (MR) signals), which are detected by the scanner. When a brain region is active it requires more oxygen, and this demand is met by an increased flow of oxygenated blood (blood with oxygen bound to the haemoglobin) to the area. Oxygenated and de-oxygenated haemoglobin have different magnetic properties, hence resulting in different MR signals, and therefore producing a distinction between active and inactive brain regions. Two brain regions are connected in a functional brain network if their time-averaged[5] BOLD time-series are highly correlated, even if the regions are not connected anatomically [5]. However, fMRI lacks a good temporal resolution, which other techniques such as electroencephalography (EEG) do possess [16].

Structural connections represent white-matter fibre tracts (anatomical connections) between brain regions, and can be measured by diffusion tensor imaging (DTI). DTI measures the diffusion of water molecules in the brain; water molecules diffuse more rapidly in the direction of white-matter fibre tracts, allowing the tracts to be detected [20].

Therefore, neuroimaging methods are necessary for the creation of brain networks; without them, we would be unable to define the relationships between brain regions in living beings. Overall, brain networks provide a way to represent the data gained from neuroimaging studies in such a way as to illuminate relationships between brain regions that would otherwise be unknown. That is, they provide a simplified representation of the brain. Given the high complexity of the brain, such a simplified representation is very valuable. We see its value in the next section when we consider generative modelling.

---

[5]Time-averaged refers to averaging the time-series over a time-period of several minutes that the data was recorded for.

3 Generative Modelling

Generative models have been widely studied in network science [2], and have a history dating back a long time in network science and mathematics [11]. However, their application to brain networks spans just the last few decades [11]. In this section, we explain what generative modelling is, and discuss the history of generative models in the context of neuroscience. Much of the understanding we present comes from [2] and [11].

3.1 What is Generative Network Modelling?

A generative model of a network is exactly what the name suggests: a model which *generates* a network. That is, a set of rules defining how nodes and edges combine to create a network. Mathematically, this amounts to defining a parameterised rule which tells you how edges form between nodes. Many generative models consist of a fixed number of nodes, so only edges are generated by the mechanism. Other generative models, referred to as growth generative models, see nodes added successively to the network, causing the network to grow in terms of both nodes and edges. In most cases, the rule defining edges is a probabilistic one: the probability an edge forms between nodes v_i and v_j is given by a probability P_{ij}, specified by the model. This probability could be constant, or a function of some features of nodes v_i and v_j. Furthermore, if the network is embedded spatially (i.e. if nodes have specified positions in space), then the positions of nodes can either be pre-specified, or probabilistically generated by the model. Once specified, the generative model is computationally implemented, and a synthetic network produced.

3.2 How is Generative Modelling Used in Neuroscience?

The goal of generative modelling in neuroscience is to illuminate mechanisms that generate brain networks. That is, to elucidate mechanisms which lead to observed patterns of relations between brain regions. In turn, this allows us to understand the biological motivation behind mechanisms that lead to the brain's incredibly complex structure and function. This is achieved by producing synthetic networks (created by computationally implementing the generative *model*) with features that match empirical brain networks (created using neuroimaging *data*) [2]. Having an understanding of the mechanisms that generate brain networks would allow us to qualitatively understand the network's behaviour as a function of the mechanism's parameters, and potentially (in the future) control the network's development to exhibit features that are biologically desirable [11].

Remark 1. This demonstrates the necessity of computation as a tool for understanding mechanisms that generate brain networks: generative mechanisms can only be understood via a computationally implemented model, since the model itself does not tell us the topological features it creates. Furthermore, this highlights the interplay between data and computation required for generative modelling: the model and synthetic networks are useless without the representation of the data (brain network) it is trying to emulate, as without the brain network, one does not know which features the generative model is aiming to produce.

Generative modelling of brain networks can be summarised in the following steps, which we first state before proceeding to explain in more detail.

1. Determine biologically plausible rules for generating a network.

2. Define a parameterised generative model that reflects these rules.

3. Choose network metrics for comparing synthetic and empirical networks (preferably orthogonal and representative of necessary biological features of the empirical networks).

4. Define an objective function that optimises for the existence of these metrics, and conduct the optimisation procedure to determine model parameters.

5. Cross-validate the model.

6. Repeat for numerous generative models, and compare them by how well they each optimise the same objective function.

3.2.1   Defining the Generative Model

Generative rules are chosen to reflect biological factors that are thought to play a role in determining brain network structure and/or function. Once determined, these rules can be defined as a probability that an edge forms between nodes. For example, if the biological factor is that long-distance connections consume a lot of energy and should be penalised, the generative model could be $P_{ij} \propto (d_{ij})^{-\gamma}$, where d_{ij} is the anatomical (Euclidean) distance between nodes v_i and v_j and γ is the (positive) model parameter. This means that the probability that an edge forms decreases as a function of distance, reflecting the biological motivation.

If the generative mechanism successfully produces a network with similar features to a brain network (i.e. small-world architecture, modular structure and a heavy-tailed degree distribution), it can be concluded that the mechanism may reflect the biological rules leading to the organisation of the brain network. That is, we can learn about the biology from the computation. However, it is important to note the word 'may': a generative mechanism that produces features displayed in empirical brain networks does not imply that the said mechanism is *the* generative mechanism of the brain network, only that it is a *possible* mechanism [11]. Therefore, as noted in [11], one must be careful to only infer possibility of, rather than claim proof of, a certain mechanism.

Remark 2. It is important to note that this step is what allows us to learn about the biology of the brain from the generative model; since the model is defined using biological motivations, the synthetic networks produced can be directly related back to them. However, it also highlights a slight subtlety, in that one can only learn about the biology from the computation by having an understanding of the biology (or at least hypotheses about the biology) prior to defining the model.

3.2.2   Comparing Synthetic and Empirical Networks

Having determined biologically motivated generative rules and used them to define a generative model, the next step is to determine how to evaluate similarity between the synthetic and empirical networks.

The most obvious way to compare the similarity between two networks is to choose network features to compare. For example, one could compare how many edges they have in common, or whether their average path lengths are similar. However, the choice of which network features to use for comparing network similarity is still an ongoing question in the literature [2], because different features can elicit very different conclusions as to how similar two networks are, as the following example from [11] illustrates.

Example 1. Consider a ring lattice and a small-world network, shown in Figure 5 (a) and (b), respectively. (The small-world network is a ring lattice with a few additional connections, and hence much lower average path length — recall 2.2.1 above.) Both of these networks have nearly exactly the same edge set, so quantifying their similarity based on this feature would lead us to conclude that the networks are highly similar. However, if we were to instead quantify similarity by average path length L, we would conclude the opposite: the ring lattice has much higher average path length than the small-world network.
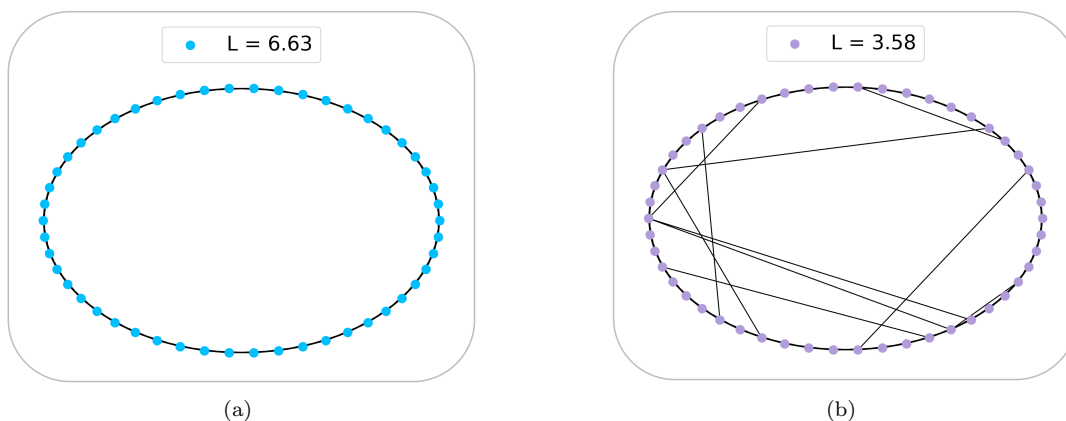

| (a) | (b) |

Figure 5: Figures to illustrate Example 1. (a) Ring lattice with $n = 50$ nodes and each node joined to $k = 4$ nearest neighbours. Shortest path length is $L = 6.63$. (b) Ten shortcuts added to the ring lattice. The shortest path length $L = 3.58$ is a lot smaller, however the edge set only differs by 10 edges.

This example demonstrates the issue of comparing goodness-of-fit using network metrics; we could miss metrics that better characterise the network, or use network metrics that lead to very incorrect conclusions for the context being considered.

Another issue with using network features to measure similarity is that many features are correlated with one another, leading to an aspect of redundancy. For example, comparing networks based on the existence of hubs and a heavy-tailed degree distribution would not serve much use over comparing them based on just one of these features; by definition, a heavy-tailed degree distribution indicates the presence of nodes with high degree, and therefore the existence of hubs.

Despite these caveats, comparing networks based on network features is the the go-to approach [11]. Therefore, care must be taken to choose features that are not correlated, and that are representative of any biological similarity that the synthetic networks are desired to possess. However, given that network neuroscience has highlighted the main set of network features characteristic of brain networks (see previous section), it is hoped that comparing networks using these features provides an accurate representation of network similarity.

3.2.3 Determining Model Parameters

Once comparison features are chosen, the next step is to determine which model parameters maximise the similarity between the synthetic networks and empirical brain networks. In the example above, this would mean determining which value of γ maximises similarity. Our report focuses on the generative models themselves, rather than parameter optimisation techniques (which would fill a whole report-worth on their own). Therefore, we refer the reader to [6, 11, 22] for a more detailed account of possible optimisation procedures, and give a brief description below.

Optimisation of parameters can be achieved by defining an objective function to minimise. The exact form of the objective function depends on the features chosen, but in general will be a function of statistics comparing the average of the features between synthetic and real brain networks (for example, t-statistics for t-tests between each feature for the synthetic and real networks, as is done in [6]). Since the objective function is a function of multiple parameters, optimisation techniques are computationally costly.

The objective function can serve a second purpose, by acting as a tool to compare different generative models [11]; the better optimised the objective function is, the more likely the model is to represent mechanisms that generated empirical networks. However, one should note that testing all possible generative mechanisms reflecting certain biological motivations is far from feasible, and therefore even the best-fitting model out of those compared may be outdone by another model not considered in the study [11].

3.2.4 Cross-validation

Cross-validation is a way of assessing how well a model generalises; in the context of generative models of brain networks, this means seeing whether the generative model and best-fitting model parameters are true representations of brain networks, rather than being specific to the dataset used to determine them. In particular, in generative modelling, cross-validation can be achieved in two ways [11]: (1) comparing synthetic networks with best-fitting parameters to empirical brain networks not used in the optimisation procedure, and (2) using the Erdos-Renyi random graph (see next section) to act as a null model, to ensure generated models perform above chance and do not overfit the data [11]. Furthermore, the generated networks can be compared to empirical networks that do not possess properties the synthetic network was desired to emulate, to see if the objective function is general enough [11].

3.3 How Can Generative Modelling Help us Understand Psychiatric Disorders?

In the same way that one can fit model parameters to empirical brain networks, model parameters can also be fit across patient groups (healthy vs psychiatric patients). Differences in these parameters can then suggest possible biological mechanisms that lead to abnormal connections. For example, suppose the generative mechanism connects two nodes v_i, v_j with probability $P_{i,j} \propto (d_{ij})^{-\gamma}$, as in our example above. If γ is much higher for an abnormal brain network than for a normal one, the distance penalty is much higher, and so this suggests there may be a higher-than-normal cost associated with long-range connections in the patient group.

Generative modelling could also be used to predict evolution of psychiatric disorders via simulation [11]: they could suggest possible mechanisms causing an individual to evolve into a disordered state, allowing a

possible way to 'forecast' the development of the illness [11]. An even more exciting possibility is the potential to use this understanding of the network's behaviour as a function of parameters to grow a network into a target configuration; this would have profound implications for the treatment of psychiatric disorders [11].

Overall, the aim is for network analysis of psychiatric disorders to involve both a description of the system-level diagnostic abnormalities, and an understanding of mechanisms that could have led to the generation of such abnormalities [6]. However, given that current understanding of generative mechanisms for normal brain networks is far from complete, this is a challenging task [6].

4 Generative Models

Having discussed what generative modelling is, and how it is done in practice in neuroscience research, we can now explore the generative models themselves. A key point to note is that in order for one to understand mechanisms leading to abnormal brain networks, one must first understand generative models of normal brain networks [4]; otherwise, a generative model that has good fit with brain networks of psychiatric patients does not tell us anything about how the mechanism differs between normal and abnormal brains, and therefore does not shed light on the disorder. Hence, in order to explore current literature relating to generative models of mental disorders, we must begin by exploring generative models of normal brain networks. Therefore, in this section we discuss generative models of normal brain networks, in preparation for the following section which addresses generative models of psychiatric disorders. Unless otherwise stated, this understanding comes from [11] and [17].

4.0.1 Erdos-Renyi Random Graph

One of the earliest generative models in network science was the Erdos-Renyi (ER) model [11]: a network consists of n nodes, and edges are formed between each pair of nodes v_i, v_j with probability $P_{ij} = p$. That is, the probability of each edge existing is equal. In Figure 6 we plot realisations of an ER random graph for a range of edge probabilities. Unfortunately, the ER model poorly approximates real-world networks; not only does the resultant synthetic network not exhibit any interesting structure, but the generative mechanism itself fails to emulate mechanisms by which networks develop, since there is no topological or spatial motivation behind the generative rule [11]. In particular, this means that the ER random graph fails to approximate brain networks: just from looking at Figure 6, is it obvious that no modular structure exists, for example.
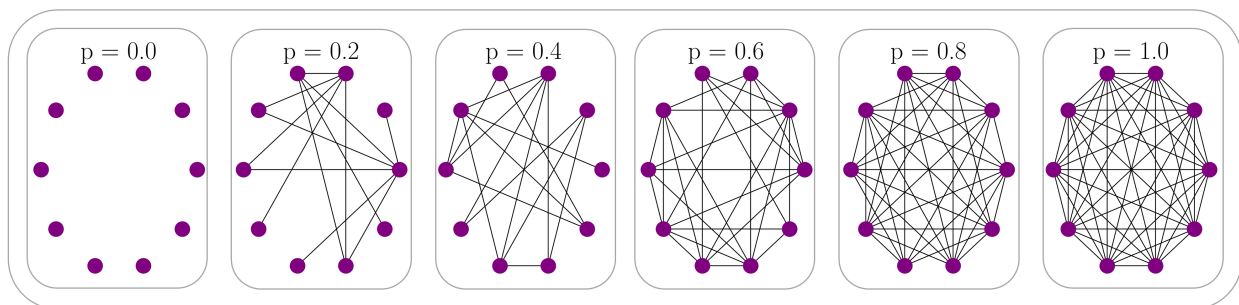

Figure 6: Realisations of an ER random graph with $n = 10$ nodes and edge probabilities $p \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$.

4.1 Topological Generative Models

This failure to emulate real-world networks led to attempts to develop mechanisms that captured topological properties of networks. In particular, mechanisms that captured small-world properties and heavy-tailed node degree distributions were sought. The Watts-Strogatz small-world model and Barabási-Albert preferential attachment models achieved just this, respectively.

4.1.1 Watts-Strogatz Small-World Model

The Watts-Strogatz small-world model generates synthetic networks with high clustering coefficient and short average path length (i.e. small-world properties) [11]. Similarly to an ER random graph, the generated network consists of a fixed number n of nodes. However, rather than the parameter p determining how likely an edge between two nodes is, p acts as a tuning parameter for rewiring of edges: the network begins in a lattice configuration, whereby each node is connected to nodes at most k lattice-spacings apart, and with probability p, each edge is 'rewired'. Rewiring means that one end of the edge is chosen at random and attached to another node in the network, chosen uniformly at random. In this way, p interpolates the model from a ring lattice with regular structure ($p = 0$) a graph with a fixed number edges positioned randomly ($p = 1$). In Figure 7, we plot realisations of a Watts-Strogatz model for $n = 10$ nodes and

increasing rewiring probability p, starting from the regular lattice, with $k = 4$. Figure 8 shows that for small rewiring probabilities, the average clustering coefficient and average path length are large, and for large rewiring probabilities, these two features are small. However, there exists a middle region with high clustering coefficient and small average path length; the two properties characterising small-world structure, and a key feature of brain networks. Therefore, the Watts-Strogatz model better approximates brain networks than the ER random graph. However, like the ER model, it lacks modular structure.
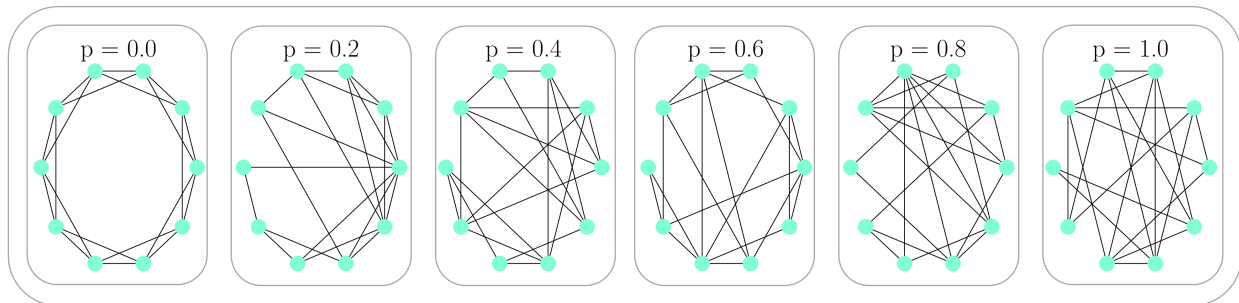

Figure 7: Realisations of a Watts-Strogatz model for $n = 10$ nodes and rewiring probabilities $p \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$, beginning with a regular $k = 4$ lattice.
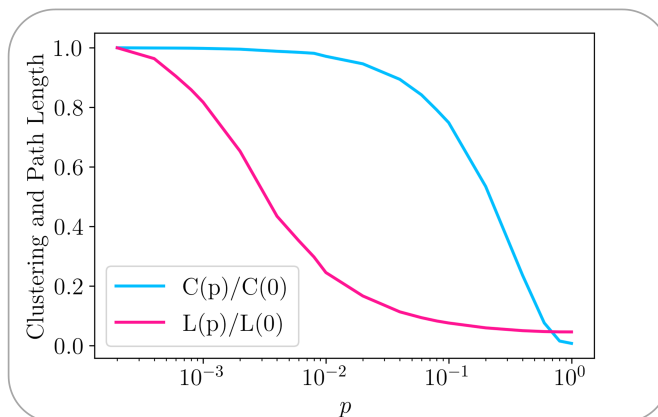

Figure 8: Normalised average clustering coefficient (C) and average shortest path length (L) for a Watts-Strogatz graph with $n = 1000$ nodes and each node initially connected to neighbours $k = 4$ lattice-spacing apart. These quantities are averaged over $N = 60$ realisations.

4.1.2 Barabási-Albert Preferential Attachment Model

The most-studied generative model in network science is the Barabási-Albert preferential attachment model (BA model) [2]. This is a growth generative model: new nodes are added to the network at each time-step. The network begins with m_0 nodes making up a complete graph (an edge exists between every pair of vertices), and at each time-step t a new node is added and forms connections with $m \leq m_0$ existing nodes, each chosen with probability $\Pi(i) \propto k_i$ (i.e. proportional to degree). In this way, new nodes connect 'preferentially' to existing nodes with high degree, hence the name of the model. This occurs until the network is made up of n nodes. We illustrate the BA model in a schematic in Figure 9. This process results in the formation of hubs. Moreover, the degree-distribution exhibits power-law behaviour (a particular heavy-tailed distribution of the form $P(k) \propto k^{-\gamma}$), which is commonly observed in brain networks [13]. However, this simple generative network model fails to display community structure [13], and therefore does not reproduce all three features of brain networks [2].
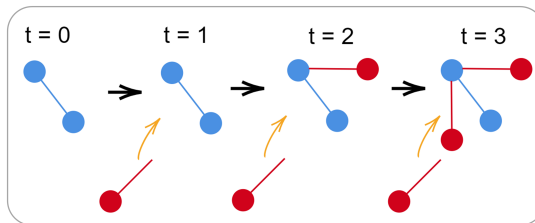
Figure 9: Schematic of the BA model with $m_0 = 2$ and $m = 1$ for three time-steps.

4.1.3 Overview

Both of these models only consider topological features, and don't involve spatial constraints [13]. Considering the brain, it seems obvious that any model failing to account for spatial embedding would fail to produce a network emulating key characteristics of the brain, since the brain has evolved in a spatially embedded environment (3-dimensional anatomical space) [2]. In particular, there are spatial constraints on growth, including the physical border of the skull, and apoptosis (cell death) [13]. In terms of functional networks, the brain is biologically expensive: the brain consumes roughly 20% of the body's metabolic energy, despite only making up 2% of the mass of the body [5]. Hence, functional brain networks are subject to spatial constraints too.

4.2 Spatial Generative Models

Given that generative models focusing on topological network features did not produce synthetic networks displaying all characteristics of brain networks, focus turned to considering generative rules that accounted for spatial features [2].

4.2.1 Biological Motivation

The main neurobiologically motivated spatial feature is a penalty for long-range connections, which reflects both functional and structural brain networks, and is a result of the brain being spatially-embedded [12]. For example, in structural networks, the concentration of growth factors that guide axon growth decays exponentially from its source [13, 23], and the material and energy costs of long-range white-matter fibres are therefore much higher than short-range connections [16]. In functional brain networks, long-distance connections are more expensive in terms of speed of signal transmission [6, 16].

4.2.2 Spatial Models

Waxman [24] introduced a generative model in which connections are established between two nodes as a decaying exponential function of distance [11, 13]. The original application for such a model was the establishment of connections between internet routers [24]. Therefore, the model was motivated by a penalty of long-range connections, much like brain networks, but instead reflected the high wiring and maintenance costs of long-range router links [13]. The model specifies the number n of nodes and the spatial positions of the nodes a $priori$, so this mechanism is not a growth mechanism, just a connection-establishment one. At each step, one edge is added, chosen to be between nodes v_i and v_j with probability $P_{ij} \propto \exp(-\beta d_{ij})$. Here, β is the model parameter to be fit, d_{ij} is the spatial distance between nodes v_i and v_j, and the normalisation factor is sum of connection probabilities over all pairs of nodes, i.e. $\sum_{i=1}^{n} \sum_{j=1}^{i-1} \exp(-\beta d_{ij})$. This model can be adapted to successively add new nodes and randomly generate their positions, with each new node being removed if no connections to it are established [11].

Since the collections of Waxman models penalise long-distance connections, they successfully create highly clustered networks; a feature the topological generative models above failed to emulate. This can be seen in Figure 10, where we generated a network with $n = 100$ nodes using the Waxman model with decay parameter $\beta = 5$. However, despite this collection of generative models having biological motivation for the wiring rules, they fail to account for many characteristic topological features of brain networks, such as rich-club structure [2, 25, 26]. This is because they penalise the long-distance connections that are present in brain networks; forming only short-range connections reduces average path length and as such reduces the efficiency of information flow between distant brain regions [6, 12].
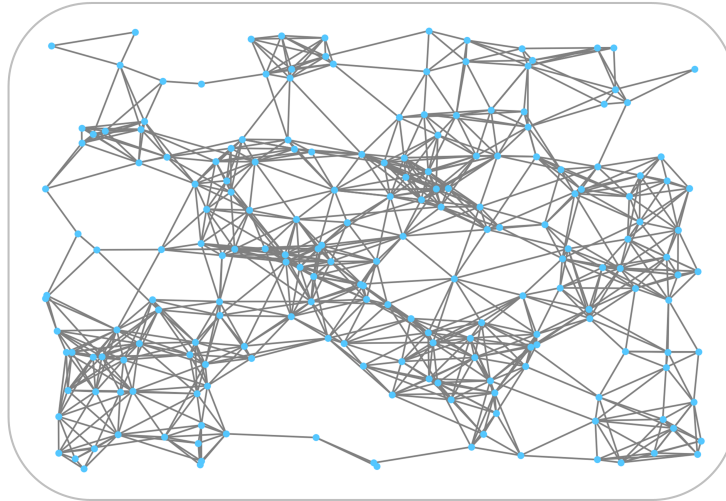
Figure 10: Network with $n = 200$ nodes (blue) positioned randomly, and $m = 3000$ edges (grey) generated using the Waxman model with decay parameter $\beta = 5$. A minimal spanning tree was used as a base to ensure the network is connected. The network is highly clustered.

4.3 Combining Spatial and Topological Rules

Since purely spatial and purely topological generative models either fail to generate clusters or fail to generate long-distance connections, and are therefore each incapable of emulating brain networks on their own, the above work led to the idea to combine topological and spatial rules in a generative model. The idea behind this was that each process would work together to create features that the other failed to reproduce on its own. This is the current state-of-the-art in generative modelling of brain networks, and applies to both functional and structural brain networks [11].

4.3.1 Biological Motivation

The biological motivation for generative models of brain networks that combine topological and spatial rules is a trade-off between the biological cost of long-range connections (described above) and the need for network features to maximise functionality (i.e. enable both integration and segregation of function with small-world architecture and rich-club structure) [12, 27, 28, 29]. It is a trade-off, because network features are preferred even if they are more costly [2].

4.3.2 Main Results

The main papers in the literature exploring this trade-off through two-parameter generative models of brain networks are 'Simple models of human brain functional networks' by Vértes et al. (which explores functional brain networks) and 'Generative models of the human connectome' by Betzel et al. (which explores structural brain networks). Both papers tested over ten different generative models that penalised long-range connections and maximised certain topological features. Vértes et al. compared synthetic networks to empirical functional networks with clustering coefficient, modularity (a measure of how modular the network structure is), degree distribution, and efficiency. In contrast, structural brain networks and synthetic networks were compared by Betzel et al. by using the degree distribution, clustering coefficient, total edge length (Euclidean distance between nodes linked by an edge) and betweenness centrality.[6] These all reflect the characteristic features of brain networks: small-world architecture, modular structure, heavy-tailed degree distribution (as a result of hubs) and high efficiency (due to rich-club structure).

Interestingly, both papers found the best-fitting model to add edges with probabilities that minimise the length of connections and maximises the overlap in two nodes' neighbours (with the only difference being that Betzel et al. normalised this measure by the total number of neighbours of the two nodes). This led to

---

[6]The betweenness centrality of a node measures the number shortest-paths a node lies in, and the average is taken over the all nodes in the network [17].

a model generating edges with probabilities

$$P_{ij} = \frac{(k_{ij})^{\gamma}(d_{ij})^{-\eta}}{\sum_{i,j}(k_{ij})^{\gamma}(d_{ij})^{-\eta}}, \tag{5}$$

where d_{ij} is the anatomical Euclidean distance between the two nodes, and k_{ij} measures the overlap in the neighbourhoods of v_i and v_j (i.e. the number of neighbours in common between nodes v_i and v_j for the functional networks, and this quantity normalised for structural networks). The normalising sum is over all pairs of nodes. Therefore, the probability that an edge forms between two nodes increases as the number of neighbours that two nodes have in common increases, and decreases the further apart the two nodes are in anatomical space. This model was called the 'economical clustering rule' in [6], and so we shall henceforth refer to (5) as such.

Biologically, for functional networks this topological rule reflects a bias of brain regions to cluster together if they share a large proportion of the same input, which is reminiscent of Hebbian learning rules [12]. For structural networks, the topological rule represents a homophilic attachment mechanism, whereby nodes of similar type connect to one another [11].

4.3.3 Functional vs Structural

As noted in [11], the fact that effectively the same generative model was the best-fitting model in both papers considered above may suggest the physical wiring and functional topology of the brain are subject to similar principles of organisation. However, this may also simply be a coincidence: both mechanisms model the same scale of brain networks, the biological motivations for generating both networks are similar (cost-efficiency and topological value), and both types of network exhibit the same network characteristics. Therefore, it seems unsurprising that the same generative models were best-fitting for both: if a generative model is optimised to emulate characteristics that both functional and structural networks possess, motivated by the same factors, one should surely expect the same model to work for both.

However, the growth mechanisms for structural networks do not translate well to functional networks: functional connectivity emerges from dynamical brain processes that are constrained by anatomical structure, anatomical distance, genetics and task-dependent activity [11]. This means that discrete and independent edge-placement in mechanistic growth models is missing biological plausibility for functional brain networks, and is far more amenable to structural brain networks [11]. That is, functional networks are not generated through an edge-addition process like structural networks are [11]. Therefore, it may be simply that since both the functional and structural networks exhibit the same network characteristics that are motivated by the same biological constraints that the mechanism that generates such properties is suited to both, rather than the mechanism itself being biologically reflected in the functional brain networks. That said, generative models of functional networks are still useful, as they can elucidate key biological motivations behind the emergence of network characteristics, even if the generative model does not reflect true generation of functional networks.

Remark 3. The relationship between structural and functional networks at the system level is not fully understood and is currently an important open question in neuroscience [6, 13, 16, 30]. Although the functional network is constrained by the structural one [5, 11, 16], the two are not directly correlated. Because of this unknown relationship, there are some caveats and limitations associated to generative models of functional brain networks (described above) [31, 32, 33], and so interpreting functional connectivity in terms of growth models needs to be done with caution [13].

4.4 Overview

We have now explored the main results concerning generative models of normal brain networks. We learnt that the best-fitting models known to date are motivated by a trade-off between the biological cost of long-range connections and the benefit of metabolically-costly topological features. This puts us in a position to explore generative models of psychiatric disorders, which we do in the next section. There, we go into more detail about how neuroimaging data is used.

5 Generative Models of Mental Disorders

The literature on generative models of mental disorders is somewhat limited, and consists mainly of review papers exploring its prospects (e.g. [2, 11]). As discussed earlier, this is a result of the requirement of prerequisite knowledge of generative models of normal brain networks, which is itself an ongoing area of research. However, there has been influential work done in the context of schizophrenia; a particularly interesting disorder to study in the context of system-level brain network generative models, since it is recognised as a disorder of dysconnected brain networks and associated with dysconnectivity at both the structural and functional level [2, 4, 16, 34, 35]. In particular, studies suggest patients suffering from childhood-onset schizophrenia (COS) have reduced modular segregation and reduced clustering [4, 36, 37, 38].

The main study into generative models of mental disorders is the paper by Vértes et al. that we considered in the previous section. They studied generative models in the context of COS. By performing exactly the same optimisation procedure with generative models on a set of 19 abnormal functional brain networks as was done on 20 healthy brain networks, they found that the best-fitting model was again that given by (5) above. This supports the notion that functional brain networks evolve to minimise wiring cost and maximise overlap in nodes' neighbourhoods. However, the best-fitting parameters were significantly different: $\eta = 2.3, \gamma = 3.33$ for volunteers with COS compared to $\eta = 2.63, \gamma = 3.17$ for healthy volunteers. The lower value of η in the volunteers with COS reflects a smaller distance penalty (i.e. lower biological cost) in connections, and therefore results in reduced modularity and clustering. The higher value of γ reflects greater propensity for connections resulting in nodes having many nearest neighbours in common. Therefore, this suggests that the abnormal connections in brain networks in schizophrenia can be explained by an 'abnormally biased trade-off' [6] between topological clustering and distance penalisation.

We now dive into the data from this paper and use our own code to produce synthetic networks using this generative model.

5.1 Simulations

The data consisted of fMRI time-series correlations between 294 regions of brain (defined by centroids[7] in anatomical space) for 20 healthy volunteers and 19 volunteers diagnosed with COS. We were also given the Euclidean distance[8] between each pair of nodes. Similarly to Vértes et al., we only consider a parcellation of the right hemisphere into 140 regions, since this allows us to draw comparisons with the paper. Furthermore, we follow the methods outlined in the paper to construct the brain networks.

Using this data, we generated the functional brain networks by adding edges between nodes if the corresponding time-series were highly correlated. In particular, denoting the correlation between nodes v_i and v_j by ρ_{ij}, each possible connection was weighted by 1 - ρ_{ij} (to give highly correlated connections with $\rho_{ij} \approx 1$ low weights, and highly uncorrelated connections with $\rho_{ij} \approx 0$ or $\rho_{ij} < 0$ high weights). Using these weights, a minimal spanning tree (a minimum-weight network connecting all the nodes) was constructed, and further edges added from lowest weight upwards until the connection density reached 4%. That is, until 4% of all possible edges existed (in this case $0.04 \times 140(140 - 1)/2 = 389$ edges). Although other methods for constructing functional brain networks exist, this technique ensures the network is fully connected (by creating a base of a minimal spanning tree), and avoids having to define a threshold constituting how correlated two nodes must be to warrant the formation of a connection [6, 37].

Synthetic networks were constructed as follows: a minimal spanning tree based on anatomical distance was constructed as a base, to ensure the generated network was connected. Then at each time-step, an edge was added, chosen to be between nodes v_i and v_j with probability P_{ij} (defined by the specific model, e.g. equation 5), until the network reached 4% of the total possible number of connections (to be consistent with the empirical brain networks).

Remark 4. A limitation with our simulations to test the goodness-of-fit with the parameters specified in the paper is that the data we are testing it on is the same data that was used to fit the model. To avoid this, we would need additional fMRI data not available to us.

---

[7]The i-th centroid is specified by 3-dimensional coordinates $\mathbf{x_i} = (x_i, y_i, z_i)$.

[8]The Euclidean distance between nodes v_i and v_j is $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$.

5.1.1 Clustering and Modularity

Before simulating the generative models, we first confirmed that the brain networks of COS patients did in fact have reduced modular segregation and clustering, as suggested by the literature. Figure 11 displays the modularity (a measure of how modular the network is) and clustering (measured by the average clustering coefficient) of the functional brain networks averaged over the 20 healthy volunteers and 19 COS patients. It is clear that the COS patients have reduced modularity and clustering, as expected.
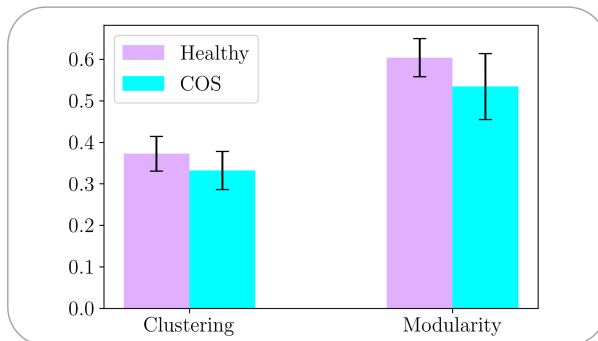

Figure 11: Clustering coefficient and modularity averaged over the 20 healthy volunteers' brain networks and 19 volunteers with COS. Both clustering and modularity are reduced in COS patients.

5.1.2 Generative Models

We then generated synthetic networks using both the economical clustering model (given by (5) above) and the exponential decay model, in order to compare good- and ill-fitting models. Since parameter optimisation (in the case of this paper, using simulated annealing) is highly computationally demanding, rather than performing the optimisation procedure, we simply use the best-fitting parameters from the paper in our algorithms. In particular, for the exponential model $P_{ij} \propto \exp(-\eta d_{ij})$ we use $\eta = 0.16$, and for the economical clustering model $P_{ij} \propto (k_{ij})^\gamma (d_{ij})^{-\eta}$ we use $\eta = 2.63, \gamma = 3.17$ for healthy individuals and $\eta = 2.3, \gamma = 3.33$ for individuals with COS.

Figure 12 gives a visual representation of the synthetic networks against the data for one healthy individual; visually, it is clear that the economical clustering model provides a good fit to the data (because the networks look similar) and the exponential decay does not, since the exponential decay model lacks modular structure. We show the same plots for an individual with COS in Figure 13; the reduced modular structure is apparent in the empirical network (a) due to an increased number of long-range connections compared to Figure 12 (a), and is reflected well in the economical clustering model (c) too. Overall, the reduced distance penalty means there are a greater number of long-range connections than in Figure 12 (c). To see this more quantitatively, in Figure 14 we plot the degree distributions of the two generative models (averaged over 20 realisations) against the data (averaged over all 19 individuals' brain networks); the economical clustering rule provides a very good fit to the degree distribution of the data.

5.2 Overview

Therefore, in this section we have considered generative models of mental disorders by focusing on the paper by Vértes et al. We produced synthetic networks using their best-fitting parameters for two generative models, and saw that the economical clustering rule fits the data quite well, whereas the exponential decay model fails to produce the necessary clusters. Furthermore, we saw that individuals with COS have reduced modular segregation, and that varying the model parameters from those of healthy individuals to reflect this provided a good fit to the data. However, as evident from the plots, the economical clustering model does not provide a perfect fit. This fit could possibly be improved by considering other forms of generative models, which we discuss under 'future directions' in the next section, alongside a discussion of the limitations of generative modelling.
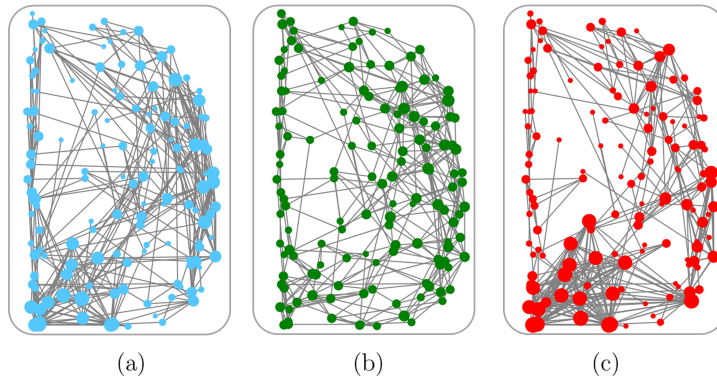
Figure 12: Right hemisphere brain network of a healthy individual. (a) Functional brain network generated from fMRI data from [6], as described above. (b) Synthetic network generated using exponential decay rule $P_{ij} \propto \exp(-0.16d_{ij})$. (c) Synthetic network generated using economical clustering rule $P_{ij} \propto (k_{ij})^{3.17}(d_{ij})^{-2.63}$. Parameters chosen to match the best-fitting parameters in [6].
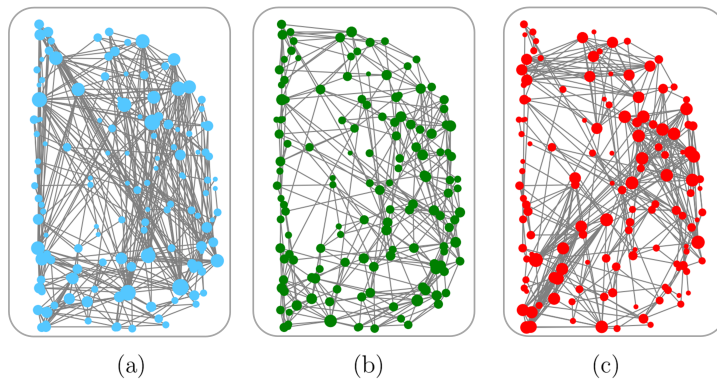

Figure 13: Right hemisphere brain network of an individual with COS. (a) Functional brain network generated from fMRI data from [6], as described above. (b) Synthetic network generated using exponential decay rule $P_{ij} \propto \exp(-0.16d_{ij})$. (c) Synthetic network generated using economical clustering rule $P_{ij} \propto (k_{ij})^{3.33}(d_{ij})^{-2.3}$. Parameters chosen to match the best-fitting parameters in [6]. The reduced modular structure is apparent in (a), and the economical clustering rule is a better fit than the exponential decay rule.
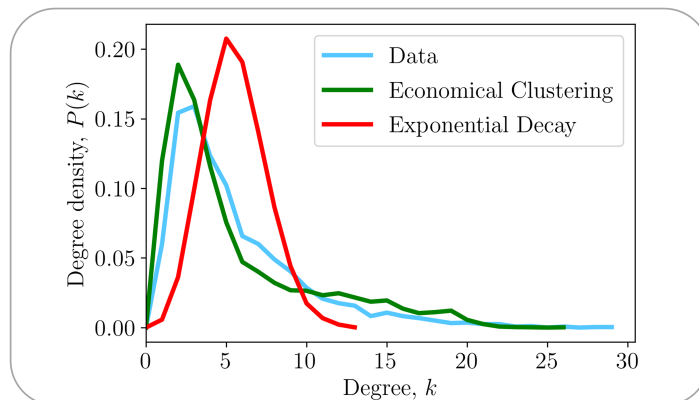

Figure 14: Degree distributions of brain networks of the right hemisphere of individuals with COS. Blue line is the average degree distribution of fMRI brain networks for 19 individuals with COS. The exponential decay rule $P_{ij} \propto \exp(-\eta d_{ij})$ has parameter $\eta = 0.16$, and the economical clustering rule $P_{ij} \propto (k_{ij})^{\gamma}(d_{ij})^{-\eta}$ has parameters $\gamma = 3.33$ and $\eta = 2.3$. The synthetic networks are averaged over 20 realisations. The economical clustering rule has a very good fit to the data.

6 Discussion

In this report, we explored generative models of mental illness. This involved first understanding normal brain networks in terms of the topological features that characterise them, before looking at generative models which produce synthetic networks that emulate these features well. With this knowledge, we then looked at generative models of childhood-onset schizophrenia (COS), focusing on a paper by Vértes et al. [6]. From the structure of this report, it is clear that a great deal of work remains to be done in the context of generative network models of mental illness; a prerequisite for such studies is an adequate understanding of normative generative models, which are themselves not completely understood. We now conclude the report with a discussion of the limitations of generative models, and a suggestion of possible future directions.

6.1 Methodological Limitations

Despite a lot of progress having been made with regards to generative models of brain networks, there still exist fundamental methodological limitations with such models.

6.1.1 Neuroimaging Limitations

Firstly, the empirical brain networks used to fit generative models are often created using noisy data [3]: fMRI data is heavily affected by small (<1mm) transient head movements [4, 5, 39], which can cause an increase in the observed connectivity across the whole network [4]. There do exist pre-processing methods[9] to minimise the impact of movement (e.g. global signal regression[10] [40]), but these methods themselves have limitations, such as biasing short-range connectivity. Therefore, the accuracy of future generative modelling is reliant on improved pre-processing methodologies for neuroimaging data in order to limit the effect of artefacts.

6.1.2 Parcellation Schemes

As explained in [12], the choice of the parcellation scheme used to define nodes could also considerably affect the fit of generative models. In particular, this could impact the determination of the best-fitting model, and in turn influence the interpretation of the biological motivation behind the the model. Modelling networks with more and more nodes greatly increases the computational complexity of the generative models (as the network grows, the number of positions for an edge to be placed grows like $O(n^2)$ [12]). Moreover, larger networks could require bigger parameter values, or even additional parameter constraints in wiring rules [12]. Therefore, future work is needed to understand the effect of different parcellations on best-fitting generative models.

6.1.3 Diagnosis of Mental Disorders

A major issue with generative models of mental disorders is the actual classification and diagnosis of the disorders themselves. The majority of neuroimaging studies employ diagnostic boundaries reflective of the Diagonstic and Statistical Manual of Mental Disorders (DSM-5) [41], which diagnoses disorders by identifying combinations of symptoms [2]. Despite being useful for patients themselves, as it gives them a 'label' for their disorder, this diagnosis procedure suffers from significant limitations when used in the scientific community: individuals with exactly the same diagnosis (name of disorder) can exhibit completely different sets of symptoms, suggesting the biological mechanisms resulting in the disorder are also completely different [2]. This therefore poses extreme challenges in elucidating causal mechanisms of mental disorders. In particular, it is possible that the paper by Vértes et al. [6] that we studied suffers from this limitation; if all COS volunteers experienced the same set of symptoms, the best-fitting model parameters could be different and would potentially even provide a better fit to the data. Even worse, the economical clustering model might not be the best-fitting model. Hence, a shift in theoretical frameworks for classifying psychiatric disorders may be necessary for the full benefit of network generative models to become apparent.

---

[9]Methods to process the fMRI data before brain networks are constructed from it.

[10]The average intensity of the MR signals over all brain regions during the course of recording data (the 'global signal') is subtracted from the time-series of each brain region by using linear regression before correlations are calculated [40].

6.2 Future directions

Given that the exploration of generative models of mental disorders is still in its infancy, there exist a multitude of future directions for their development.

6.2.1 Multiple-Parameter Models

As we saw above, the current best generative models consider wiring rules based only on two parameters. Although this reflects the two biological motivations behind the models, it may be the case that other biological factors play a role too. This is suggested by the fact that even networks generated using the best-fitting models do not exactly fit empirical brain networks. Therefore, future work should aim to postulate further biological factors that shape the generation of brain networks, and use them to define additional wiring rules. Having said this, since it is not possible to test an exhaustive list of two-parameter models, it may be the case that a generative model defined by just two parameters produces synthetic networks with an even better fit than the ones discussed above [2].

6.2.2 Longitudinal Studies

One issue not considered in this report is that real brain networks are not fixed; they evolve both over task-dependent activity and over an entire lifetime [5]. Therefore, it would be beneficial for generative models to account for this continuous evolution of brain networks by using longitudinal data in studies. This would allow model parameters to be understood as a function of the evolution of the brain networks. However, obtaining such data is impractical, and very few studies have yet managed to obtain such longitudinal data [11]. As suggested by Bassett et al. [11], one way to overcome this could be to use other species with shorter life-spans as model-organisms for constructing brain networks, from which longitudinal generative models could be created.

6.2.3 Non-Euclidean Distance

So far, generative models of brain networks have measured distance between brain regions as the Euclidean distance in anatomical space. However, as pointed out by Betzel et al. [12], other measures of distance based on neurobiology may make more sense and reveal more interesting results. For example, distance could be a measure of the fibre length linking two brain regions (i.e. the length of white-matter fibre tracts, rather than the straight-line Euclidean distance between centroids) [12]. This would be of interest, because although Euclidean distance and fibre length are correlated, some fibre lengths are far longer than the Euclidean distance [12]. It has been suggested that models using this measure of distance may have improved fit for other generative mechanisms [12]. Having said this, despite DTI being the state of the art technology for identifying the structural connections in the brain, it does lead to many false positive and false negatives [42], which would significantly affect any results concerning generative models [12].

6.2.4 Higher-order Models

Generative models of brain networks that currently exist only consider dyadic interactions between brain regions. However, networks which account for higher-order interactions may be of merit, since beyond-pairwise interactions are posited to play a non-trivial role in brain dynamics [2, 4, 11]. One way to achieve this could be to model the brain as an hypergraph, which is a generalisation of a network whereby edges can contain any number of nodes (compared to a normal network in which each edge contains only two nodes). The study of generative models of hypergraphs is gaining fruition [43, 44, 45], and hypergraphs have recently been applied to describe brain networks [46]. Therefore, we think it would be interesting to consider such hypergraph generative mechanisms in the context of generative models of brain networks.

6.3 Addressing the NIMH goal

The motivation behind exploring generative models is to address the NIMH goal of defining brain mechanisms underlying mental health disorders. As demonstrated by this report, progress towards such a goal has so far been limited. We now suggest two ways in which generative modelling could be used to make further progress towards this goal.

6.3.1 Other Scales

Generative models of brain networks are helpful in elucidating key mechanisms driving the growth of brain networks, but they only provide insight at the system-level of the brain [2]. The complexity of the brain means that alterations in brain networks of individuals with psychiatric disorders occur at multiple spatial and temporal scales, suggesting that mechanistic models at other scales may be helpful too [2]. This may be of particular use, since many current therapeutic interventions, such as drugs, target cellular and molecular levels of the brain. However, understanding the system-level impact of such interventions is currently infeasible in humans [2].

6.3.2 Other Mental Health Disorders

Research into mental health disorders in the field of network neuroscience has seen a large focus on schizophrenia. However, there exist many disorders for which generative modelling could be of merit, but have so far received very little attention. One disorder we believe would particularly benefit from such exploration, but has eluded the network neuroscience community so far, is Anorexia Nervosa (AN); an eating disorder characterised by severe food restriction and low body weight [47]. As the psychiatric disorder with the highest mortality rate [48], a better understanding of the biological mechanisms behind the disorder is crucial. Given that there are clear energy deficits associated to the disorder, this would likely affect parameters of generative models that represent energy costs. Therefore, the disorder would be of particular interest to study in the context of generative models, to see how brain networks vary as a function of biological cost parameters over the course of the illness. However, this would require longitudinal data which, as discussed above, is hard to obtain.

6.4 Clinical Interventions

Although understanding mechanisms behind mental disorders is interesting, the end goal is for this understanding to translate to utility in clinical practice. Unfortunately, such translation has so far been almost non-existent, since determining mechanistic rules governing (patho-)physiological network growth does not solve the question of how to clinically intervene [2, 4].

There do exist therapies being developed that target system-level dysfunction explicitly, such as deep brain stimulation (DBS) and neurofeedback, that could benefit from insight from generative models [2]. These therapies involve electrically stimulating the brain to induce functional network changes. However, there is a lack of understanding of how electrical perturbations to the brain affect dynamics [49], so such treatments are not well-understood. One potential solution to increase understanding of these treatments, as suggested in [4], is to use network control theory (NCT). NCT explains how alterations to a network affects the global network dynamics; it supports the motivation behind using DBS, but it has not been used to explain how neuromodulation impacts brain networks [4].

Generative modelling could also be used to explore the effect of treatment interventions *in silico*[11] [11], by testing how the topological characteristics of a synthetic network evolve over time when the generative model used is varied [11]. By perturbing the model parameters or wiring rules, one could affect the trajectory of the network, and attempt to prevent it from developing topological features that are biologically maladaptive [11]. As stated in [11], this goal is in line with NCT. However, knowing which treatment interventions could cause such network perturbations is another hurdle in itself.

6.5 Conclusion

Overall, generative models of mental disorders remain incompletely understood [12]. However, the little progress made to date is not a reflection of the utility of the insights that can be gained generative models. Indeed, without such insights, we are unable to understand mechanisms leading to abnormal large-scale connections in brain networks of individuals with psychiatric disorders. Furthermore, without insight from generative models we are very far away from a point where an understanding of mechanisms behind mental disorders could translate into clinical practice. Therefore, we look forward to future work shedding light on the system-level mechanisms behind mental health disorders, and further progress towards the NIMH goal being made.

---

[11]By computer simulation.

References

[1] National Institute of Mental Health (NIMH). Strategic planning reports, 2020.

[2] Urs Braun, Axel Schaefer, Richard F Betzel, Heike Tost, Andreas Meyer-Lindenberg, and Danielle S Bassett. From maps to multi-dimensional network mechanisms of mental disorders. *Neuron*, 97(1):14–31, 2018.

[3] Alex Fornito and Andrew Zalesky. Computational approaches to understanding mental dysfunction: Progress, challenges, and new frontiers. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(9):728–730, 2018.

[4] Danielle S Bassett, Cedric Huchuan Xia, and Theodore D Satterthwaite. Understanding the emergence of neuropsychiatric disorders with network neuroscience. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(9):742–753, 2018.

[5] Petra E Vértes and Edward T Bullmore. Annual research review: growth connectomics–the organization and reorganization of brain networks during normal and abnormal development. *Journal of Child Psychology and Psychiatry*, 56(3):299–320, 2015.

[6] Ed Bullmore and Olaf Sporns. The economy of brain network organization. *Nature Reviews Neuroscience*, 13(5):336–349, 2012.

[7] Frank Pillmann. Carl wernicke and the neurobiological paradigm in psychiatry. *Acta Neuropsychologica*, 5:246–260, 01 2007.

[8] Danielle S Bassett and Olaf Sporns. Network neuroscience. *Nature neuroscience*, 20(3):353–364, 2017.

[9] David M Lydon-Staley and Danielle S Bassett. Network neuroscience: a framework for developing biomarkers in psychiatry. *Biomarkers in Psychiatry*, pages 79–109, 2018.

[10] Alex Fornito and Edward T Bullmore. Connectomic intermediate phenotypes for psychiatric disorders. *Frontiers in psychiatry*, 3:32, 2012.

[11] Richard F Betzel and Danielle S Bassett. Generative models for network neuroscience: prospects and promise. *Journal of The Royal Society Interface*, 14(136):20170623, 2017.

[12] Richard F Betzel, Andrea Avena-Koenigsberger, Joaquín Goñi, Ye He, Marcel A De Reus, Alessandra Griffa, Petra E Vértes, Bratislav Mišic, Jean-Philippe Thiran, Patric Hagmann, et al. Generative models of the human connectome. *Neuroimage*, 124:1054–1064, 2016.

[13] Claus C Hilgetag and Marcus Kaiser. Organization and function of complex cortical networks. In *Lectures in Supercomputational Neurosciences*, pages 107–133. Springer, 2007.

[14] Karl J Friston. Functional and effective connectivity: a review. *Brain connectivity*, 1(1):13–36, 2011.

[15] *The Structure and Dynamics of Networks:*. Princeton University Press, 2006.

[16] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews neuroscience*, 10(3):186–198, 2009.

[17] Renaud Lambiotte and Naoki Masuda. *A guide to temporal networks*, volume 4. World Scientific, 2016.

[18] Bryan Ek, Caitlin VerSchneider, and Darren A Narayan. Global efficiency of graphs. *AKCE International Journal of Graphs and Combinatorics*, 12(1):1–13, 2015.

[19] Gary H Glover. Overview of functional magnetic resonance imaging. *Neurosurgery Clinics*, 22(2):133–139, 2011.

[20] Denis Le Bihan, Jean-François Mangin, Cyril Poupon, Chris A Clark, Sabina Pappata, Nicolas Molko, and Hughes Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 13(4):534–546, 2001.

[21] Dae-Jin Kim and Byoung-Kyong Min. Rich-club in the brain's macrostructure: Insights from graph theoretical analysis. *Computational and Structural Biotechnology Journal*, 2020.

[22] Peter JM Van Laarhoven and Emile HL Aarts. Simulated annealing. In *Simulated annealing: Theory and applications*, pages 7–15. Springer, 1987.

[23] Marcus Kaiser and Sreedevi Varier. Evolution and development of brain networks: from caenorhabditis elegans to homo sapiens. *Network: Computation in Neural Systems*, 22(1-4):143–147, 2011.

[24] Bernard M Waxman. Routing of multipoint connections. *IEEE journal on selected areas in communications*, 6(9):1617–1622, 1988.

[25] Marcus Kaiser and Claus C Hilgetag. Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Comput Biol*, 2(7):e95, 2006.

[26] L da F Costa, Francisco A Rodrigues, Gonzalo Travieso, and Paulino Ribeiro Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in physics*, 56(1):167–242, 2007.

[27] Giulio Tononi, Olaf Sporns, and Gerald M Edelman. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, 91(11):5033–5037, 1994.

[28] Yuhan Chen, Shengjun Wang, Claus C Hilgetag, and Changsong Zhou. Trade-off between multiple constraints enables simultaneous formation of modules and hubs in neural systems. *PLoS Comput Biol*, 9(3):e1002937, 2013.

[29] Yuhan Chen, Shengjun Wang, Claus C Hilgetag, and Changsong Zhou. Trade-off between multiple constraints enables simultaneous formation of modules and hubs in neural systems. *PLoS Comput Biol*, 9(3):e1002937, 2013.

[30] Yi-Ping Lo, Reuben O'Dea, Jonathan J Crofts, Cheol E Han, and Marcus Kaiser. A geometric network model of intrinsic grey-matter connectivity of the human brain. *Scientific reports*, 5(1):1–14, 2015.

[31] Petra E Vértes, Aaron F Alexander-Bloch, Nitin Gogtay, Jay N Giedd, Judith L Rapoport, and Edward T Bullmore. Simple models of human brain functional networks. *Proceedings of the National Academy of Sciences*, 109(15):5868–5873, 2012.

[32] Catalina Obando and Fabrizio de Vico Fallani. A statistical model for brain networks inferred from large-scale electrophysiological signals. *Journal of The Royal Society Interface*, 14(128):20160940, 2017.

[33] Timothy O Laumann, Evan M Gordon, Babatunde Adeyemo, Abraham Z Snyder, Sung Jun Joo, Mei-Yen Chen, Adrian W Gilmore, Kathleen B McDermott, Steven M Nelson, Nico UF Dosenbach, et al. Functional system and areal organization of a highly sampled individual human brain. *Neuron*, 87(3):657–670, 2015.

[34] Alex Fornito, Andrew Zalesky, Christos Pantelis, and Edward T Bullmore. Schizophrenia, neuroimaging and connectomics. *Neuroimage*, 62(4):2296–2314, 2012.

[35] Klaas E Stephan, Karl J Friston, and Chris D Frith. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophrenia bulletin*, 35(3):509–527, 2009.

[36] Aaron Alexander-Bloch, Renaud Lambiotte, Ben Roberts, Jay Giedd, Nitin Gogtay, and Ed Bullmore. The discovery of population differences in network community structure: new methods and applications to brain functional networks in schizophrenia. *Neuroimage*, 59(4):3889–3900, 2012.

[37] Aaron F Alexander-Bloch, Nitin Gogtay, David Meunier, Rasmus Birn, Liv Clasen, Francois Lalonde, Rhoshel Lenroot, Jay Giedd, and Edward T Bullmore. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Frontiers in systems neuroscience*, 4:147, 2010.

[38] Justin T Baker, Avram J Holmes, Grace A Masters, BT Thomas Yeo, Fenna Krienen, Randy L Buckner, and Dost Öngür. Disruption of cortical association networks in schizophrenia and psychotic bipolar disorder. *JAMA psychiatry*, 71(2):109–118, 2014.

[39] František Váša, Rafael Romero-Garcia, Manfred G Kitzbichler, Jakob Seidlitz, Kirstie J Whitaker, Matilde M Vaghi, Prantik Kundu, Ameera X Patel, Peter Fonagy, Raymond J Dolan, et al. Conservative and disruptive modes of adolescent change in human brain functional connectivity. *Proceedings of the National Academy of Sciences*, 117(6):3248–3253, 2020.

[40] Kevin Murphy, Rasmus M Birn, Daniel A Handwerker, Tyler B Jones, and Peter A Bandettini. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage*, 44(3):893–905, 2009.

[41] Fifth Edition et al. Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc*, 21, 2013.

[42] Cibu Thomas, Q Ye Frank, M Okan Irfanoglu, Pooja Modi, Kadharbatcha S Saleem, David A Leopold, and Carlo Pierpaoli. Anatomical accuracy of brain connections derived from diffusion mri tractography is inherently limited. *Proceedings of the National Academy of Sciences*, 111(46):16574–16579, 2014.

[43] Jian-Wei Wang, Li-Li Rong, Qiu-Hong Deng, and Ji-Yong Zhang. Evolving hypernetwork model. *The European Physical Journal B*, 77(4):493–498, 2010.

[44] Dajie Liu, Norbert Blenn, and Piet Van Mieghem. A social network model exhibiting tunable overlapping community structure. *Procedia Computer Science*, 9:1400–1409, 2012.

[45] Feng Hu, Jin-Li Guo, Fa-Xu Li, and Hai-Xing Zhao. Hypernetwork models based on random hypergraphs. *International Journal of Modern Physics C*, 30(08):1950052, 2019.

[46] Shi Gu, Muzhi Yang, John D Medaglia, Ruben C Gur, Raquel E Gur, Theodore D Satterthwaite, and Danielle S Bassett. Functional hypergraph uncovers novel covariant structures over neurodevelopment. *Human brain mapping*, 38(8):3823–3835, 2017.

[47] Stephan Zipfel, Katrin E Giel, Cynthia M Bulik, Phillipa Hay, and Ulrike Schmidt. Anorexia nervosa: aetiology, assessment, and treatment. *The Lancet Psychiatry*, 2(12):1099–1111, 2015.

[48] Patrick F Sullivan et al. Mortality in anorexia nervosa. *American Journal of Psychiatry*, 152(7):1073–1074, 1995.

[49] Gustavo Deco, Joana Cabral, Victor M Saenger, Melanie Boly, Enzo Tagliazucchi, Helmut Laufs, Eus Van Someren, Beatrice Jobst, Angus Stevner, and Morten L Kringelbach. Perturbation of whole-brain dynamics in silico reveals mechanistic differences between brain states. *Neuroimage*, 169:46–56, 2018.

A   Appendix

A.1   Acknowledgements

First and foremost, I would like to thank Jotun for putting so much work into this course, and making it thoroughly enjoyable. Also, I would like to thank my class tutor for the much-valued feedback given on my essays throughout the term. Finally, I would like to thank Dr Petra Vértes for providing me with data to allow me to explore generative models of mental disorders in more depth in my report.

A.2   A note on figures

All figures original (drawn on Powerpoint/Mathcha.io).