# Complex Network Analysis

## Applications To Human Brain Functional Networks

Hoang Le

MASTER THESIS - UPF / Year 2012-2013

**Supervisors**
Xerxes D. Arsiwalla, Riccardo Zucca, Paul Verschure
**Department**
Synthetic, Perceptive, Emotive and Cognitive Systems Group
Department of Information and Communication Technologies
Universitat Pompeu Fabra
Barcelona, Spain

# Acknowledgement

I would like to thank my thesis advisors Xerxes D. Arsiwalla, Riccardo Zucca and Paul Verschure for their continuing support and guidance throughout my research. Without them this thesis would not have been completed.

# Abstract

We investigate the topology of human brain functional networks, using fMRI data. We re-examine the question of whether the degree distribution of these networks really scale as power-law (scale-free) and which statistical tests are better suited to answer such questions. Earlier studies have all been based on least-square estimation, which is not a reliable estimator of power-law distributions. Degree distribution of brain functional networks from 10 healthy individuals were analyzed using rigorous statistical analysis. The statistics do not support a power-law, but rather the generalized Pareto distribution. We propose methods to construct synthetic random and power-law networks from our empirical networks as a way to compare efficiency among these different models, using graph-theoretic measures.



**Keywords:** brain functional networks, graph theory, network science, scale-free networks, generalized Pareto distribution.

# Table of Contents

# List of Figures

# List of Tables

# Chapter 1

# INTRODUCTION

## 1.1 Problem Statement

Much interest in theoretical neuroscience has revolved around graph-theoretic scaling properties of the network of correlations in the human brain. Some authors have attempted to show that the degree distribution of nodes in brain functional networks are scale-free; that is, they obey power-law degree distributions $P(k) \sim k^{-\alpha}$ (see for example [3], [5], [6], and [7]). The scale-free model is theoretically attractive for several reasons, among which are (i) the seemingly ubiquitous presence of scale-free networks in nature, as claimed by a large body of work from other fields related to scale-free networks, including both non-biological and biological networks (see for example [8]); (ii) the link of scale-free networks to self-organized criticality, as can be seen in [9], [10], [11], and [12]; (iii) the existence of a fat tail, implying larger number of brain hubs compared to random or other small-world network models, ensuring efficiency of information processing and resilience (see [13] and [14]). However, some other authors have claimed that instead of being scale-free, brain functional networks follow a power law with truncated exponential distribution (see [4]).

We noticed a systematic methodological weakness of previous works in the literature ( [3], [5], [6], [7], [4]), as they mainly examined the structure of the brain functional networks based on either visual assessment [6], [4], or using least square error fitting on a log-scale to establish their claims [3], [5], [7]. Least square fitting in many case does not give good estimate of the scaling parameter $\alpha$. And even when it does, the errors are no longer normally distributed under log-log scale, thus the coefficient of determination $R^2$, frequently used to assess the goodness of fit in linear regression, cannot be a reliable goodness of fit indicator in this context.

We deploy rigorous statistical techniques to verify these claims. In addition,

structural and dynamical consequences of brain functional networks are further investigated in light of our results. Concretely:

(i) We disprove the scale-free hypothesis of the brain functional networks, which has been prevalent in the literature to date

(ii) We offer our own framework of the brain functional networks structure, verified through rigorous statistical analysis. We argue why our model is competitive with the scale-free models from a efficiency/cost perspective

(iii) We develop a hubs map of the brain functional networks, including both positive correlation hubs and negative correlation hubs. This map will serve as a basis for our dynamical modeling

## 1.2 State of the Art

### 1.2.1 Contemporary Network Science

The study of networks in the form of mathematical graph theory is one of the fundamental pillars of discrete mathematics. Euler's celebrated 1735 solution of the Konigsberg bridge problem is often cited as the first true proof in the theory of networks, and during the twentieth century graph theory has developed into a substantial body of knowledge ( [15], [16]).

Networks have also been studied extensively in the social sciences. Typical network studies in sociology involve the circulation of questionnaires, asking respondents to detail their interactions with others ( [17]). One can then use the responses to reconstruct a network in which vertices represent individuals and edges the interactions between them. Typical social network studies address issues of centrality (which individuals are best connected to others or have most influence) and connectivity (whether and how individuals are connected to one another through the network).

The last decade has witnessed the birth of a new movement of interest and research in the study of complex networks, i.e. networks whose structure is irregular, complex and dynamically evolving in time, with the main focus moving from the analysis of small networks to that of systems with thousands or millions of nodes, and with a renewed attention to the properties of networks of dynamical units. This flurry of activity, triggered by two seminal papers, that by Watts and Strogatz on small-world networks [18], and that by Barabasi and Albert on scale-free networks appeared one year later in Science [19], has been certainly induced by the increased computing powers and by the possibility to study the properties of a plenty of large databases of real networks. These include transportation

networks, phone call networks, the Internet and the World Wide Web, the actors collaboration network in movie databases, scientific co-authorship and citation networks from the Science Citation Index, but also systems of interest in biology and medicine, as neural networks or genetic, metabolic and protein networks. Within neuroscience, the interest in studying human brain from the perspective of network science is rapidly increasing, thanks to concepts and techniques developed from other disciplines, the development of brain imaging technologies, and the wealth of available data sets.

### 1.2.2 Basic Definitions and Notations

Graph theory is the natural framework for the exact mathematical treatment of complex networks and formally, a complex network can be represented as a graph. Within the scope of this report, we will use the term *network* and *graph* interchangeably. A undirected graph $G = (V, E)$ consists of two sets $V$ and $E$, such that $V \neq \emptyset$, and $E$ is a set of unordered (order) pairs of elements of $V$. $V$ is called set of *vertices* (also commonly known as *nodes*) and $E$ is a set of *edges* (also commonly called *links*), where elements consist of pair $u, v$ of distinct vertices $u, v \in V$.

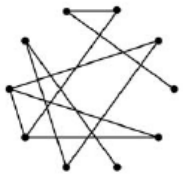
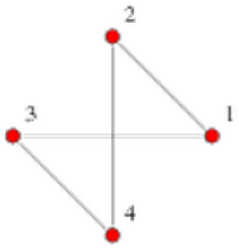


Figure 1.1: *illustration of a graph*

A graph can be *undirected* or *directed*, depending on whether the edges in $E$ have an ordering to its vertices (i.e., so that $u, v$ is distinct from $v, u$, for $u, v \in V$). Also, a graph can be *simple*, or a *multi-graph*, if there are multiples edges connecting two vertices. Within the scope of this report, we will deal mainly with simple, undirected graph.

More importantly, a graph can be *binary*, or *weighted*. A binary graph is one which a link indicates the presence of a relationship between two nodes (relationship either exists or does not exist). A weighted graph also incorporates connection strength into the links among vertices. Any simple, undirected graph can be uniquely represented in the form of an *adjacency* matrix. A graph with $n$ vertices and be isomorphically mapped to a $n \times n$ square *adjacency* matrix $A$, with each row (column) represents a vertex. The entry at row $i$ and column $j$ of the adjacency matrix $A$, or $a_{ij}$ indicates the connection strength between nodes

$i$ and $j$. It then follows that a binary graph can be represented by a *symmetric, binary* adjacency matrix, whereas a weighted graph is represented by a *symmetric, weighted* adjacency matrix. This one-to-one mapping between graphs and adjacency matrices greatly facilitates the study of network graphs with the help of formal mathematical tools, such as linear algebra.




$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Figure 1.2: *example of the correspondence between a graph and its adjacency matrix*

The *degree* $k_i$ of a node $i$ is the number of edges incident with the node, and is defined in terms of the adjacency matrix $A$ as: $k_i = \sum_{j \in V} a_{ij}$. The degree distribution of a network is an important feature in studying network topology. The *weighted degree* of a node is defined similarly.

### 1.2.3 Graph-theoretic Measures

#### Degree Distribution

The most basic topological characterization of a graph $G$ is its *degree distribution* $P(k)$, defined as the probability that a node chosen uniformly at random has degree $k$ or equivalently, as the fraction of nodes in the graph having degree $k$. The degree distribution provides a natural summary of the connectivity in the graph. During the past decade, it has been found that approximate power-law distributions appears to be ubiquitous in networks across many areas of the sciences [20]. This discovery was originally quite unexpected, as such structure is in contrast to that of networks studied throughout much of the 20th century [16],

such as traditional random graphs. In the case of random graphs, vertex degree is of a fairly similar order of magnitude across the graph, homogeneous instead of heterogeneous. The corresponding degree distribution are thus concentrated, and typically decay exponentially fast, rather than like a power-law. A power-law distribution is characterized by a "fat-tail", implying the existence of numerous network hubs, compared to other wise random networks. Due to its seemingly ubiquitous presence in nature, networks with power-law degree distributions have been the focus of a great deal of attention in the literature [21]. They are also referred to as *scale-free networks* [20]. Formally, the probability density function of a scale-free network takes the form $P(k) \sim k^{-\alpha}$. The term scale-free refers to any functional form $f(x)$ that remains unchanged to within a multiplicative factor under a rescaling of the independent variable $x$. The earliest published example of a scale-free network is Price's network of scientific citations [22], where the value of scaling parameter $\alpha$ is between 2.5 and 3. More recently, power-law degree distributions have been observed in a wide range of other networks, including other citation networks ( [23], [24]), the World Wide Web ( [25], [26], [27]), the Internet ( [28], [29], [30]), metabolic networks ( [31], [32]), telephone call networks ( [33], [34]), and the network of human sexual contacts ( [35], [36]). Other common functional forms for the degree distribution are exponentials, such that those seen in the power grid [37] and railway networks [38], and power laws with exponential cutoffs, such as those seen in the networks of movie actors [37] and some collaboration networks [39].

**Degree Correlation and Mixing Patterns**

The degree distribution is useful as a composite summary of how degree varies across nodes in the network, but it does not provide any information on precisely which nodes are connected to which others. To capture information of this sort, it is helpful to establish summaries that describe the patterns of association among nodes of similar degrees. Traditionally in the context of social network analysis, a pattern of selective linking where highly connected nodes tend to be connected to each other has been studied under the term *homophily*. Recently, a similar concept of *assortative mixing* has been explored for different types of networks [40]. A network is said to be *assortative* if high-degree vertices have a preference to attach to other high-degree vertices, and *disassortative* if high-degree vertices tend to connect to low-degree ones. This mixing pattern in networks can be summarized through *assortativity coefficient*, defined as:

$$r = \frac{l^{-1}\sum_{(i,j)\in E} k_i k_j - [l^{-1}\sum_{(i,j)\in E} \frac{1}{2}(k_i + k_j)]^2}{l^{-1}\sum_{(i,j)\in E} \frac{1}{2}k_i^2 + k_j^2 - [l^{-1}\sum_{(i,j)\in E} \frac{1}{2}(k_i + k_j)]^2}$$

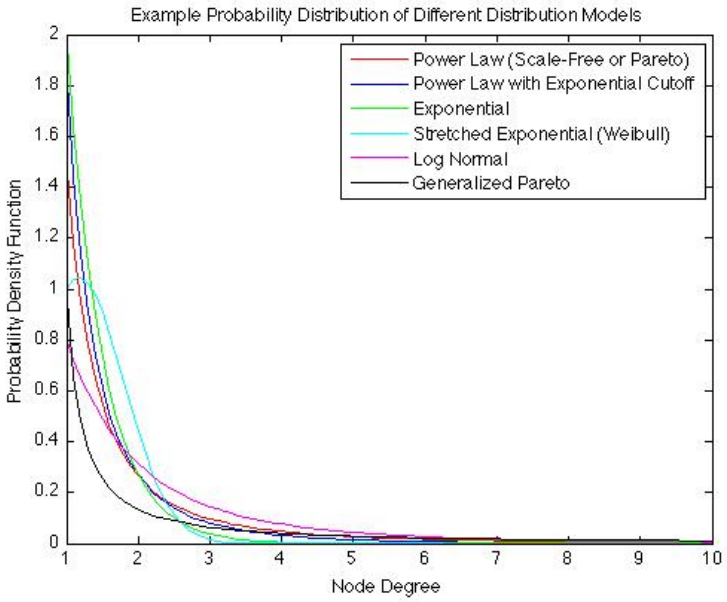



Figure 1.3: *illustrative probability density functions of popular distribution models*

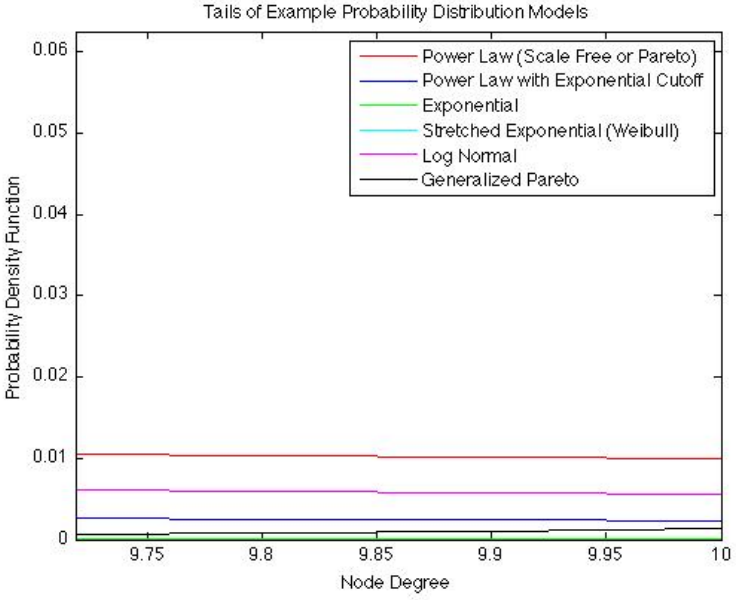



Figure 1.4: *illustrative probability density functions of the tails of popular distribution models, signifying the structure of network hubs*

with $l = \sum_{(i,j)\in V} a_{ij}$ is the number of links in the network. For weighted networks, the weighted assortativity coefficient can be defined similarly as [41]:

$$r = \frac{l^{-1}\sum_{(i,j)\in E} w_{ij}k_i^w k_j^w - [l^{-1}\sum_{(i,j)\in E}\frac{1}{2}(w_{ij}(k_i^w + k_j^w)]^2}{l^{-1}\sum_{(i,j)\in E}\frac{1}{2}w_{ij}(k_i^2 + k_j^2) - [l^{-1}\sum_{(i,j)\in E}\frac{1}{2}w_{ij}(k_i^w + k_j^w)]^2}$$

with $w_{ij}$ represents connection weight of link $(i, j)$, and $k_i^w = \sum_{i\in V} w_{ij}$ is the weighted degree of $i$. Networks with a positive assortativity coefficient are likely to have a resilient core of mutually interconnected high-degree hubs. On the other hand, networks with a negative assortativity coefficient are likely to have widely distributed and consequently vulnerable high-degree hubs. Some examples of assortative networks include scientific coauthorship and film actor collaboration networks [40], while the Internet, World Wide Web, protein interaction networks and networks of food web have been shown to be disassortative [40]. Notably, random networks and the scale-free networks defined by preferential attachment growth model of Barabasi and Albert have assortativity coefficient of 0 [40]. Related measure of assortativity computed on individual nodes is the *average neighbor degree* [42] $k_{nn,i} = \frac{\sum_{j\in V} a_{ij}k_j}{k_i}$ for binary networks and $k_{nn,i}^w = \frac{\sum_{j\in V} w_{ij}k_j^w}{k_i^w}$ for weighted networks.

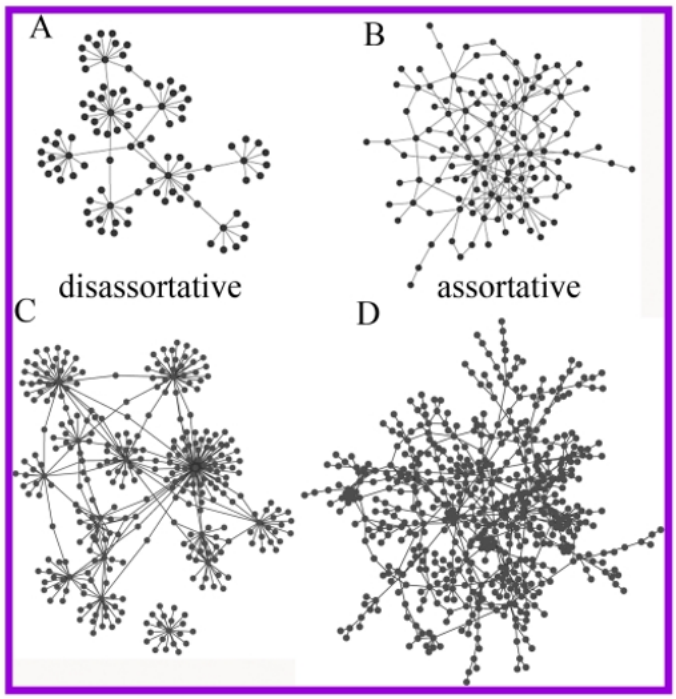



Figure 1.5: *illustrative assortative and disassortative networks, from [1]*

**Characterizing Network Cohesion and Connectivity, Small-World Properties**

A clear deviation from the behavior of random graphs can be seen in the property of network clustering. In many networks it is found that if vertex A is connected to vertex B, and vertex B to vertex C, then there is increased probability that vertex A will also be connected to vertex C. In the context of social networks, the friend of your friend is also likely to be your friend. This is exhibited in network topology through the number of triangles in the network. It can be quantified by defining a *clustering coefficient* $C$ such as [18]: $C = \frac{1}{n}\sum_{i\in V} C_i = \frac{1}{n}\sum_{i\in V}\frac{2t_i}{k_i(k_i-1)}$ where $t_i = \frac{1}{2}\sum_{j,h\in V} a_{ij}a_{ih}a_{jh}$ denotes the number of triangles around node $i$, and $C_i$ is the clustering coefficient of node $i$. This notion of clustering coefficient can be generalized for weighted networks as [43]:

$$C^w = \frac{1}{n}\sum_{i\in V} C_i^w = \frac{1}{n}\sum_{i\in V}\frac{2t_i^w}{k_i(k_i-1)}$$

where $t_i^w = \frac{1}{2}\sum_{j,h\in V}(w_{ij}w_{ih}w_{jh})^{\frac{1}{3}}$ is the weighted geometric mean of triangles around nodes $i$. Various higher-order clustering coefficients have also been proposed, among which are the k-clustering coefficient that accounts for k-neighbors ( [44], [45]), or other measures based on the internal structure of cycles of order four ( [46]), or on the number of cycles of a generic order [47]. Some definitions of clustering coefficients without bias of degree correlation have also been proposed ( [48], [49]). In general, regardless of which definition of the clustering coefficient is used, the values of real-world networks tend to be considerably higher than those of a random graph with similar number of vertices and edges.

Another important concept that characterizes the cohesion of a network is its *shortest path lengths* among different nodes. Shortest paths play an important role in the communication within a network. The idea has long been explored in the study of graph theory [50]. Shortest path between two nodes in a graph is typically determined computationally through the use of standard Dijkstra's algorithm, or the breadth-first search method. The *efficiency* of the internal structure of a network can be examined by looking at shortest paths among all vertices. A measure of the typical separation between two nodes in the graph is given by the *average shortest path length*, also known as the *characteristic path length* of a network, formally defined as [18]:

$$L = \frac{1}{n}\sum_{i\in V} L_i = \frac{1}{n}\sum_{i\in V}\frac{\sum_{j\in V, j\neq i} d_{ij}}{n-1}$$

where $d_{ij}$ is the shortest path length between $i$ and $j$. Note that for weighted networks, $d_{ij}^w = \sum_{a_{uv} \in g_{i \to j^w}} f(w_{uv})$, where $f$ is a map (typically an inverse) from weight to length, and $g_{i \to j^w}$ is the shortest weighted path between $i$ and $j$. For unconnected network, $d_{ij}$ can be $\infty$, thus it is sometimes more convenient to looks at the *global efficiency* of a network [14]:

$$E = \frac{1}{n} \sum_{i \in V} E_i = \frac{1}{n} \sum_{i \in V} \frac{\sum_{j \in V, j \neq i} d_{ij}^{-1}}{n-1}$$

where $E_i$ is the efficiency of node $i$. The definition of *global efficiency* for weighted networks can be derived similarly.

An explosion of interest in network science emerged after a seminal paper from Duncat Watts and Steven Strogatz came out in 1998, in which they studied a set of so-called *small-world* networks [18]. The small-world effect was first studied by Stanley Milgram in the 1960s [51], in which letters passed from person to person were able to reach a designated target individual in only a small number of steps (around 6 in the published case). Watts and Strogatz proposed to define a class of small-world networks as those having both a small value of characteristic path length $L$, like random graph, and a high clustering coefficient $C$, like regular lattices. Such a definition corresponds to networks efficient in exchanging information both at a global and local scale. Built on this characterization of small-world networks, recently a quantitive measure of *small-world-ness* was suggested by [52], in which network small-worldness $S = \frac{C/C_{rand}}{L/L_{rand}}$ where $C$ and $C_{rand}$ are the clustering coefficients, and $L$ and $L_{rand}$ are the characteristic path lengths of the respective tested network and a random network. Small-world networks often have $S \gg 1$.

### Characterizing Network Hubs, Centrality

Many questions that might be asked about a node in a network essentially seek to understand its importance in the network. This importance can be expressed through how well it is integrated into the rest of the network, or vice versa, the potential impact of deleting this node from the network. A similar concept can also be defined for the importance of a certain link in a network. In network science, measures of centrality are designed to quantify such notion of importance. The most obvious measure of node centrality is its degree $k$. Later when we define functional brain networks, however, the degree of a node can take on different meanings, depending on how exactly the network is constructed. Particularly in the context of weighted networks that allow for both positive and negative links,

the definition of hubs will depend on whether we look at positive and negative links as a whole in the network, or treat them separately. Regardless of the treatment, developing a map of network hubs can be very useful in capturing main functionalities, while allowing a certain degree of simplification to take place. This network hubs characterization is helpful especially in the context of modeling dynamical processes in a given network.

Beside degree of nodes, two other measures of node centrality can be used to examine the prominence of nodes in the network. *Closeness centrality* is defined as the inverse of the average shortest path length from one node to all other nodes in the network. A related and often more sensitive measure is *betweenness centrality*, defined as the fraction of all shortest paths in the network that pass through a given node. Formally

$$L_i^{-1} = \frac{n-1}{\sum\limits_{j \in V, j \neq i} d_{ij}}$$

denotes the closeness centrality of node $i$ and betweenness centrality of node $i$ is defined as:

$$b_i = \frac{1}{(n-1)(n-2)} \sum_{h,j \in V, h \neq j, h \neq i, j \neq i} \frac{\rho_{hj}^{(i)}}{\rho_{hj}}$$

where $\rho_{hj}$ is the number of shortest paths between $h$ and $j$, and $\rho_{hj}^{(i)}$ is the number of shortest paths between $h$ and $j$ that pass through $i$

Bridging nodes that connect disparate parts of the network often have a high betweenness centrality. The notion of betweenness centrality is naturally extended to links and could therefore also be used to detect important connections within a network.

### 1.2.4 Brain Functional Networks

#### Overview of Brain Networks

We now switch our discussion to recent works on various types of brain networks, with a special focus on functional networks. Two main factors contributed to the recent wave of interest in studying the brain through the lens of network science. First, the development of technical tools from graph theory, some of which described above, and increased computational power have reached a point of cross-fertilization where data-rich fields such as computational neuroscience can be meaningfully studied with the help of these new techniques. Second, modern brain mapping techniques, such as diffusion MRI, functional MRI, EEG, and MEG produce increasingly large data sets of anatomical and functional connectivity patterns, gradually allowing researchers for the first time meaningfully map

the entire brain in the form of massive networks, also known as Connectome ( [53], [54], [55]), in increasingly high level of resolution. Brain connectivity data sets comprise networks of brain regions connected by anatomical tracts or by functional associations. The three main types of brain networks can be broadly classified as follows [56]:

- **Structural networks:** structural brain networks correspond to fiber density of white matter tracts between pairs of brain regions. Diffusion magnetic resonance imaging allows the mapping of the diffusion process of molecules, mainly water, in biological tissues, in vivo and non-invasively. The results from diffusion MRI can be used to build a tractography of whole brain, providing an estimate of axonal trajectories across the entire white matter [54]. Together with a parcellation of the brain into different regions of interests (ROIs), connection weight between each pair of ROIs can be computed to build a structural network of the brain. Initial studies of structural brain networks showed that individual brain networks have an exponential node degree distribution and their global organization is in the form of a small-world [53]. However, it should be noted that construction of structural networks is still in relatively early phase and thus not many data sets are currently available for more in-depth studies.

- **Functional networks:** functional brain networks correspond to magnitudes of temporal correlations in activity between pairs of brain regions. Two main methods are typically used to construct functional networks.

    1. The functional networks can be derived by calculating cross-correlations between BOLD signals from different brain regions throughout a fMRI session. The smallest unit of brain regions is called brain "voxel" (of dimension 3x3.475x3.475 $mm^3$). Magnetic resonance brain activity is measured in each voxel at each time step. Two brain sites are functionally connected if their Pearson temporal correlation exceeds a threshold value $r_c$, regardless of their anatomical connectivity. No clear rule exists for the choice of threshold. However, most studies have considered positive thresholds of at least 0.5
    2. Alternatively, some researchers applied discrete wavelet transform to fMRI time series to estimate frequency-dependent correlation matrices characterizing functional connectivity between brain regions [4]. Here, wavelet transform effects a time-scale decomposition that partitions the total energy of a signal over a set of compactly supported basis functions, or little waves, each of which is uniquely scaled in frequency and located in time [4]. The result is correlation matrices corresponding to each range of frequency.

Either method arrives at a correlation matrix of connectivity that is dependent on the choice of threshold. After thresholding, connectivity matrices are typically binarized, with functional connection between pairs of brain regions assigned values of 1, with the rest being 0.

- **Effective networks:** effective brain networks represent direct or indirect causal influences of one region on another and may be estimated from observed perturbations [57]. Causal interactions are computed using transfer entropy, a measure of directed information flow. Thus effective brain networks take the form of directed graphs.

The techniques to construct structural and effective brain networks are evolving and still in relatively early stages. An illustration of the standard method to construct structural and functional brain networks is provided below in figure 1.6.The focus of this thesis will primarily be in the context of functional networks. For the remainder of the report, brain networks will imply functional networks.

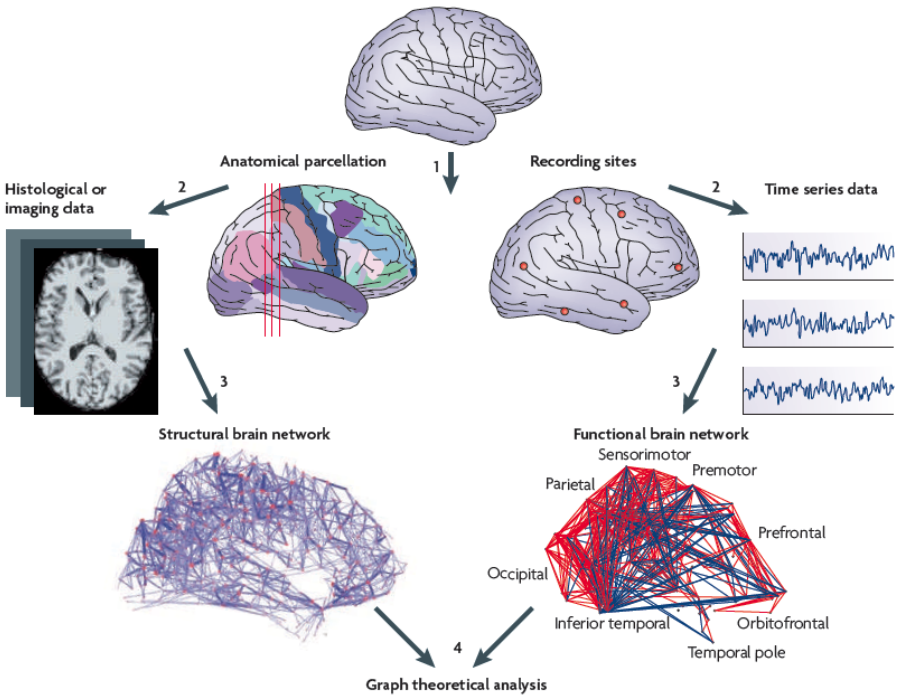



Figure 1.6: *example of structural and functional networks construction, image from [2]*

## Graph-theoretic Measures in the Context of Brain Networks

An individual network measure may characterize one or several aspects of global and local brain connectivity. At a macro level, measurement values of all individual degree elements comprise the distribution of brain networks. The *degree distribution* is an important marker of network development and resilience. The mean network degree is most commonly used as a measure of *density*, or the total wiring cost of the network [58]. Degree distribution of brain networks may hold important clues to the dynamical processes on the networks, as the high-level of interest on the hypothesized scale-free properties of brain functional networks indicate [10], [6], [12], [9]. Furthermore, degree distribution also indicates to a certain extent the resilience of the network. For instance, complex networks with power-law degree distributions may be resilient to gradual random deterioration, but highly vulnerable to disruption of high-degree central nodes [13]. Another useful measure of resilience is the *assortativity coefficient*. Networks with a positive assortativity coefficient are likely to have a comparatively resilient core of mutually interconnected hubs [56]. The effect of lesions of human brains or the effect of neuro-degeneration can be quantified by looking at these macro measures.

At a lower level, functional segregation in the brain is the ability for specialized processing to occur within densely interconnected groups of brain regions [56]. Measures of segregation, such as *clustering coefficient*, quantify the presence of such groups, known as clusters and modules, within the network. More sophisticated measures of segregation not only describe the presence of densely interconnected groups of regions, but also find the exact size and composition of these individual groups [59]. At the same time, functional integration in the brain is the ability to rapidly combine specialized information from distributed brain regions. Measures of integration, such as the *characteristic path length*, characterize the ease with which brain regions communicate. Lengths of path consequently estimate the potential for functional integration between brain regions, with shorter path implying stronger potential for integration. Paths in functional networks represent sequences of statistical associations and may not correspond to information flow on anatomical connections [56], and thus provide another dimension for analysis. A combined balance of functional integration and segregation in the form of small-world networks was hypothesized to be a well-design structure, allowing the brain to simultaneously reconcile the opposing demands of processing the information efficiently at both the global and local level. Such a design appears to be a feature of anatomical connectivity [60]. In addition, several studies examining functional networks also report varying degree of small-worldness [4]. Given the more abstract nature of functional paths, a more complete understanding of the relationship between structural dynamics

and functional connectivity will help clarify this issue [61].

The characterization of brain hubs provides a way to study the simplified brain networks, while allowing for the capture of main structural / functional properties. The *degree* is the most common indication of brain hubs. Other measures of centrality are based on the idea that central nodes participate in many short paths within a network and consequently act as important controls of information flows [62]. Measures of centrality may have different interpretations in structural and functional networks. For instance, anatomically central nodes often facilitate integration, and consequently enable functional links between anatomically unconnected regions. Such links in turn may make central nodes less prominent and so reduce the sensitivity of centrality measures in functional networks [56].

### Topological Properties of Brain Functional Networks

Much interest in the analysis of brain networks has revolved around topological structure of the networks, especially scaling properties of the network. In the first report on large-scale topology of brain functional networks, Eguiluz et al claimed that functional networks are scale-free, with scaling parameter $\alpha \approx 2$ [3]. The constructed network in this case came from 36x64x64 brain voxels, each measured at 400 time steps, with 2.5 seconds spacing. The method of establishing scale-free properties, as displayed in figure 1.1, is to create a histogram of degree frequency on a log-log scale, across several different levels of correlation thresholds ($r_c$ = 0.5, 0.6, 0.7 or 0.8). The best fit straight line is then fitted through the data at different thresholds, and the estimated scaling parameter $\alpha$ would be the slope of this line, as can be seen in figure 1.7. In addition, they claimed that char-

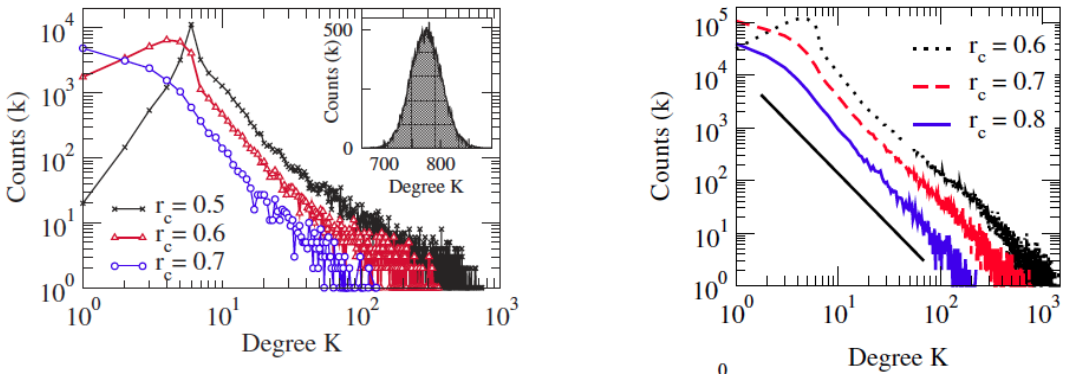



Figure 1.7: *Log-log plot of degree distribution from Eguiluz et al. [3]*

acteristic path length is small and comparable with those of equivalent random networks, and the clustering coefficient is orders of magnitude larger than those of equivalent random networks. After this study, several other have come out in

agreement with the hypothesis that the brain functional networks are scale-free, notably [6], [5], and [7]. The scale-free model is theoretically attractive for several reasons, among which are (i) the seemingly ubiquitous presence of scale-free networks in nature, as claimed by a large body of work from other fields related to scale-free networks, including both non-biological and biological networks [8]; (ii) the link of scale-free networks to self-organized criticality, which is a property of dynamical systems which have a critical point as an attractor [9], [10], [11], and [12].Their macroscopic behaviour thus displays the spatial and/or temporal scale-invariance characteristic of the critical point of a phase transition; (iii) the existence of a fat tail, implying larger number of brain hubs compared to random or other small-world network models, ensuring efficiency of information processing and resilience (see [13] and [14]).

However, some other authors have claimed that instead of being scale-free, brain functional networks follow a power law with truncated exponential distribution (see [4], [63], [64]). Formally, a power law with exponential cutoff can be expressed by the probability distribution function $P(k) \sim k^{-\alpha} e^{-\lambda k}$. Archard et al studied the brain functional networks through constructing networks of 90 nodes, and estimated scaling parameter $\alpha = 1.8$ and cutoff degree $\lambda = 0.2$ [4]. Again, log-scale plots of histogram of degree frequency were examined in their study. One such plot is displayed in figure 1.8 below. In addition, they observed global mean path length of 2.49, which is approximately equivalent to a comparable random network, whereas clustering coefficient of 0.53 is two times greater. They concluded that low-frequency functional networks have a small-world architecture, but are not scale-free. In addition, the network is more resilient to targeted attack on its hubs than a comparable scale-free network, but about equally resilient to random error.

### 1.2.5 Methodological Weakness of Previous Studies

Several methodological issues exist with prior studies on topological structural of functional networks. Here we address several key points which motivate our current study.

- The commonly used method for analyzing the degree distribution of functional networks in prior studies is either least-square fitting or visual assessment. First, using visual assessment is not a reliable way of establishing the power-law relationship, as many heavy-tail distributions can share the feature of noisy data towards the tail of the distribution on a log-log scale, thus can be highly misleading. Second, least-square fitting can produce substantially inaccurate estimates of parameters for power-law distributions. And

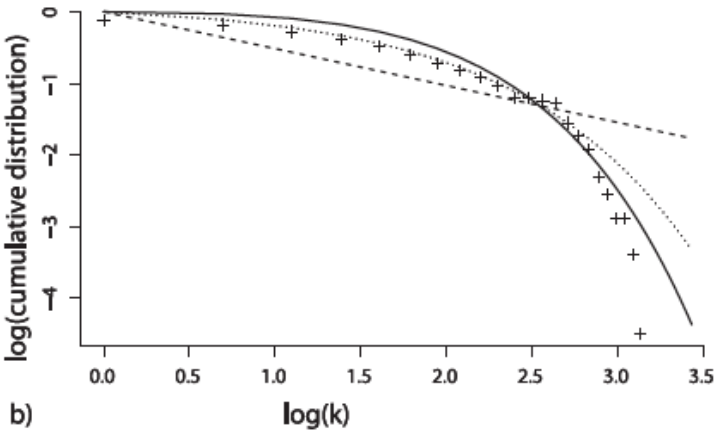



Figure 1.8: *Log-log plot of degree distribution from Achard et al. [4]. The plus sign indicates observed data, the solid line is the best-fitting exponentially truncated power law, the dotted line is an exponential, and the dashed line is a power law.*

even when in cases where such methods return accurate answers, such methods will not be able to give satisfactory indication of whether the data obey a power law [65]. To date, the most common approach for testing empirical data against a hypothesized power-law distribution is to transform the distribution $P(k) \sim k^{-\alpha}$ into the log form $\log P(k) = c - \alpha \log k$. The probability density $P(k)$ can be estimated by constructing a histogram of the data and the resulting function can then be fitted to the linear form by least-square linear regression. The slope of the fit is interpreted as the estimated $\hat{\alpha}$ of the scaling parameter, and $r^2$ is taken as an indicator of the quality of the fit. Table 1.1 illustrates how least-square fitting can wildly mis-calculate the scaling parameter, by generating synthetic data sets generated from a priori known power-law distribution curve with $\alpha = 2$.

| Number of Synthetic Data Points | LSQ Estimated $\alpha$ |
|---|---|
| 50000 | 1.0589 |
| 100000 | 1.1691 |

Table 1.1: *Inaccuracy of LSQ estimation on a priori known power-law distribution with $\alpha$ = 2*

More seriously, a fit to a power-law distribution can account for a large fraction of the variance even when the fitted data do not follow a power-law, and hence high values of $r^2$ cannot be taken as evidence in favor of

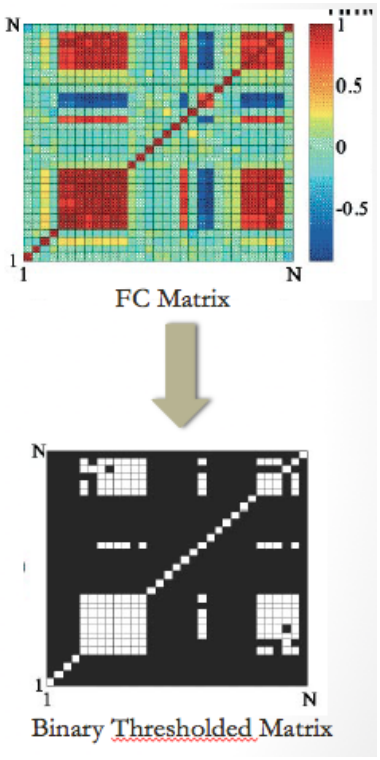



Figure 1.9: *typical FC matrix and binary thresholded adjacency matrix, image from Eguiluz et al. [3]*

the power-law form. In addition, the fits extracted by regression methods usually do not satisfy basic requirements on probability distributions such as normalization, and hence can be incorrect.

- Previous studies frequently transformed extracted functional correlation (FC) matrices into binary adjacency matrices. This has the potential to introduce more errors into the constructed networks, and ignore important gradient of functional relationships. A binary network with threshold $r_c = 0.4$, for example, would view cross-correlation of 0.5 and 0.9 to be functionally equivalent. A typical example can be seen in figure 1.9.

- By considering only positive thresholds, potentially important information regarding anti-correlation relationships in functional networks are ignored by previous studies. Several authors have argued for potential relevance of anti-correlation in functional brain context [66], [67], [68], [69].

# Chapter 2

# METHODOLOGY

## 2.1 Data and Construction of Brain Functional Networks

### 2.1.1 Data Acquisition

#### 1000 Functional Connectomes Project

Our primary data source comes from the 1000 Functional Connectomes Project, an open-access repository of resting-state functional MRI datasets [55]. The 1000 Functional Connectomes Project (http://www.nitrc.org/projects/fcon_1000/) is an international open-access repository of resting-state functional connectivity MRI datasets with subjects recruited in different cohorts across the world. For consistency and due to computational limit, we select 10 healthy, right-handed male subjects from the Ann Arbor, Michigan cohort with age ranging from 18 to 33 for our study. All datasets were reoriented to RPI. Also, the first 5 time points of each time series were discarded, leaving each dataset with 295 time points across 64x64x40 brain voxels. Different levels of resolution were considered. At the lowest resolution level V1, collection of neighboring 4x4x2 brain voxels were combined by taking the average fMRI signal at each time step, effectively transforming the original data into 16x16x20 brain sites. At resolution level V2, collection of neighboring 4x4x1 brain voxels were combined to obtain time series of 16x16x40 brain sites. Similarly at level V3, neighboring 2x2x2 brain voxels were combined to obtain time series data of 32x32x20 brain sites.

#### Task-based fMRI data from Eguiluz et al.

Original task-based fMRI data used in [3] consist of four healthy, right-handed subjects. Subjects were studied using a Siemens-Trio 3.0 Tesla imaging system

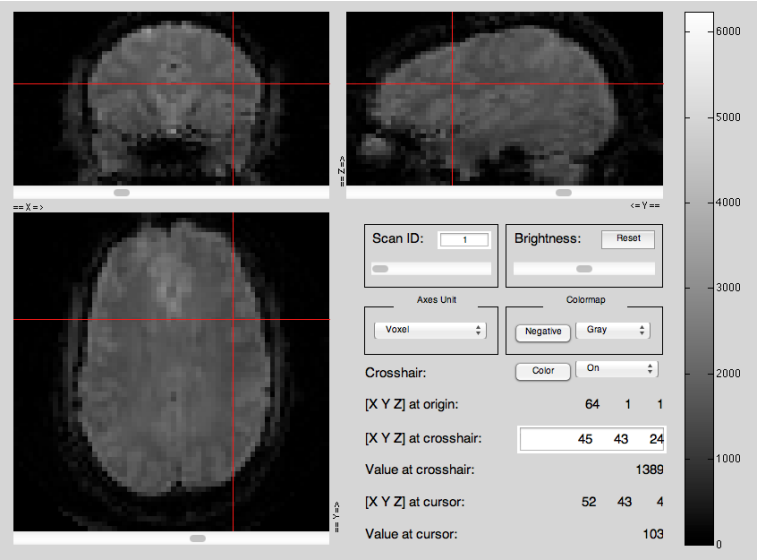



Figure 2.1: *sampled image of fMRI session from subject 34781*

using a birdcage radio-frequency head coil. The data were preprocessed using the package FSL (http://www.fmrib.ox.ac.uk/fsl). Subjects performed on-off finger tapping with threes different protocols. In one case they were instructed verbally to start and stop tapping, in the other one the start or stop cue was a small green or red dot in a video screen, and in the last one the start or stop cue was the entire screen turning green or red [3].

#### Human Connectome Project

Recently, a consortium of universities led by Washington University at St. Louis and University of Minnesota has initiated a new effort called Human Connectome Project (http://www.humanconnectome.org/) to collect structural and functional MRI data. The data acquisition and processing are still underway. At this stage, we have received sampled, one subject sets of both resting-state fMRI data set and task-based fMRI data sets for language, emotional, gambling, motor skill, and working memory task.

### 2.1.2 Network Construction

We adopt the standard approach to construct brain functional networks from fMRI data as presented in the literature [3], [5], [63]. To derive correlation matrix, Pearson correlation coefficient between any pair of brain regions $x_1$ and $x_2$ is

defined as:

$$r(x_1, x_2) = \frac{\langle V(x_1,t)V(x_2,t)\rangle - \langle V(x_1,t)\rangle\langle V(x_2,t)\rangle}{\sigma(V(x_1))\sigma(V(x_2))}$$

where the activity of brain region $x$ at time $t$ is denoted as $V(x,t)$, $\sigma^2(V(x)) = \langle V(x,t)^2\rangle - \langle V(x,t)\rangle^2$, and $\langle . \rangle$ represents temporal averages. Figure 2.2 displays the resulting correlation matrix for subject 34781 from the 1000 Functional Connectome Project. From the FC matrix, functional networks can be extracted by looking at a range of different thresholds. For each threshold $r > 0$, the *weighted adjacency matrix* is obtained by keeping all values in the correlation matrix that are greater than or equal to $r$, while other entries become 0. For each threshold $r < 0$, such adjacency matrix is obtained by keeping the values that are less than or equal to $r$, while other entries go to 0. We consider 17 different thresholds for each subject, corresponding to 17 different extracted weighted networks. The threshold values range from $r = -0.7$ to $r = 0.8$, with each increment of 0.1. Note that for $r > 0.8$ and $r < -0.7$, extracted network will become too sparse for meaningful analysis. An example of thresholded matrix corresponding to $r = 0.4$ for subject 34781 at resolution level V2 can be seen in figure 2.3. This adjacency matrix has one-to-one relationship with a functional brain network at the given threshold and resolution level. Each row (column) represents a node in the network. Node degree is simply the sum of each corresponding row (column) of the adjacency matrix. Graph-theoretic measures of the constructed network can be performed on the corresponding weighted adjacency matrix.

Computationally, the construction of brain functional networks and all the analysis were performed using Matlab R2009 (Mathworks Inc.). In many cases where functional networks are sufficiently large (20,000 nodes or more), the network construction process can become computationally expensive. An efficient strategy comprises of following steps is needed to reduce running time: (i) storage of correlation matrix in single format instead of double format to ensure the matrix can be efficiently loaded into random access memory for processing. Note that using single format does not compromise the analysis of the data, given single format in Matlab can be accurate up to 7 decimal digits (ii) utilization of parallel processing toolbox in Matlab. This could help reduce running time by approximately 25 percent (iii) building correlation matrix by dividing the original group of brain sites into multiple blocks. As an example, when the original set of brain sites is divided into 2 blocks (block 1 and block 2), a full correlation matrix can be constructed by concatenating 4 different sub-matrices: M1 = block 1 x block 1, M2 = block 1 x block 2, M3 = block 2 x block 2, M4 = block 2 x block 1. This strategy is needed when the number of brain sites becomes too large to hold

the entire correlation matrix in random access memory. In practice, M1 and M3 are symmetrical, and M4 is the transpose of M2, thus we can effectively calculate only half of the pair-wise correlation coefficients to build the full correlation matrix.

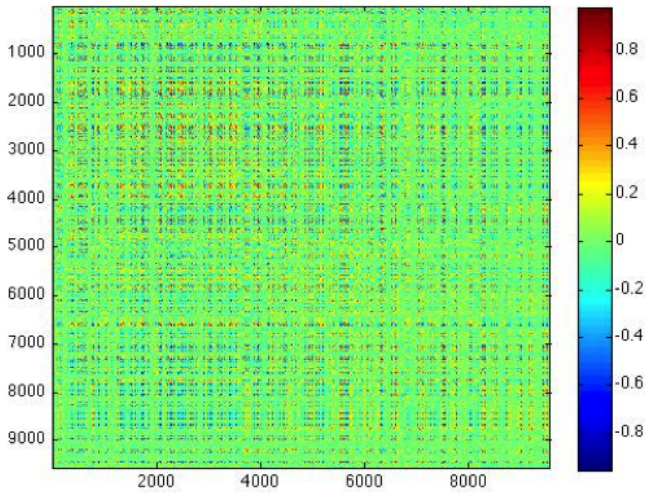



Figure 2.2: *constructed correlation matrix from subject 34781 at resolution level V2*

## 2.2 Degree Distribution Testing

### 2.2.1 Testing Power-Law Distribution

From the previous section, assume that we have constructed a *weighted functional network* with $n$ nodes and a weighted degree sequence $x_1 \leq x_2 \leq ... \leq x_n$. To test whether the degree distribution of functional networks follows a power-law, we follow the method suggested by [65], which advocates using Maximum Likelihood Estimator to estimate the scaling parameter $\alpha$ and a parametric bootstrap method to test the goodness of fit.

#### Estimate the parameters $x_{min}$ and $\alpha$ of the power-law model using method of *maximum likelihood* and Kolmogorov-Smirnov statistic

In practice, few empirical phenomena obey power laws for all values of $x$. More often the power law applied only for values greater than some minimum $x_{min}$. In

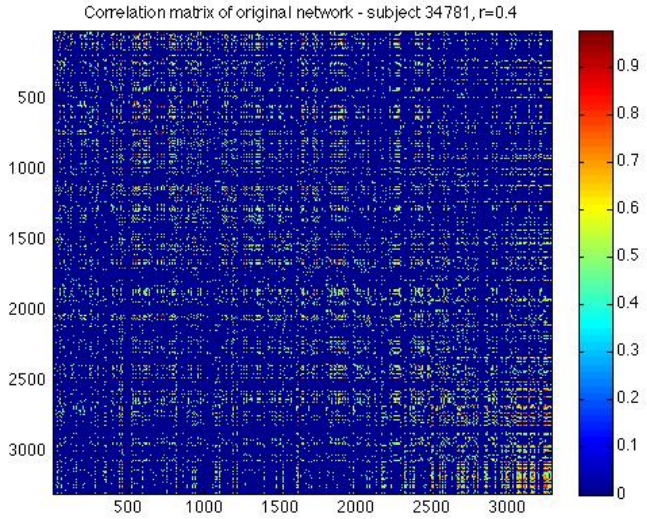



Figure 2.3: *thresholded correlation matrix from subject 34781,* $r = 0.4$*, resolution level V1*

such cases the tail of the distribution follows a power law. For each $\alpha$ and $x_{min}$, the probability density function for power-law distribution is given by

$$p(x) = \frac{\alpha - 1}{x_{min}}(\frac{x}{x_{min}})^{-\alpha}$$

. The *likelihood* of the data given the model is the conditional probability that the data were drawn from the model given $\alpha$:

$$p(x \mid \alpha) = \prod_{x_i \geq x_{min}} p(x_i \mid \alpha)$$

The data are most likely to have been generated by the model with scaling parameter $\alpha$ that maximizes this function. Note that given our set of degree sequence, this is a single variable function of $\alpha$. Thus finding the *maximum likelihood estimate*, or MLE of scaling parameter $\hat{\alpha}$ becomes the task of solving for the maxima of this likelihood function. In case of power-law distribution, there is a closed-form solution for $\hat{\alpha}$, which is given by:

$$\hat{\alpha} = 1 + k(\sum_{x_i \geq x_{min}} \ln \frac{x_i}{x_{min}})^{-1} \tag{2.1}$$

with $k$ is the number of $x_i \geq x_{min}$. Thus for each possible value of $x_{min}$, the estimated scaling parameter is uniquely determined by equation 2.1. There remains

the question of how we should go about choosing $x_{min}$. Clauset et al. suggested that we choose $\hat{x_{min}}$ that gives the best possible power-law fit out of all possible $x_{min}$. Kolmogorov-Statistic, which measures the distance between the *cumulative density functions* (CDFs) of the data and the fitted model, is commonly used to quantify this degree of fitness. For each $x_{min}$, Kolmogorov-Smirnov (KS) statistic is given by:

$$D = \max_{x \geq x_{min}} \mid S(x) - P(x) \mid$$

where $S(x)$ is the CDF of the data for the observations with value at least $x_{min}$, and $P(x)$ is the CDF for the power-law model that best fits the data in the region $x \geq x_{min}$. The estimated $\hat{x_{min}}$ is the value of $x_{min}$ that minimizes D.

**Calculate the goodness-of-fit between the data and the power-law using a parametric bootstrap based on the parameters estimated from previous step**

*Goodness-of-fit* test is conducted by generating a large number of power-law distributed synthetic data sets with scaling parameter $\alpha$ and lower bound $x_{min}$ equal to those of the distribution that best fits the observed data. We then fit each synthetic data set individually using its own power-law model and calculate the KS statistic for each one relative to its own model. Then we count the fraction of the time the resulting KS statistic is larger than the KS statistic value for the empirical data (as determined from the previous step). This fraction is our p-value. If this p-value is very small (less than 0.1), then power-law distribution is not a good model for our observed data. Otherwise if p-value is greater than 0.1, then we do not reject the hypothesis that the observed data follows a power-law distribution. The Matlab code to conduct the power-law test according to the procedure described here was developed by Clauset [65]. We adopt the code with some minor modifications.

It is very important to note that failure to reject power-law distribution is no guarantee for the power law being the best model for the empirical data. It is entirely possible that other families of distribution may be able to explain the data better. Test for other families of distribution, as well as for model selection, are discussed in the following sections

### 2.2.2 Other Families of Distribution

A *goodness-of-fit* test as laid out in the previous section can be used to rule out distribution hypothesis in the event the calculated p-value does not satisfy certain critical value threshold. It does not, however, guarantees the tested model to be the best model for the observed data. As such, we expand our analysis to consider

other popular 1-parameter and 2-parameter models that have appeared in the literature. The expression for each of the considered distribution families is given in table 2.1 as $p(x) = Cf(x)$, with $C$ being a constant.

| Distribution Name | $f(x)$ | $C$ | Parameter Condition |
|---|---|---|---|
| power law with cutoff | $x^{-\alpha}e^{-\lambda x}$ | $\frac{\lambda^{1-\alpha}}{\Gamma(1-\alpha,\lambda x_{min})}$ | $\alpha > 0, \lambda > 0$ |
| exponential | $e^{-\lambda x}$ | $\lambda e^{\lambda x_{min}}$ | $\lambda > 0$ |
| log-normal | $\frac{1}{x}exp[-\frac{(\ln x-\mu)^2}{2\sigma^2}]$ | $\sqrt{\frac{2}{\pi\sigma^2}}[erfc(\frac{\ln x_{min}-\mu}{\sqrt{2}\sigma})]^{-1}$ | $\mu, \sigma \in \Re$ |
| Weibull | $x^{\beta-1}e^{-\lambda x^\beta}$ | $\beta\lambda e^{\lambda x_{min}^\beta}$ | $\lambda > 0, \beta > 0$ |
| generalized Pareto | $(1 + k\frac{x-x_{min}}{\sigma})^{-1-\frac{1}{k}}$ | $\frac{1}{\sigma}$ | $\sigma > 0$ |

Table 2.1: *Other families of distribution, $f(x)$ is the functional form of probability density function (pdf), C is the normalizing constant of the pdf, such that* $\int_{x_{min}}^{\infty} Cf(x) = 1$

Note that in general, the testing procedures described in previous section can also be applied to each of these distribution families. We adapt the Matlab program for power-law distribution test for other distribution families. The implementation, however, is more challenging, as solutions for the *maximum likelihood estimate* do not exist in closed-form expression, with the exception of exponential distribution. Numerical solutions are thus required using Matlab's optimization toolbox.

### 2.2.3 Model Selection

In the event that two or more distribution families "pass" the statistical test, or more precisely, cannot be ruled out based on criteria described in the previous section, we use *likelihood ratio test* first suggested by [70] to determine which one is a better model for the observed data. The basic idea behind the likelihood ratio test is to compare the likelihood of the data under two competing distributions. The one with the higher likelihood is then the better fit. Alternatively one can calculate the ratio of the two likelihoods, or equivalently the logarithm $\Re$ of the ratio, which is positive or negative depending on which distribution is better or zero in the event of a tie. The sign of the log likelihood ratio, however, will not definitely indicate which model is the better fit because like other quantities, it is

subject to statistical fluctuation. If its true value, meaning its expected value over many independent data sets drawn from the same distribution, is close to zero, then the fluctuations could change the sign of the ratio and hence the results of the test cannot be trusted. In order to make a firm choice between distributions we need a log likelihood ratio that is sufficiently positive or negative that it could not plausibly be the result of a chance fluctuation from a true result that is close to zero. To make a quantitative judgement about whether the observed value of $\Re$ is sufficiently far from zero, we use the results from [70] to calculate the standard deviation $\sigma$ of $\Re$. This method gives us a p-value that tells us whether the observed sign of $\Re$ is statistically significant.

In technical terms, consider two candidate distributions of observed data with density function $p_1(x)$ and $p_2(x)$ respectively. The *log likelihood ratio* can be derived as:

$$\Re = \sum_{i=1}^{n} [\ln p_1(x_i) - \ln p_2(x_i)] = \sum_{i=1}^{n} [\ell_i^{(1)} - \ell_i^{(2)}]$$

where $\ell_i^{(j)} = \ln p_j(x_i)$. The variance of the difference $\ell_i^{(1)} - \ell_i^{(2)}$ can be approximated as:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} [(\ell_i^{(1)} - \ell_i^{(2)}) - (\bar{\ell}^{(1)} - \bar{\ell}^{(2)})]^2$$

with $\bar{\ell}^{(1)} = \frac{1}{n} \sum_{i=1}^{n} \ell_i^{(1)}$ and $\bar{\ell}^{(2)} = \frac{1}{n} \sum_{i=1}^{n} \ell_i^{(2)}$. The critical p-value, or the probability that the measured log likelihood ratio has a magnitude as large or larger than the observed value $|\Re|$, is given by:

$$\begin{aligned} p &= \frac{1}{\sqrt{2\pi n\sigma^2}} [\int_{-\infty}^{-|\Re|} e^{-t^2/2n\sigma^2} dt + \int_{|\Re|}^{\infty} e^{-t^2/2n\sigma^2} dt] \\ &= |erfc(\Re/\sqrt{2n}\sigma)| \end{aligned}$$

where $erfc(z) = 1 - erf(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} \exp -t^2 dt$ is the complementary Gaussian error function, which can be calculated using Matlab.

If this p-value is small ($p < 0.1$) then it is unlikely that the observed sign is a chance result of fluctuations and the sign is a reliable indicator of which model is the better fit to the data. If p is large on the other hand, the sign is not reliable and the test does not favor either model over the other.

## 2.3 Networks Model Comparison

Having rigorously tested different network models, we will incorporate graph-theoretic measures introduced in chapter 1 to compare the topological features between our empirical networks, and equivalent networks of hypothesized models. We are especially interested in examining measures that relate to efficiency of networks, from different graph-theoretic angles. This requires creating equivalent networks of different hypothetical distributions from our empirical network. In other words, from our empirical network, we want to derive null network model for comparison purposes. This issue has not been addressed widely in the literature ( [71]). So far, most previous studies have dealt with binary, sparse networks ( [72], [71]). The basic idea is to randomly select four distinct nodes $A$,$B$,$C$ and $D$ in a binary, sparse network so that there is a connection from $A$ to $B$, and from $C$ to $D$. In addition, the selection criterion is such that no connection exists between $A$ to $D$, and $B$ to $C$. At each step, we can replace the connection $A \longleftrightarrow B$ and $C \longleftrightarrow D$ with those of $A \longleftrightarrow D$ and $C \longleftrightarrow B$ [72]. The random, binary network is obtained by repeating this process over many iterations. It is easy to see that for binary, sparse networks, this procedure will preserve the degree distribution of the original network, since the degree of each node does not change after each rewiring operation.


Figure 2.4: *Binary Rewiring Algorithm in the Literature*

The method of rewiring binary network has been the standard by which null model is created so that graph-theoretic measures such as small-world properties are calculated [56]. However, when dealing with weighted networks, this rewiring method does not work since the connection weight of each pair of nodes can be different. We further develop several methods to work with weighted networks as described in what follows.

### 2.3.1 Creating Null Model by Rewiring

One way to create a null model for a weighted functional network in a manner similar to binary rewiring so that the degree distribution is preserved, we modify the connection strength among different nodes in a way that preserves the degree

of each node at each step of iteration. This method works for both dense and sparse networks.


Figure 2.5: *Our algorithm for rewiring weighted networks in order to randomize connections, while preserving the degree distribution. 4 random nodes are chosen and their inter-connection strengths are modified by a randomly generated number* $\gamma$

- Step 1: choose 4 different arbitrary nodes in the network, call these $A$, $B$, $C$, and $D$. Denote the connection strength $AD = x$, $BC = y$, $AB = w$, and $CD$ = t. If there is no connection between two nodes, then the connection strength is 0.

- Step 2: generate a random number $\gamma$ such that $-1 \leq \gamma \leq 1$. Modify connection strengths among the four nodes as follows: $AD = x + \gamma$, $BC = y + \gamma$, $AB = w - \gamma$, and $CD = t - \gamma$.

- Step 3: check to see if any of connection strength $AB$, $BC$, $AD$, $CD$ has absolute value exceeding 1. If yes, repeat step 2

- Step 4: repeat step $1 \rightarrow 3$ over many iterations (the number of iterations should be at least the number of links in the network)

Figure 2.6 illustrates the outcome of this rewiring strategy after 1,000,000 iterations. It can be seen easily from the algorithm described above that each node maintains its weight degree after each iteration. Thus, this algorithm provides a way to randomize original network without changing its degree distribution. Graph-theoretic measures can then be applied to examine the efficiency of the original network compared to a random network. This method, although perfectly preserves individual degrees, has the weakness of altering the connection strength among nodes in the network. Also, it will be difficult to keep the connection strengths in randomized network to be in the same range as the original network, especially when the functional network becomes sparse due to high correlation threshold.

### 2.3.2 Creating Null Model by Bootstrap Method

We develop another method, called the bootstrap method, to create null model by transforming the original network into an equivalent network of any degree distribution. Furthermore, the equivalent network can be obtained by preserving the individual connection strength in the original network as well. The trade-off, compared to the previous method, is that the result will be approximate, in the sense that the resulting degree sequence will not be 100 percent coincident with a targeted degree sequence. However, this approximation can work quite well over many iterations (despite being more computationally expensive). Let $G = (V, E)$ be the original weighted network. Let $\hat{G}$ denote the (dynamic) synthetic network, $\hat{E}$ be the (dynamic) set of links in $\hat{G}$, and $S = E \setminus \hat{E}$ be the dynamic stack of links that contain all the connections that are in $E$ but not in $\hat{E}$. Initially $\hat{G}, \hat{E} = \emptyset$, and $S = E$. The bootstrap algorithm can be carried out as follows:

- Step 1: Design a target degree sequence in decreasing magnitude $\hat{w_1}, \hat{w_2}, ..., \hat{w_n}$ such that $\sum_{i=1}^{n} \hat{w_i} = \sum_{i=1}^{n} w_i$, with $w_i's$ represent the degree sequence of original network. In other words, design a target degree sequence that preserves the sum of individual degrees of original network. In two special cases, the target degree sequence can be exactly the same as the original degree sequence, or it can follow a scale-free distribution. We discuss how to create a scale-free degree sequence later in this section.

- Step 2: Starting from the highness target degree $\hat{w_1}$ to lowest target degree $\hat{w_n}$, pick random links from the stack $S$, and attach these links to node $i$ in $\hat{G}$, the other ends can be attached to other nodes in $\hat{G}$ at random. At the same time, remove these links from stack $S$. Do this until the constructed degree of node $i$ in $\hat{G}$ is within 0.5 of the target degree $\hat{w_i}$, and then move on to node $i + 1$ in $\hat{G}$. Update the dynamic degree of $\hat{G}$.

- Step 3: At node $i+1$, if the current degree of node $i+1$ in $\hat{G}$ already exceeds $\hat{w_{i+1}}$, start choosing random links attached to nodes $i + 1$ and remove these selected links by throwing them back into the dynamic stack $S$. If not, continue adding links to node $i + 1$ similar to step 2. With either case, stop when the dynamic degree of node $i+1$ in $\hat{G}$ is within 0.5 of the target degree $\hat{w_{i+1}}$. Update the dynamic degree of $\hat{G}$.

- Step 4: After all $n$ nodes have been cycled through, due to the random nature of assigning links to nodes, it should be expected that the dynamic degree of $\hat{G}$ will differ from the target degree sequence $\hat{w_i}$. We then go back to node 1 and repeat step $2 \rightarrow 3$. One iteration is considered complete when all $n$ nodes have been cycled through by operations in step $2 \rightarrow 3$. The algorithm

can terminate when the all the degrees in the dynamic $\hat{G}$ are within 0.5 of the target degree sequence, or the maximum number of iterations has been reached.

It can be seen that when the above algorithm terminates, there may still be some links left over in the dynamic stack $S$, or the degree sequence of $\hat{G}$ may still differ from the target degree sequence $\hat{w}_i$ by an amount greater than 0.5 at some points along the degree sequence. However, this difference is reduced with each iteration. Hence, despite being an approximate method, this strategy can work quite well to achieve a synthetic network with any degree sequence of our choosing.

We now return to the specific question of how to design a scale-free degree sequence $\hat{w}_i$. Note that for any $n$ random numbers $r_1, r_2, ..., r_n$ uniformly distributed on $[0, 1]$, and any given $\alpha$, the series $\{x_i\}$ such that $x_i = (1 - r_i)^{-1/(\alpha-1)}$ are drawn from a scale-free distribution with scaling parameter $\alpha$. Thus, we simply need to rescale $\{x_i\}$ by a constant parameter into $\{\hat{w}_i\}$ so that $\sum_{i=1}^{n} \hat{w}_i = \sum_{i=1}^{n} w_i$. The choice of $\alpha$, however, needs to be coordinated with other topological features of the constructed scale-free network to ensure a randomness factor meaningful enough for our comparison purposes.

Figure 2.7 displays the resulting CDFs of synthetic bootstrap networks versus CDF of original network for subject 34781 with correlation threshold $r = 0.4$. The left hand side chart shows the result for a synthetic network that preserves the degree sequence of the original network. The chart on the right hand side shows the result for a synthetic, scale-free network corresponding to the scaling parameter $\alpha = 5$. Both synthetic networks were constructed using 10 iterations. Although the degree sequence is not perfectly preserved, it can be clearly seen that synthetic networks constructed by the bootstrap method can serve as good null models for our comparison goals. Figures 2.8 and 2.9 display the transformation of FC matrix using the algorithm described above to achieve a random network of the same degree distribution as the original network, and a random network with scale-free degree distribution with $\alpha = 5$, respectively.

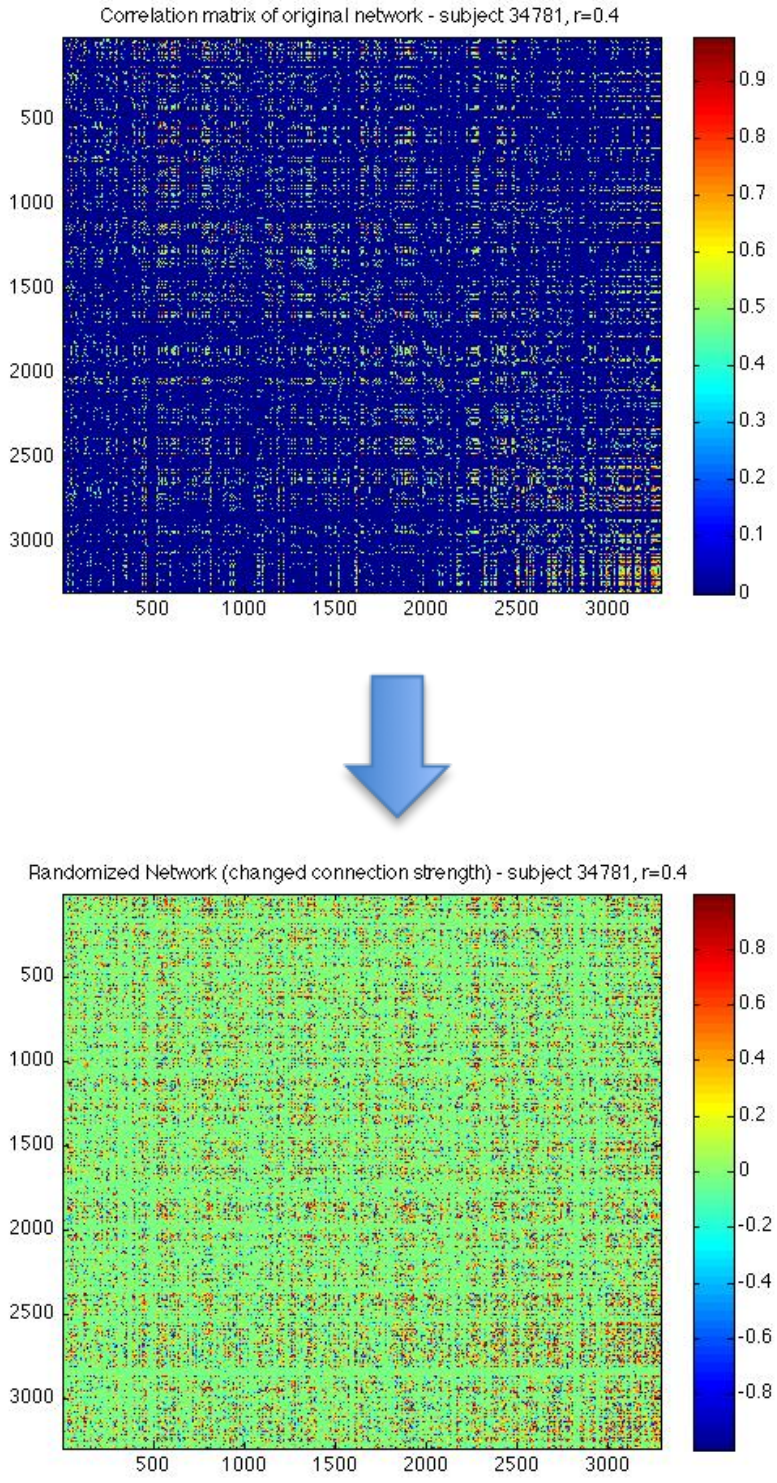



Figure 2.6: *Randomized Network by "Rewiring", subject 34781,* $r = 0.4$

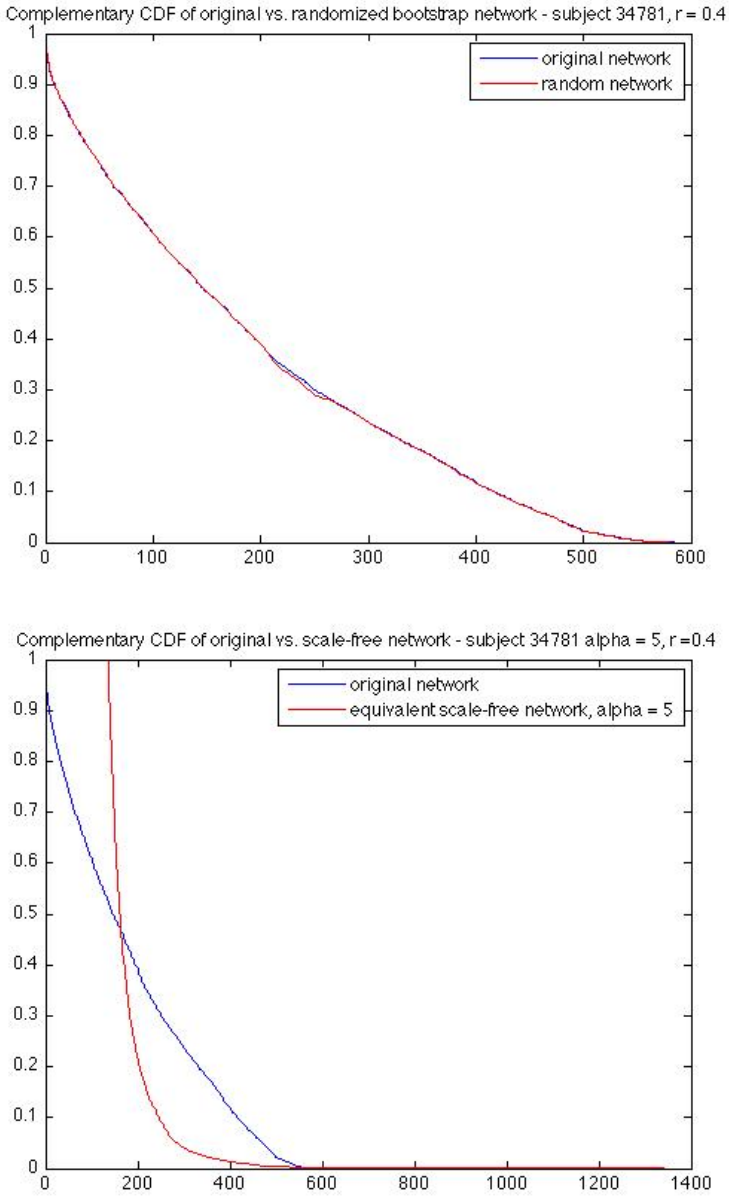



Figure 2.7: *CDF of original vs. synthetic bootstrap networks for constant degree sequence and scale-free degree sequence with* $\alpha = 5$*, respectively, subject 34781,* $r = 0.4$

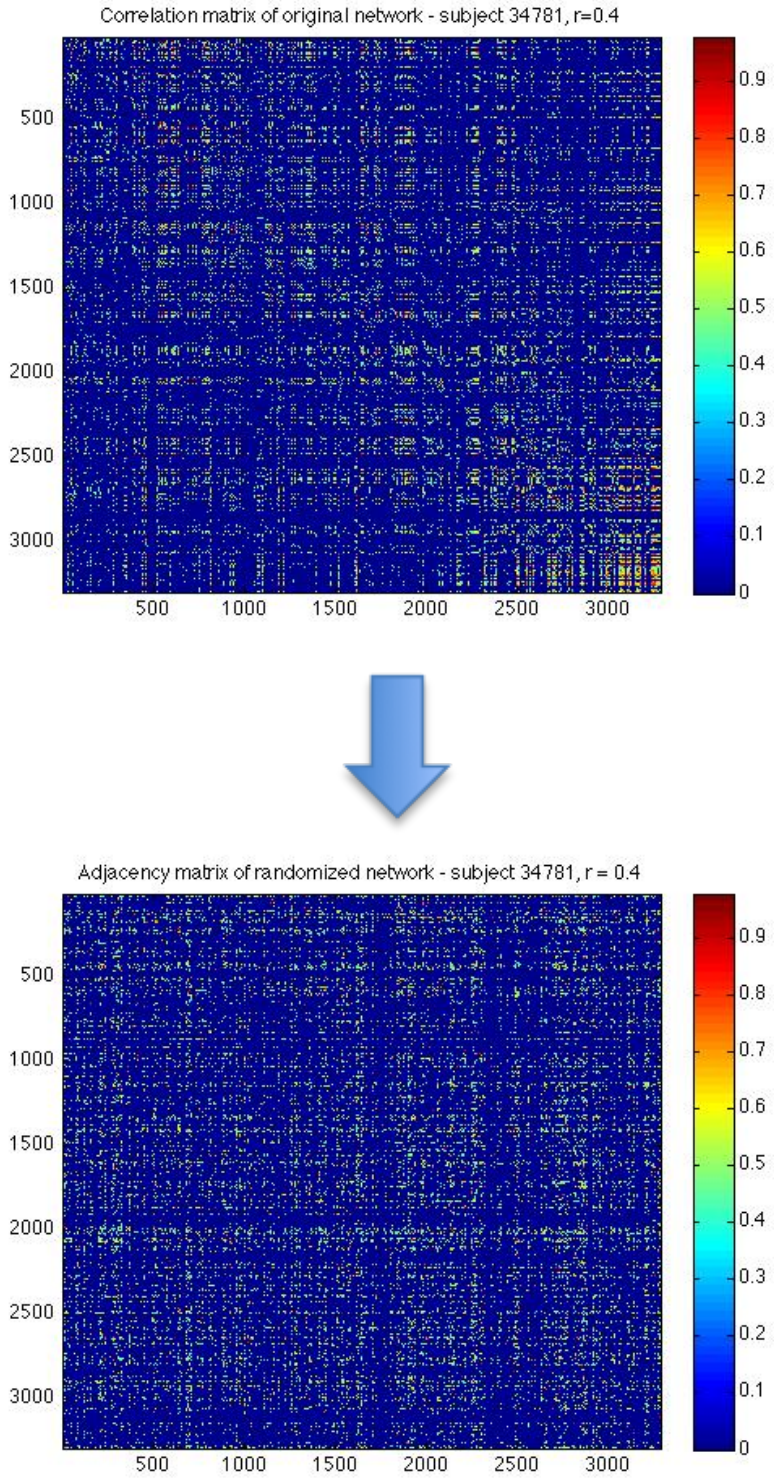



Figure 2.8: *transformation of adjacency matrix that preserves degree distribution using bootstrap method, subject 34781,* $r = 0.4$

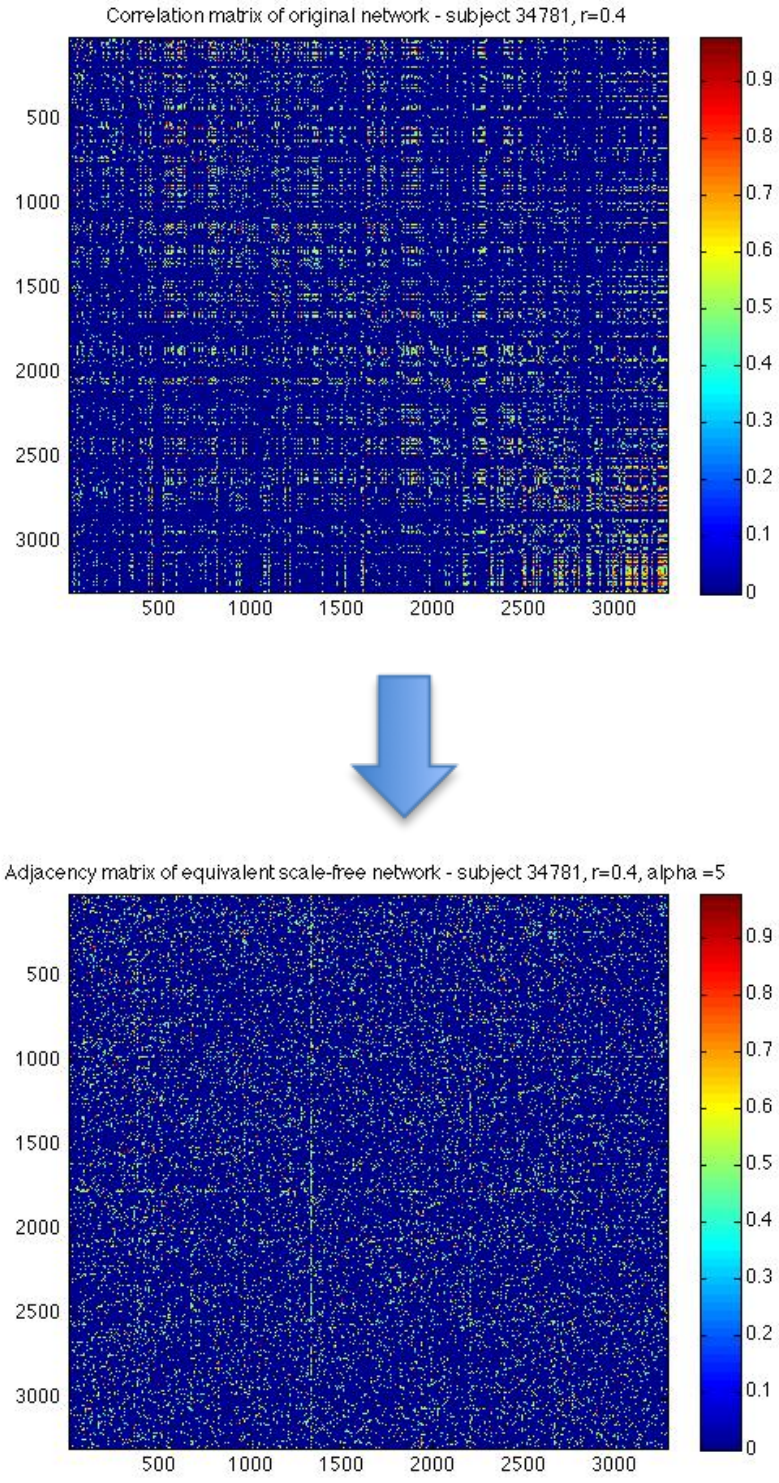



Figure 2.9: *transformation of adjacency matrix using bootstrap method to achieve scale-free degree sequence with* $\alpha = 5$*, subject 34781,* $r = 0.4$

# Chapter 3

# RESULTS

## 3.1 Test Results for Different Families of Distribution

Procedures to test the power-law distribution hypothesis were carried out in Matlab (Mathworks Inc.) with Statistics and Optimization toolboxes.. Each test for each subject was conducted over 17 different thresholds, from $r = -0.7$ to $r = 0.8$, with increment of $0.1$. Note that outside of this range, the constructed network becomes too sparse for meaningful analysis. Per [65], if we wish the calculated p-value to be accurate within about $\epsilon$, then we should generate at least $\frac{1}{4}\epsilon^{-2}$ synthetic data sets. Based on this, the parametric goodness-of-fit test was conducted over 1000 repetitions, ensuring precision of p-value up to 2 decimal digits. Table 3.2 and table 3.3 illustrate the results of power-law distribution test for 10 chosen subjects from the 1000 Functional Connectome Project at resolution level V2. Note that *number of tail data* indicates the number of nodes with the degree exceeding the cut-off point with which we can establish the best possible fit for a given data set. In our implementation, the *number of tail data* is ensured to be greater than 50 data points, and also greater than $5\%$ of the total number of non-zero data points. This is to prevent trivial scenarios where there are too few data points left at the tail, effectively causing the fit to be less reliable. It can be seen that p-values are consistently below $10\%$ across different thresholds, meaning a synthetically generated data set from the estimated $\alpha$ parameter tends to always fit better than the empirical data sets from the KS-statistic point of view. This also holds true with other levels of resolution. This means power-law distribution is not suitable for the distribution of brain functional networks.

In a similar, though less straight-forward fashion, other families of distributions can also be tested against the set of empirical data from our constructed networks. Unlike power-law distribution, solving for the best fit parameters using

maximum likelihood estimation method for other families of distributions typically requires the use of numerical/optimization methods, due to the lack closed-form expression for the maximum likelihood estimation solution. We used a standard numerical method in Matlab (Mathworks Inc.) that finds zeros of functions based on an algorithm originated by T. Dekker ( [74]). Note that this optimization method can significantly increase the computation time, thus fully carrying out all tests at resolution level higher than V2 for all subjects was not always practical, especially for power-law with exponential cutoff, Weibull, and generalized Pareto distribution. However, we conducted the test at resolution level V3 for 2-3 subjects for each distribution to confirm that the results stay consistent across resolutions, which is indeed the case with our 10 subjects. Tables 3.4 and 3.5 report the results for the power-law with exponential cutoff distribution and the exponential distribution tests. Similar to above, p-values are consistently low for both distributions. Tables 3.6 and 3.7 display the results for the log-normal distribution and Weibull distribution tests. Here both distributions exhibit high p-values across different thresholds, though not in all cases, implying that we cannot reject the log-normal and Weibull model for the given empirical data sets. Finally, tables 3.8 and 3.9 display test results for the generalized Pareto distribution. Again, we observe high p-values across different thresholds, and consistently negative shape parameter $k$, which we will discuss in the next chapter.

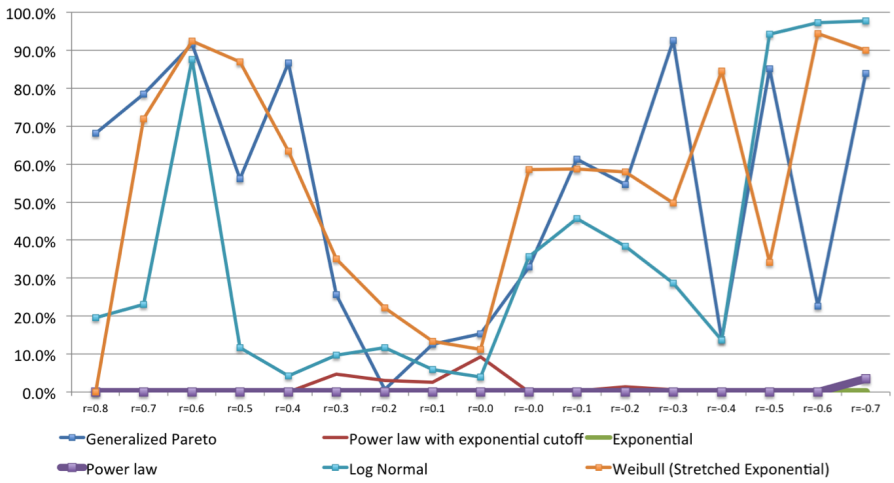



Figure 3.1: *p-values across different thresholds and different distributions for subject 34781*

Figure 3.1 graphically displays the results across different thresholds and distributions for one of the 10 subjects. In aggregate, table 3.1 shows the test statistics for all 10 subjects. A single distribution test is considered ”pass” if corresponding

p-value is greater than 10%. In summary, the power-law, power-law with exponential cutoff and exponential distributions can be rejected due to the consistently low p-values, while other distributions deserve further considerations.

Table 3.1: *Summary of results for 10 subjects from 1000 Functional Connectome project*

| | Distribution Test | | | | | |
|---|---|---|---|---|---|---|
| | generalized Pareto | Power Law with Exp Cutoff | Power Law | Exponential | Log Normal | Weibull |
| % pass out of 170 tests | 65.3% | 12.4% | 0.6% | 0.0% | 50.6% | 69.4% |
| % of tail data versus total data | 35.4% | 49.0% | 15.2% | 86.1% | 22.2% | 27.9% |

## 3.2 Test Results for Model Selection

As discussed in chapter 2, a low p-value can serve as a basis for rejection of certain hypothesized distribution, but a high p-value is not a guarantee for the hypothesized distribution to be the best possible distribution to explain the data. Based on this, it becomes clear from the results presented in the previous section that the power-law model, together with the power-law with exponential cut-off and the exponential model can be rejected as topological model for brain functional networks. The three remaining plausible models are the **log-normal** distribution, the **Weibull** (a.k.a stretched exponential) distribution and the **generalized Pareto** distribution, which, as previously indicated, is the generalized version of both the power-law and the exponential model. The next step is to use **likelihood ratio test** as laid out in chapter 2 to select which model is the most plausible for our various data sets.

One subtlety deserves some mentioning before we present the results. Note that the likelihood ratio test implicitly assumes the two competing models to be applied to the same set of data, thus implying the comparisons of two data sets with the same number of data points. Frequently, however, as can be seen from the results of distribution test, the *number of tail data* of one model for one particular data set, at one particular threshold, is different from that of another model. In order to enable a fair comparison between two different models in this case, we would truncate the model with the ”longer tail” to ensure the equality in tail length.

Table 3.2: *Results of power-law test, 10 subjects from 1000 Functional Connectome project, part 1*

| | Subject 4111 | | | | Subject 4619 | | | |
|---|---|---|---|---|---|---|---|---|
| **Threshold** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** |
| r=0.8 | 2,033 | 7.97 | 127 | 0.0% | 2,011 | 9.16 | 188 | 0.0% |
| r=0.7 | 3,131 | 6.72 | 336 | 0.0% | 3,159 | 7.82 | 438 | 0.0% |
| r=0.6 | 4,101 | 9.22 | 370 | 0.0% | 4,132 | 7.08 | 829 | 0.0% |
| r=0.5 | 4,986 | 11.82 | 413 | 0.0% | 4,931 | 8.43 | 903 | 0.0% |
| r=0.4 | 5,876 | 13.30 | 456 | 0.0% | 5,732 | 8.43 | 1,133 | 0.0% |
| r=0.3 | 7,208 | 13.63 | 589 | 0.0% | 7,012 | 9.79 | 1,108 | 0.0% |
| r=0.2 | 9,457 | 15.28 | 618 | 0.0% | 9,468 | 10.40 | 1,171 | 0.0% |
| r=0.1 | 9,560 | 18.85 | 580 | 0.0% | 9,560 | 10.00 | 1,330 | 0.0% |
| r=0.0 | 9,560 | 19.51 | 625 | 0.0% | 9,560 | 9.76 | 1,409 | 0.0% |
| r=-0.7 | 2,431 | 9.31 | 196 | 0.0% | 2,159 | 8.42 | 201 | 0.0% |
| r=-0.6 | 3,438 | 15.64 | 176 | 0.0% | 2,948 | 7.91 | 409 | 0.0% |
| r=-0.5 | 4,434 | 6.55 | 795 | 0.0% | 3,857 | 8.47 | 529 | 0.0% |
| r=-0.4 | 5,515 | 13.23 | 422 | 0.0% | 4,918 | 9.67 | 578 | 0.0% |
| r=-0.3 | 6,834 | 15.01 | 453 | 0.0% | 6,372 | 10.66 | 608 | 0.0% |
| r=-0.2 | 9,324 | 16.71 | 475 | 0.0% | 9,034 | 11.34 | 639 | 0.0% |
| r=-0.1 | 9,560 | 12.23 | 883 | 0.0% | 9,560 | 11.41 | 697 | 0.0% |
| r=-0.0 | 9,560 | 12.87 | 904 | 0.0% | 9,560 | 11.53 | 710 | 0.0% |

| | Subject 13636 | | | | Subject 13959 | | | |
|---|---|---|---|---|---|---|---|---|
| **Threshold** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** |
| r=0.8 | 1,522 | 11.00 | 95 | 1.0% | 666 | 6.30 | 67 | 0.0% |
| r=0.7 | 2,432 | 7.15 | 273 | 0.0% | 1,311 | 8.06 | 198 | 4.0% |
| r=0.6 | 3,420 | 10.06 | 229 | 0.0% | 2,394 | 3.90 | 568 | 0.0% |
| r=0.5 | 4,545 | 12.77 | 261 | 0.0% | 4,083 | 3.72 | 682 | 0.0% |
| r=0.4 | 5,864 | 4.42 | 1,435 | 0.0% | 6,448 | 4.24 | 642 | 0.0% |
| r=0.3 | 7,269 | 4.66 | 1,661 | 0.0% | 8,408 | 3.42 | 1,302 | 0.0% |
| r=0.2 | 9,399 | 4.82 | 1,874 | 0.0% | 9,539 | 3.39 | 2,850 | 0.0% |
| r=0.1 | 9,560 | 4.83 | 2,175 | 0.0% | 9,560 | 4.92 | 2,108 | 0.0% |
| r=0.0 | 9,560 | 4.98 | 2,268 | 0.0% | 9,560 | 5.41 | 2,158 | 0.0% |
| r=-0.7 | 2,030 | 5.03 | 353 | 0.0% | 595 | 3.03 | 144 | 0.0% |
| r=-0.6 | 3,011 | 4.46 | 641 | 0.0% | 1,409 | 5.42 | 129 | 34.5% |
| r=-0.5 | 4,233 | 4.23 | 987 | 0.0% | 2,650 | 4.27 | 247 | 0.0% |
| r=-0.4 | 5,597 | 4.30 | 1,218 | 0.0% | 4,954 | 4.14 | 389 | 0.0% |
| r=-0.3 | 7,112 | 4.62 | 1,307 | 0.0% | 7,496 | 3.22 | 921 | 0.0% |
| r=-0.2 | 9,334 | 4.76 | 1,522 | 0.0% | 9,466 | 2.79 | 2,253 | 0.0% |
| r=-0.1 | 9,560 | 4.90 | 1,740 | 0.0% | 9,560 | 2.76 | 4,661 | 0.0% |
| r=-0.0 | 9,560 | 5.05 | 1,862 | 0.0% | 9,560 | 2.86 | 5,424 | 0.0% |

| | Subject 18698 | | | | Subject 28433 | | | |
|---|---|---|---|---|---|---|---|---|
| **Threshold** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** |
| r=0.8 | 474 | 1.87 | 210 | 0.0% | 504 | 1.48 | 423 | 0.0% |
| r=0.7 | 1,271 | 1.43 | 1,223 | 0.0% | 1,514 | 1.50 | 1,189 | 0.0% |
| r=0.6 | 2,377 | 6.68 | 132 | 0.0% | 3,376 | 12.36 | 171 | 0.0% |
| r=0.5 | 3,505 | 6.56 | 198 | 0.5% | 5,890 | 4.02 | 617 | 0.0% |
| r=0.4 | 4,429 | 8.15 | 224 | 0.0% | 8,079 | 5.18 | 409 | 0.0% |
| r=0.3 | 5,602 | 4.80 | 876 | 0.0% | 9,326 | 5.09 | 661 | 0.0% |
| r=0.2 | 9,287 | 5.69 | 992 | 0.0% | 9,559 | 5.17 | 945 | 0.0% |
| r=0.1 | 9,560 | 6.84 | 1,013 | 0.0% | 9,560 | 4.60 | 1,641 | 0.0% |
| r=0.0 | 9,560 | 2.18 | 9,488 | 0.0% | 9,560 | 3.73 | 4,896 | 0.0% |
| r=-0.7 | 671 | 3.38 | 103 | 0.0% | 370 | 1.86 | 192 | 0.0% |
| r=-0.6 | 1,407 | 5.20 | 74 | 3.5% | 975 | 5.03 | 87 | 0.0% |
| r=-0.5 | 2,448 | 4.34 | 195 | 0.0% | 1,875 | 5.31 | 158 | 0.0% |
| r=-0.4 | 3,673 | 4.49 | 293 | 0.0% | 3,074 | 5.90 | 216 | 0.0% |
| r=-0.3 | 5,231 | 3.98 | 812 | 0.0% | 4,973 | 3.74 | 730 | 0.0% |
| r=-0.2 | 9,269 | 4.54 | 905 | 0.0% | 8,928 | 4.38 | 703 | 0.0% |
| r=-0.1 | 9,560 | 5.30 | 1,024 | 0.0% | 9,560 | 4.30 | 912 | 0.0% |
| r=-0.0 | 9,560 | 5.96 | 1,041 | 0.0% | 9,560 | 3.62 | 2,316 | 0.0% |

Table 3.3: *Results of power-law test, 10 subjects from 1000 Functional Connectome project, part 2*

| | Subject 30421 | | | | Subject 34781 | | | |
|---|---|---|---|---|---|---|---|---|
| **Threshold** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** |
| r=0.8 | 1,103 | 3.63 | 193 | 0.0% | 1,644 | 5.06 | 163 | 0.0% |
| r=0.7 | 2,110 | 7.30 | 190 | 0.0% | 2,948 | 5.98 | 296 | 0.0% |
| r=0.6 | 3,039 | 11.79 | 201 | 0.0% | 4,185 | 10.40 | 210 | 0.0% |
| r=0.5 | 3,922 | 17.12 | 200 | 0.0% | 5,266 | 7.36 | 529 | 0.0% |
| r=0.4 | 4,747 | 15.25 | 335 | 0.0% | 6,221 | 10.38 | 471 | 0.0% |
| r=0.3 | 5,775 | 16.94 | 292 | 0.0% | 7,291 | 12.79 | 474 | 0.0% |
| r=0.2 | 9,254 | 13.56 | 512 | 0.0% | 9,394 | 14.41 | 506 | 0.0% |
| r=0.1 | 9,560 | 14.81 | 505 | 0.0% | 9,560 | 15.98 | 569 | 0.0% |
| r=0.0 | 9,560 | 15.62 | 530 | 0.0% | 9,560 | 17.50 | 499 | 0.0% |
| r=-0.7 | 1,483 | 13.63 | 76 | 0.5% | 2,273 | 7.41 | 123 | 3.5% |
| r=-0.6 | 2,349 | 8.66 | 221 | 0.0% | 3,464 | 4.93 | 418 | 0.0% |
| r=-0.5 | 3,271 | 13.54 | 209 | 0.0% | 4,747 | 6.47 | 408 | 0.0% |
| r=-0.4 | 4,288 | 10.99 | 365 | 0.0% | 5,972 | 7.29 | 508 | 0.0% |
| r=-0.3 | 5,501 | 13.84 | 277 | 0.0% | 7,192 | 8.72 | 557 | 0.0% |
| r=-0.2 | 9,181 | 9.87 | 588 | 0.0% | 9,333 | 9.64 | 652 | 0.0% |
| r=-0.1 | 9,560 | 11.37 | 544 | 0.0% | 9,560 | 10.81 | 712 | 0.0% |
| r=-0.0 | 9,560 | 12.23 | 540 | 0.0% | 9,560 | 11.93 | 693 | 0.0% |

| | Subject 47659 | | | | Subject 75922 | | | |
|---|---|---|---|---|---|---|---|---|
| **Threshold** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** | **Number of non zeros of original sample** | **alpha (scale parameter)** | **Number of tail data** | **p value** |
| r=0.8 | 1,023 | 1.59 | 611 | 0.0% | 1,537 | 4.09 | 271 | 0.0% |
| r=0.7 | 2,306 | 1.37 | 2,306 | 0.0% | 2,629 | 2.62 | 828 | 0.0% |
| r=0.6 | 3,581 | 8.47 | 193 | 0.0% | 3,577 | 4.45 | 522 | 0.0% |
| r=0.5 | 4,500 | 9.44 | 228 | 0.0% | 4,569 | 4.93 | 686 | 0.0% |
| r=0.4 | 5,277 | 8.63 | 349 | 0.0% | 5,494 | 4.67 | 975 | 0.0% |
| r=0.3 | 6,565 | 8.18 | 633 | 0.0% | 6,666 | 4.27 | 1,344 | 0.0% |
| r=0.2 | 9,455 | 9.46 | 722 | 0.0% | 9,367 | 4.28 | 1,558 | 0.0% |
| r=0.1 | 9,560 | 12.47 | 660 | 0.0% | 9,560 | 4.33 | 1,990 | 0.0% |
| r=0.0 | 9,560 | 14.85 | 575 | 0.0% | 9,560 | 4.50 | 2,212 | 0.0% |
| r=-0.7 | 1,046 | 5.47 | 85 | 6.5% | 1,595 | 4.29 | 258 | 0.0% |
| r=-0.6 | 1,939 | 5.70 | 149 | 0.0% | 2,518 | 6.49 | 248 | 0.0% |
| r=-0.5 | 3,189 | 7.12 | 183 | 0.0% | 3,581 | 7.20 | 351 | 0.0% |
| r=-0.4 | 4,634 | 7.28 | 233 | 0.0% | 4,855 | 6.55 | 486 | 0.0% |
| r=-0.3 | 6,359 | 5.52 | 603 | 0.0% | 6,328 | 5.44 | 842 | 0.0% |
| r=-0.2 | 9,326 | 6.65 | 597 | 0.0% | 9,235 | 5.13 | 1,132 | 0.0% |
| r=-0.1 | 9,560 | 8.64 | 574 | 0.0% | 9,560 | 5.26 | 1,278 | 0.0% |
| r=-0.0 | 9,560 | 9.82 | 586 | 0.0% | 9,560 | 5.40 | 1,403 | 0.0% |

Table 3.4: *Results of power-law with exponential cutoff and exponential distribution tests, 10 subjects from 1000 Functional Connectome project, part 1*

| | Subject 4111 | | | | Subject 4619 | | | |
|---|---|---|---|---|---|---|---|---|
| | Power Law with exp. cutoff | | Exponential | | Power Law with exp. cutoff | | Exponential | |
| Threshold | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value |
| r=0.8 | 1,410 | 0.0% | 1,583 | 0.0% | 194 | 5.3% | 1,704 | 0.0% |
| r=0.7 | 2,359 | 0.0% | 2,660 | 0.0% | 526 | 0.0% | 2,742 | 0.0% |
| r=0.6 | 3,278 | 0.0% | 3,632 | 0.0% | 659 | 0.7% | 3,741 | 0.0% |
| r=0.5 | 627 | 0.3% | 4,675 | 0.0% | 925 | 0.3% | 4,839 | 0.0% |
| r=0.4 | 919 | 0.7% | 5,736 | 0.0% | 1,173 | 0.7% | 5,732 | 0.0% |
| r=0.3 | 603 | 5.0% | 6,885 | 0.0% | 1,147 | 3.0% | 6,861 | 0.0% |
| r=0.2 | 608 | 3.0% | 8,462 | 0.0% | 1,296 | 0.0% | 8,125 | 0.0% |
| r=0.1 | 9,557 | 0.0% | 9,560 | 0.0% | 9,498 | 0.0% | 9,560 | 0.0% |
| r=0.0 | 9,491 | 0.0% | 9,560 | 0.0% | 9,380 | 0.0% | 9,560 | 0.0% |
| r=-0.7 | 1,950 | 0.0% | 2,169 | 0.0% | 216 | 10.7% | 2,143 | 0.0% |
| r=-0.6 | 174 | 22.7% | 3,118 | 0.0% | 377 | 1.0% | 2,948 | 0.0% |
| r=-0.5 | 3,772 | 0.0% | 4,179 | 0.0% | 579 | 0.0% | 3,857 | 0.0% |
| r=-0.4 | 4,845 | 0.0% | 5,249 | 0.0% | 614 | 1.7% | 4,752 | 0.0% |
| r=-0.3 | 459 | 2.7% | 6,401 | 0.0% | 624 | 1.3% | 5,819 | 0.0% |
| r=-0.2 | 7,369 | 0.0% | 7,897 | 0.0% | 642 | 1.3% | 7,220 | 0.0% |
| r=-0.1 | 9,557 | 0.0% | 9,560 | 0.0% | 704 | 3.7% | 9,560 | 0.0% |
| r=-0.0 | 9,491 | 0.0% | 9,560 | 0.0% | 9,558 | 0.0% | 9,560 | 0.0% |

| | Subject 13636 | | | | Subject 13959 | | | |
|---|---|---|---|---|---|---|---|---|
| | Power Law with exp. cutoff | | Exponential | | Power Law with exp. cutoff | | Exponential | |
| Threshold | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value |
| r=0.8 | 1,200 | 0.0% | 1,339 | 0.0% | 483 | 0.0% | 565 | 0.0% |
| r=0.7 | 272 | 11.0% | 2,188 | 0.0% | 207 | 69.0% | 1,006 | 0.0% |
| r=0.6 | 730 | 0.7% | 3,035 | 0.0% | 591 | 0.3% | 1,614 | 0.0% |
| r=0.5 | 1,218 | 0.0% | 4,019 | 0.0% | 719 | 4.0% | 2,457 | 0.0% |
| r=0.4 | 1,592 | 0.0% | 5,140 | 0.0% | 833 | 7.0% | 4,224 | 0.0% |
| r=0.3 | 1,806 | 0.3% | 6,609 | 0.0% | 5,291 | 3.7% | 6,756 | 0.0% |
| r=0.2 | 7,605 | 0.0% | 8,175 | 0.0% | 3,917 | 0.0% | 9,377 | 0.0% |
| r=0.1 | 9,531 | 0.0% | 9,560 | 0.0% | 2,561 | 0.0% | 9,560 | 0.0% |
| r=0.0 | 9,162 | 0.0% | 9,560 | 0.0% | 2,566 | 0.0% | 9,560 | 0.0% |
| r=-0.7 | 398 | 26.0% | 1,948 | 0.0% | 173 | 55.0% | 476 | 0.1% |
| r=-0.6 | 762 | 0.7% | 2,745 | 0.0% | 129 | 67.0% | 990 | 0.0% |
| r=-0.5 | 1,142 | 0.0% | 3,731 | 0.0% | 2,065 | 0.0% | 1,983 | 0.0% |
| r=-0.4 | 1,413 | 0.0% | 4,832 | 0.0% | 2,280 | 9.7% | 3,442 | 0.0% |
| r=-0.3 | 1,535 | 0.0% | 6,272 | 0.0% | 4,097 | 18.0% | 5,871 | 0.0% |
| r=-0.2 | 1,647 | 0.0% | 8,016 | 0.0% | 4,825 | 0.0% | 8,198 | 0.0% |
| r=-0.1 | 2,053 | 0.0% | 9,560 | 0.0% | 5,707 | 0.0% | 9,560 | 0.0% |
| r=-0.0 | 9,511 | 0.0% | 9,560 | 0.0% | 6,252 | 0.0% | 9,560 | 0.0% |

| | Subject 18698 | | | | Subject 28433 | | | |
|---|---|---|---|---|---|---|---|---|
| | Power Law with exp. cutoff | | Exponential | | Power Law with exp. cutoff | | Exponential | |
| Threshold | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value |
| r=0.8 | 328 | 0.3% | 357 | 0.0% | 384 | 0.0% | 245 | 0.0% |
| r=0.7 | 1,088 | 0.0% | 778 | 0.0% | 1,309 | 0.3% | 833 | 0.0% |
| r=0.6 | 2,122 | 0.0% | 1,490 | 0.0% | 2,966 | 0.0% | 2,106 | 0.0% |
| r=0.5 | 276 | 71.0% | 2,463 | 0.0% | 5,395 | 0.0% | 3,729 | 0.0% |
| r=0.4 | 2,921 | 0.0% | 3,579 | 0.0% | 7,408 | 0.0% | 5,494 | 0.0% |
| r=0.3 | 4,273 | 0.0% | 4,805 | 0.0% | 7,944 | 0.0% | 7,069 | 0.0% |
| r=0.2 | 5,805 | 0.0% | 6,371 | 0.0% | 8,894 | 0.0% | 8,729 | 0.0% |
| r=0.1 | 7,767 | 0.0% | 9,560 | 0.0% | 8,852 | 0.0% | 9,560 | 0.0% |
| r=0.0 | 9,540 | 0.0% | 9,560 | 0.0% | 6,766 | 1.0% | 9,560 | 0.0% |
| r=-0.7 | 400 | 34.7% | 528 | 0.0% | 313 | 1.7% | 302 | 0.1% |
| r=-0.6 | 850 | 48.0% | 1,079 | 0.0% | 449 | 17.0% | 700 | 0.0% |
| r=-0.5 | 603 | 52.3% | 1,855 | 0.0% | 1,383 | 0.0% | 1,292 | 0.0% |
| r=-0.4 | 768 | 16.3% | 2,792 | 0.0% | 1,813 | 0.0% | 2,175 | 0.0% |
| r=-0.3 | 898 | 41.3% | 4,047 | 0.0% | 2,731 | 0.0% | 3,260 | 0.0% |
| r=-0.2 | 941 | 10.3% | 5,882 | 0.0% | 976 | 3.7% | 5,057 | 0.0% |
| r=-0.1 | 1,041 | 12.3% | 9,560 | 0.0% | 5,094 | 0.0% | 9,560 | 0.0% |
| r=-0.0 | 9,361 | 0.0% | 9,560 | 0.0% | 4,725 | 25.0% | 9,560 | 0.0% |

Table 3.5: *Results of power-law with exponential cutoff and exponential distribution tests, 10 subjects from 1000 Functional Connectome project, part 2*

| | Subject 30421 | | | | Subject 34781 | | | |
|---|---|---|---|---|---|---|---|---|
| | Power Law with exp. cutoff | | Exponential | | Power Law with exp. cutoff | | Exponential | |
| Threshold | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value |
| r=0.8 | 226 | 11.7% | 796 | 0.0% | 1,359 | 0.0% | 1,279 | 0.0% |
| r=0.7 | 1,270 | 0.0% | 1,545 | 0.0% | 2,764 | 0.0% | 2,300 | 0.0% |
| r=0.6 | 201 | 2.7% | 2,395 | 0.0% | 3,526 | 0.0% | 3,434 | 0.0% |
| r=0.5 | 444 | 2.0% | 3,329 | 0.0% | 4,254 | 0.0% | 4,678 | 0.0% |
| r=0.4 | 356 | 3.0% | 4,378 | 0.0% | 5,400 | 0.0% | 5,948 | 0.0% |
| r=0.3 | 295 | 8.3% | 5,538 | 0.0% | 480 | 4.7% | 7,249 | 0.0% |
| r=0.2 | 538 | 3.7% | 6,871 | 0.0% | 526 | 3.0% | 8,790 | 0.0% |
| r=0.1 | 528 | 6.3% | 9,560 | 0.0% | 591 | 2.7% | 9,560 | 0.0% |
| r=0.0 | 9,556 | 0.0% | 9,560 | 0.0% | 556 | 9.3% | 9,560 | 0.0% |
| r=-0.7 | 1,159 | 0.0% | 1,291 | 0.0% | 2,189 | 0.0% | 1,874 | 0.0% |
| r=-0.6 | 1,849 | 0.0% | 2,059 | 0.0% | 2,853 | 0.0% | 2,930 | 0.0% |
| r=-0.5 | 417 | 0.7% | 2,882 | 0.0% | 3,785 | 0.0% | 4,134 | 0.0% |
| r=-0.4 | 525 | 1.0% | 3,818 | 0.0% | 5,111 | 0.0% | 5,499 | 0.0% |
| r=-0.3 | 306 | 5.0% | 5,005 | 0.0% | 643 | 0.7% | 6,874 | 0.0% |
| r=-0.2 | 723 | 0.0% | 6,418 | 0.0% | 569 | 1.3% | 8,630 | 0.0% |
| r=-0.1 | 6,727 | 0.0% | 9,560 | 0.0% | 2,019 | 0.3% | 9,560 | 0.0% |
| r=-0.0 | 9,546 | 0.0% | 9,560 | 0.0% | 2,369 | 0.0% | 9,560 | 0.0% |

| | Subject 47659 | | | | Subject 75922 | | | |
|---|---|---|---|---|---|---|---|---|
| | Power Law with exp. cutoff | | Exponential | | Power Law with exp. cutoff | | Exponential | |
| Threshold | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value | Number of tail data | p value |
| r=0.8 | 715 | 0.0% | 646 | 0.0% | 310 | 7.3% | 1,086 | 0.0% |
| r=0.7 | 2,079 | 0.0% | 1,432 | 0.0% | 835 | 0.0% | 1,881 | 0.0% |
| r=0.6 | 3,437 | 0.0% | 2,526 | 0.0% | 2,190 | 0.0% | 2,709 | 0.0% |
| r=0.5 | 4,352 | 0.0% | 3,827 | 0.0% | 4,386 | 0.0% | 3,674 | 0.0% |
| r=0.4 | 4,503 | 0.0% | 4,857 | 0.0% | 1,169 | 0.0% | 4,813 | 0.0% |
| r=0.3 | 638 | 1.3% | 5,936 | 0.0% | 5,247 | 0.0% | 6,051 | 0.0% |
| r=0.2 | 803 | 1.3% | 7,517 | 0.0% | 6,949 | 0.0% | 7,632 | 0.0% |
| r=0.1 | 659 | 2.3% | 9,560 | 0.0% | 9,551 | 0.0% | 9,560 | 0.0% |
| r=0.0 | 585 | 3.7% | 9,560 | 0.0% | 6,675 | 0.0% | 9,560 | 0.0% |
| r=-0.7 | 872 | 0.0% | 838 | 0.1% | 284 | 8.7% | 1,248 | 0.0% |
| r=-0.6 | 1,164 | 0.0% | 1,457 | 0.0% | 344 | 13.3% | 2,040 | 0.0% |
| r=-0.5 | 1,908 | 0.0% | 2,255 | 0.0% | 361 | 13.3% | 2,956 | 0.0% |
| r=-0.4 | 2,533 | 0.0% | 3,327 | 0.0% | 706 | 1.0% | 4,020 | 0.0% |
| r=-0.3 | 579 | 2.0% | 4,885 | 0.0% | 907 | 0.0% | 5,300 | 0.0% |
| r=-0.2 | 4,749 | 0.0% | 7,086 | 0.0% | 6,732 | 0.0% | 7,023 | 0.0% |
| r=-0.1 | 7,990 | 0.0% | 9,560 | 0.0% | 8,273 | 0.0% | 9,560 | 0.0% |
| r=-0.0 | 8,035 | 0.0% | 9,560 | 0.0% | 9,221 | 0.0% | 9,560 | 0.0% |

Implementing the *log likelihood ratio* for the generalized Pareto model versus the log normal model, and then for the generalized Pareto model versus the Weibull model across all data sets presented in the previous section (170 different tests), the results are as follows:

- The log likelihood of the generalized Pareto model is greater than the log likelihood of the Weibull model for 168 out of 170 tests, 154 of which are significant (meaning the probability that the observed positive sign of the difference between the log likelihood being a chance result of fluctuations is less than 10%)
- The log likelihood of the generalized Pareto model is greater than the log likelihood of the log normal model for 170 out of 170 tests, 164 of which are significant

These results clearly demonstrate that the generalized Pareto model is the best model among the popular models in consideration.

Table 3.6: *Results of log-normal and Weibull distribution tests, 10 subjects from 1000 Functional Connectome project, part 1*

| | Subject 4111 | | | | | | Subject 4619 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Log Normal | | | Weibull | | | Log Normal | | | Weibull | | |
| Threshold | Number data | of tail | p value | Number data | of tail | p value | Number data | of tail | p value | Number data | of tail | p value |
| r=0.8 | | 232 | 25.0% | | 321 | 50.5% | | 430 | 25.7% | | 478 | 71.7% |
| r=0.7 | | 176 | 84.3% | | 919 | 1.9% | | 975 | 0.7% | | 1,280 | 6.6% |
| r=0.6 | | 717 | 2.3% | | 1,243 | 6.6% | | 1,396 | 0.0% | | 1,562 | 0.0% |
| r=0.5 | | 1,161 | 0.3% | | 1,227 | 10.4% | | 1,728 | 0.0% | | 2,146 | 0.0% |
| r=0.4 | | 1,146 | 0.0% | | 1,285 | 10.0% | | 2,123 | 0.0% | | 2,497 | 0.0% |
| r=0.3 | | 1,235 | 2.0% | | 1,298 | 11.9% | | 594 | 1.0% | | 2,375 | 0.0% |
| r=0.2 | | 1,383 | 0.0% | | 1,337 | 7.7% | | 2,424 | 0.0% | | 633 | 17.4% |
| r=0.1 | | 1,350 | 0.3% | | 1,355 | 4.2% | | 2,595 | 0.0% | | 683 | 15.8% |
| r=0.0 | | 1,317 | 0.3% | | 1,460 | 6.6% | | 2,757 | 0.0% | | 688 | 20.6% |
| r=-0.7 | | 345 | 74.3% | | 421 | 92.0% | | 535 | 70.7% | | 590 | 80.7% |
| r=-0.6 | | 218 | 66.3% | | 339 | 54.2% | | 961 | 9.3% | | 1,219 | 34.5% |
| r=-0.5 | | 239 | 86.7% | | 454 | 32.6% | | 1,258 | 13.0% | | 1,416 | 13.3% |
| r=-0.4 | | 931 | 0.3% | | 1,105 | 0.2% | | 1,375 | 1.0% | | 253 | 99.0% |
| r=-0.3 | | 1,010 | 0.0% | | 1,412 | 0.2% | | 1,399 | 3.0% | | 1,917 | 6.1% |
| r=-0.2 | | 992 | 0.0% | | 1,275 | 0.9% | | 1,430 | 2.3% | | 1,971 | 9.6% |
| r=-0.1 | | 839 | 0.0% | | 1,178 | 1.0% | | 1,735 | 1.3% | | 2,496 | 11.9% |
| r=-0.0 | | 824 | 0.3% | | 1,166 | 3.2% | | 1,890 | 2.0% | | 2,594 | 7.4% |

| | Subject 13636 | | | | | | Subject 13959 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Log Normal | | | Weibull | | | Log Normal | | | Weibull | | |
| Threshold | Number data | of tail | p value | Number data | of tail | p value | Number data | of tail | p value | Number data | of tail | p value |
| r=0.8 | | 124 | 74.7% | | 562 | 2.7% | | 299 | 3.3% | | 372 | 27.1% |
| r=0.7 | | 697 | 72.3% | | 840 | 24.9% | | 269 | 71.3% | | 269 | 56.7% |
| r=0.6 | | 1,193 | 0.3% | | 1,423 | 5.8% | | 223 | 88.7% | | 301 | 99.9% |
| r=0.5 | | 526 | 6.0% | | 539 | 18.5% | | 1,071 | 14.3% | | 1,194 | 14.0% |
| r=0.4 | | 533 | 12.3% | | 736 | 25.6% | | 1,280 | 6.7% | | 1,637 | 0.4% |
| r=0.3 | | 619 | 9.3% | | 744 | 31.6% | | 3,172 | 1.0% | | 4,123 | 38.2% |
| r=0.2 | | 693 | 5.0% | | 701 | 27.1% | | 5,875 | 0.0% | | 7,149 | 0.0% |
| r=0.1 | | 679 | 3.3% | | 677 | 22.3% | | 5,813 | 0.7% | | 8,234 | 5.3% |
| r=0.0 | | 664 | 5.3% | | 641 | 16.9% | | 5,537 | 0.0% | | 9,551 | 5.9% |
| r=-0.7 | | 740 | 34.3% | | 1,146 | 61.8% | | 225 | 71.0% | | 294 | 74.6% |
| r=-0.6 | | 1,400 | 4.0% | | 1,721 | 9.7% | | 275 | 28.0% | | 511 | 26.4% |
| r=-0.5 | | 1,971 | 0.3% | | 2,384 | 0.2% | | 616 | 72.3% | | 653 | 95.3% |
| r=-0.4 | | 2,334 | 0.0% | | 2,998 | 0.1% | | 809 | 8.7% | | 1,979 | 18.3% |
| r=-0.3 | | 2,836 | 0.0% | | 3,645 | 0.0% | | 2,435 | 0.7% | | 5,354 | 28.2% |
| r=-0.2 | | 3,326 | 0.0% | | 4,313 | 0.0% | | 6,502 | 0.0% | | 5,013 | 0.0% |
| r=-0.1 | | 3,984 | 0.0% | | 5,029 | 0.0% | | 8,497 | 0.0% | | 974 | 14.3% |
| r=-0.0 | | 4,353 | 0.0% | | 5,096 | 0.0% | | 6,445 | 0.0% | | 982 | 12.1% |

| | Subject 18698 | | | | | | Subject 28433 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Log Normal | | | Weibull | | | Log Normal | | | Weibull | | |
| Threshold | Number data | of tail | p value | Number data | of tail | p value | Number data | of tail | p value | Number data | of tail | p value |
| r=0.8 | | 57 | 94.4% | | 48 | 93.2% | | 83 | 54.6% | | 95 | 90.7% |
| r=0.7 | | 128 | 77.8% | | 203 | 24.1% | | 85 | 82.9% | | 116 | 85.3% |
| r=0.6 | | 142 | 99.0% | | 655 | 12.6% | | 2,936 | 0.0% | | 2,849 | 0.0% |
| r=0.5 | | 412 | 41.6% | | 729 | 37.3% | | 371 | 11.7% | | 318 | 62.3% |
| r=0.4 | | 264 | 85.4% | | 688 | 77.1% | | 575 | 87.1% | | 681 | 75.1% |
| r=0.3 | | 429 | 76.6% | | 468 | 98.3% | | 1,407 | 12.9% | | 1,450 | 58.1% |
| r=0.2 | | 473 | 61.6% | | 564 | 96.3% | | 2,042 | 5.1% | | 2,313 | 12.2% |
| r=0.1 | | 483 | 31.2% | | 544 | 97.3% | | 3,274 | 0.0% | | 3,352 | 0.2% |
| r=0.0 | | 479 | 32.2% | | 548 | 88.0% | | 8,386 | 0.0% | | 7,367 | 0.0% |
| r=-0.7 | | 217 | 39.2% | | 406 | 23.6% | | 47 | 87.2% | | 205 | 12.0% |
| r=-0.6 | | 226 | 84.0% | | 842 | 37.5% | | 168 | 44.9% | | 430 | 13.0% |
| r=-0.5 | | 922 | 36.7% | | 1,053 | 69.9% | | 378 | 39.6% | | 444 | 83.7% |
| r=-0.4 | | 1,130 | 63.4% | | 1,425 | 70.0% | | 463 | 39.9% | | 523 | 77.5% |
| r=-0.3 | | 1,592 | 35.3% | | 1,966 | 50.0% | | 1,671 | 1.6% | | 2,085 | 29.3% |
| r=-0.2 | | 1,721 | 21.3% | | 2,115 | 65.4% | | 2,067 | 10.2% | | 2,671 | 20.2% |
| r=-0.1 | | 1,815 | 12.8% | | 3,224 | 46.8% | | 3,278 | 4.9% | | 4,099 | 6.6% |
| r=-0.0 | | 2,262 | 23.3% | | 3,232 | 65.1% | | 4,301 | 0.8% | | 4,598 | 1.2% |

Table 3.7: *Results of log-normal and Weibull distribution tests, 10 subjects from 1000 Functional Connectome project, part 2*

| | Subject 30421 | | | | | | Subject 34781 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Log Normal | | | Weibull | | | Log Normal | | | Weibull | | |
| Threshold | Number of data | tail | p value | Number of data | tail | p value | Number of data | tail | p value | Number of data | tail | p value |
| r=0.8 | | 292 | 52.2% | | 306 | 53.5% | | 277 | 19.6% | | 1,142 | 0.0% |
| r=0.7 | | 254 | 58.8% | | 289 | 88.8% | | 389 | 23.1% | | 425 | 71.9% |
| r=0.6 | | 277 | 88.0% | | 287 | 97.5% | | 369 | 87.5% | | 559 | 92.4% |
| r=0.5 | | 560 | 8.8% | | 599 | 34.8% | | 730 | 11.7% | | 804 | 86.9% |
| r=0.4 | | 662 | 26.6% | | 662 | 87.9% | | 954 | 4.2% | | 938 | 63.5% |
| r=0.3 | | 687 | 41.6% | | 770 | 97.8% | | 1,018 | 9.8% | | 1,051 | 35.0% |
| r=0.2 | | 1,020 | 47.2% | | 1,111 | 87.4% | | 1,040 | 11.7% | | 1,076 | 22.2% |
| r=0.1 | | 1,059 | 27.3% | | 1,398 | 70.7% | | 1,043 | 6.0% | | 1,055 | 13.4% |
| r=0.0 | | 1,464 | 8.8% | | 1,671 | 51.4% | | 1,053 | 3.9% | | 1,137 | 11.2% |
| r=-0.7 | | 112 | 99.7% | | 707 | 0.4% | | 205 | 97.7% | | 232 | 89.9% |
| r=-0.6 | | 802 | 0.0% | | 994 | 0.2% | | 225 | 97.2% | | 274 | 94.3% |
| r=-0.5 | | 546 | 2.1% | | 636 | 16.2% | | 257 | 94.2% | | 1,088 | 34.2% |
| r=-0.4 | | 741 | 1.3% | | 649 | 25.6% | | 1,050 | 13.6% | | 1,095 | 84.5% |
| r=-0.3 | | 487 | 71.4% | | 910 | 29.2% | | 1,017 | 28.7% | | 1,162 | 49.8% |
| r=-0.2 | | 819 | 26.1% | | 1,575 | 16.5% | | 1,024 | 38.4% | | 1,110 | 57.9% |
| r=-0.1 | | 1,331 | 5.0% | | 1,545 | 56.7% | | 1,052 | 45.7% | | 1,183 | 58.7% |
| r=-0.0 | | 1,356 | 6.4% | | 1,535 | 37.4% | | 1,048 | 35.7% | | 1,231 | 58.6% |

| | Subject 47659 | | | | | | Subject 75922 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Log Normal | | | Weibull | | | Log Normal | | | Weibull | | |
| Threshold | Number of data | tail | p value | Number of data | tail | p value | Number of data | tail | p value | Number of data | tail | p value |
| r=0.8 | | 68 | 94.8% | | 374 | 4.1% | | 467 | 0.8% | | 549 | 0.2% |
| r=0.7 | | 428 | 0.1% | | 1,226 | 0.0% | | 1,060 | 0.0% | | 510 | 20.7% |
| r=0.6 | | 362 | 1.9% | | 421 | 15.5% | | 274 | 32.7% | | 294 | 45.8% |
| r=0.5 | | 451 | 85.2% | | 387 | 90.4% | | 1,453 | 0.0% | | 280 | 76.5% |
| r=0.4 | | 751 | 11.7% | | 921 | 50.5% | | 1,893 | 0.0% | | 287 | 84.6% |
| r=0.3 | | 1,509 | 1.3% | | 1,640 | 13.6% | | 2,508 | 0.0% | | 3,009 | 0.0% |
| r=0.2 | | 1,675 | 4.8% | | 2,637 | 6.5% | | 3,358 | 0.0% | | 3,766 | 0.0% |
| r=0.1 | | 1,396 | 18.3% | | 1,835 | 44.3% | | 4,280 | 0.0% | | 5,464 | 0.0% |
| r=0.0 | | 1,397 | 51.0% | | 2,928 | 59.3% | | 4,917 | 0.0% | | 7,685 | 0.0% |
| r=-0.7 | | 143 | 98.7% | | 248 | 95.5% | | 515 | 27.8% | | 619 | 48.8% |
| r=-0.6 | | 298 | 49.4% | | 559 | 56.0% | | 716 | 39.9% | | 888 | 44.3% |
| r=-0.5 | | 305 | 38.5% | | 616 | 56.4% | | 824 | 50.8% | | 1,044 | 51.7% |
| r=-0.4 | | 317 | 92.3% | | 676 | 89.3% | | 1,288 | 46.0% | | 1,409 | 60.8% |
| r=-0.3 | | 387 | 95.1% | | 1,496 | 30.6% | | 1,810 | 1.0% | | 2,024 | 9.4% |
| r=-0.2 | | 1,217 | 22.6% | | 1,422 | 66.7% | | 2,139 | 0.0% | | 2,450 | 0.2% |
| r=-0.1 | | 1,219 | 24.2% | | 1,183 | 58.7% | | 2,390 | 0.0% | | 2,643 | 0.0% |
| r=-0.0 | | 1,153 | 17.5% | | 1,231 | 58.6% | | 2,710 | 0.0% | | 9,350 | 0.0% |

Table 3.8: *Results of generalized Pareto distribution tests, 10 subjects from 1000 Functional Connectome project, part 1*

| | Subject 4111 - Generalized Pareto | | | | Subject 4619 - Generalized Pareto | | | |
|---|---|---|---|---|---|---|---|---|
| Threshold | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value |
| r=0.8 | -0.40 | 119.9 | 894 | 35.1% | -0.59 | 312.8 | 1,335 | 5.7% |
| r=0.7 | -0.52 | 259.8 | 1,532 | 50.9% | -0.61 | 278.7 | 1,019 | 22.0% |
| r=0.6 | -0.52 | 152.5 | 568 | 97.6% | -0.70 | 187.1 | 512 | 96.3% |
| r=0.5 | -0.52 | 139.2 | 537 | 91.5% | -0.70 | 325.2 | 1,312 | 26.7% |
| r=0.4 | -0.53 | 191.9 | 943 | 86.2% | -0.70 | 317.6 | 1,363 | 76.7% |
| r=0.3 | -0.55 | 152.4 | 635 | 99.9% | -0.68 | 270.1 | 1,180 | 96.3% |
| r=0.2 | -0.55 | 177.4 | 916 | 96.2% | -0.66 | 315.6 | 1,529 | 30.0% |
| r=0.1 | -0.75 | 936.6 | 5,980 | 2.7% | -0.79 | 1254.5 | 6,605 | 1.0% |
| r=0.0 | -0.74 | 811.5 | 5,429 | 5.5% | -0.75 | 1104.3 | 6,330 | 1.7% |
| r=-0.7 | -0.50 | 248.0 | 1,412 | 10.8% | -0.46 | 221.5 | 762 | 34.0% |
| r=-0.6 | -0.58 | 420.5 | 2,234 | 4.3% | -0.48 | 231.8 | 812 | 25.3% |
| r=-0.5 | -0.63 | 565.9 | 2,970 | 0.0% | -0.68 | 185.1 | 275 | 87.3% |
| r=-0.4 | -0.67 | 693.6 | 3,591 | 0.2% | -0.56 | 375.3 | 1,582 | 0.7% |
| r=-0.3 | -0.68 | 789.8 | 4,348 | 0.3% | -0.55 | 388.5 | 1,758 | 1.0% |
| r=-0.2 | -0.69 | 792.1 | 4,543 | 0.0% | -0.51 | 308.1 | 1,378 | 8.3% |
| r=-0.1 | -0.69 | 740.5 | 4,472 | 1.0% | -0.50 | 340.6 | 1,613 | 7.7% |
| r=-0.0 | -0.69 | 665.6 | 4,113 | 1.2% | -0.49 | 330.7 | 1,587 | 4.0% |

| | Subject 13636 - Generalized Pareto | | | | Subject 13959 - Generalized Pareto | | | |
|---|---|---|---|---|---|---|---|---|
| Threshold | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value |
| r=0.8 | -0.51 | 162.9 | 787 | 91.2% | -0.46 | 52.7 | 402 | 85.0% |
| r=0.7 | -0.53 | 344.7 | 1,666 | 0.1% | -0.23 | 29.7 | 269 | 39.0% |
| r=0.6 | -0.53 | 219.7 | 581 | 99.2% | -0.35 | 55.6 | 161 | 99.3% |
| r=0.5 | -0.60 | 471.1 | 1,728 | 9.4% | -0.37 | 149.7 | 342 | 88.8% |
| r=0.4 | -0.65 | 841.4 | 3,582 | 1.4% | -0.21 | 298.9 | 2,078 | 0.6% |
| r=0.3 | -0.65 | 916.8 | 4,222 | 0.9% | -0.08 | 370.9 | 4,426 | 13.8% |
| r=0.2 | -0.61 | 221.5 | 539 | 99.7% | -0.51 | 574.8 | 1,058 | 33.1% |
| r=0.1 | -0.59 | 861.3 | 4,808 | 0.1% | -0.31 | 562.4 | 4,904 | 0.6% |
| r=0.0 | -0.63 | 225.7 | 564 | 84.0% | -0.31 | 541.4 | 4,973 | 0.0% |
| r=-0.7 | -0.39 | 288.0 | 1,607 | 2.0% | -0.13 | 34.5 | 316 | 74.4% |
| r=-0.6 | -0.64 | 265.3 | 339 | 88.3% | -0.25 | 93.3 | 628 | 55.0% |
| r=-0.5 | -0.68 | 256.0 | 301 | 86.4% | -0.17 | 135.9 | 419 | 96.8% |
| r=-0.4 | -0.69 | 244.9 | 299 | 89.8% | -0.05 | 218.6 | 2,108 | 32.4% |
| r=-0.3 | -0.43 | 724.7 | 4,948 | 0.0% | 0.13 | 231.0 | 5,242 | 1.0% |
| r=-0.2 | -0.40 | 755.1 | 5,956 | 0.0% | -0.51 | 698.4 | 883 | 62.2% |
| r=-0.1 | -0.36 | 737.9 | 7,409 | 0.0% | -0.47 | 486.6 | 702 | 51.9% |
| r=-0.0 | -0.33 | 644.6 | 6,625 | 0.0% | -0.47 | 471.5 | 713 | 63.1% |

| | Subject 18698 - Generalized Pareto | | | | Subject 28433 - Generalized Pareto | | | |
|---|---|---|---|---|---|---|---|---|
| Threshold | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value |
| r=0.8 | -0.44 | 29.3 | 132 | 17.1% | -0.56 | 25.7 | 71 | 59.4% |
| r=0.7 | -0.32 | 61.4 | 410 | 0.8% | 1.25 | 5.7 | 1,268 | 0.0% |
| r=0.6 | -0.27 | 105.9 | 817 | 30.4% | 0.49 | 24.3 | 2,237 | 0.0% |
| r=0.5 | -0.29 | 181.8 | 1,418 | 0.3% | -0.39 | 97.5 | 358 | 100.0% |
| r=0.4 | -0.37 | 288.6 | 1,981 | 10.9% | -0.35 | 179.2 | 1,005 | 36.0% |
| r=0.3 | -0.41 | 397.6 | 3,118 | 1.3% | -0.39 | 272.8 | 1,943 | 19.2% |
| r=0.2 | -0.45 | 495.4 | 4,164 | 4.9% | -0.39 | 359.9 | 3,179 | 8.6% |
| r=0.1 | -0.46 | 529.7 | 5,220 | 72.7% | -0.34 | 353.7 | 3,573 | 4.3% |
| r=0.0 | -0.44 | 441.0 | 4,325 | 76.6% | -0.47 | 280.6 | 1,066 | 92.8% |
| r=-0.7 | -0.30 | 38.8 | 145 | 93.3% | -0.35 | 22.9 | 136 | 48.4% |
| r=-0.6 | -0.03 | 66.7 | 834 | 39.4% | -0.44 | 62.3 | 189 | 96.1% |
| r=-0.5 | -0.12 | 127.1 | 1,133 | 62.8% | -0.38 | 117.5 | 527 | 90.0% |
| r=-0.4 | -0.17 | 193.1 | 1,783 | 71.9% | -0.30 | 175.6 | 1,171 | 50.1% |
| r=-0.3 | -0.23 | 270.2 | 2,740 | 21.4% | -0.28 | 240.4 | 2,189 | 48.4% |
| r=-0.2 | -0.27 | 334.6 | 3,830 | 41.3% | -0.25 | 291.2 | 3,167 | 33.6% |
| r=-0.1 | -0.27 | 347.9 | 4,642 | 50.2% | -0.16 | 291.7 | 4,240 | 4.8% |
| r=-0.0 | -0.22 | 296.1 | 4,507 | 69.5% | -0.08 | 250.6 | 4,372 | 0.5% |

Table 3.9: *Results of generalized Pareto distribution tests, 10 subjects from 1000 Functional Connectome project, part 2*

| | Subject 30421 - Generalized Pareto | | | | Subject 34781 - Generalized Pareto | | | |
|---|---|---|---|---|---|---|---|---|
| Threshold | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value |
| r=0.8 | -0.55 | 65.1 | 113 | 94.7% | -0.45 | 77.6 | 358 | 68.2% |
| r=0.7 | -0.56 | 212.3 | 859 | 84.7% | -0.35 | 72.7 | 258 | 78.5% |
| r=0.6 | -0.67 | 380.5 | 1,558 | 9.7% | -0.29 | 70.9 | 210 | 91.6% |
| r=0.5 | -0.50 | 108.5 | 461 | 54.0% | -0.39 | 142.8 | 571 | 56.1% |
| r=0.4 | -0.39 | 86.8 | 397 | 80.3% | -0.42 | 148.5 | 604 | 86.6% |
| r=0.3 | -0.41 | 129.1 | 642 | 69.0% | -0.56 | 680.0 | 5,873 | 25.6% |
| r=0.2 | -0.42 | 155.0 | 802 | 61.3% | -0.56 | 589.4 | 4,784 | 0.7% |
| r=0.1 | -0.63 | 707.1 | 4,863 | 0.0% | -0.54 | 480.4 | 4,003 | 12.6% |
| r=0.0 | -0.45 | 183.5 | 1,054 | 14.7% | -0.53 | 450.9 | 4,055 | 15.3% |
| r=-0.7 | -0.50 | 175.8 | 693 | 88.3% | -0.40 | 115.0 | 422 | 83.9% |
| r=-0.6 | -0.62 | 320.3 | 1,281 | 42.7% | -0.41 | 195.3 | 920 | 22.6% |
| r=-0.5 | -0.72 | 492.2 | 2,005 | 36.7% | -0.42 | 257.7 | 1,371 | 85.1% |
| r=-0.4 | -0.59 | 206.2 | 808 | 66.3% | -0.41 | 379.4 | 3,308 | 13.9% |
| r=-0.3 | -0.55 | 218.0 | 861 | 75.3% | -0.47 | 483.7 | 4,334 | 92.5% |
| r=-0.2 | -0.63 | 652.6 | 3,640 | 11.3% | -0.46 | 416.4 | 3,348 | 54.7% |
| r=-0.1 | -0.55 | 278.7 | 1,231 | 52.3% | -0.46 | 401.7 | 3,568 | 61.3% |
| r=-0.0 | -0.60 | 537.7 | 3,263 | 4.0% | -0.44 | 368.4 | 3,462 | 32.9% |

| | Subject 47659 - Generalized Pareto | | | | Subject 75922 - Generalized Pareto | | | |
|---|---|---|---|---|---|---|---|---|
| Threshold | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value | k (shape parameter) | sigma (scale parameter) | Number of tail data | p value |
| r=0.8 | -0.38 | 61.4 | 337 | 13.0% | -0.29 | 84.5 | 730 | 0.5% |
| r=0.7 | -0.55 | 112.0 | 424 | 64.5% | -0.50 | 169.2 | 510 | 94.1% |
| r=0.6 | -0.44 | 162.9 | 1,063 | 0.4% | -0.57 | 296.8 | 1,150 | 38.1% |
| r=0.5 | -0.26 | 71.6 | 228 | 80.0% | -0.57 | 313.3 | 1,207 | 29.9% |
| r=0.4 | -0.43 | 319.8 | 2,755 | 11.4% | -0.58 | 518.1 | 2,500 | 0.3% |
| r=0.3 | -0.53 | 520.3 | 4,784 | 0.0% | -0.54 | 588.8 | 3,357 | 0.0% |
| r=0.2 | -0.46 | 261.0 | 1,927 | 7.4% | -0.50 | 605.4 | 3,914 | 0.0% |
| r=0.1 | -0.43 | 191.9 | 1,341 | 30.2% | -0.75 | 463.2 | 1,052 | 15.8% |
| r=0.0 | -0.41 | 152.2 | 1,045 | 61.9% | -0.29 | 542.0 | 9,396 | 0.0% |
| r=-0.7 | -0.21 | 42.9 | 142 | 99.7% | -0.37 | 144.0 | 807 | 91.7% |
| r=-0.6 | -0.37 | 118.4 | 704 | 91.6% | -0.47 | 263.9 | 1,422 | 1.6% |
| r=-0.5 | -0.36 | 174.6 | 1,077 | 73.7% | -0.37 | 159.0 | 655 | 21.1% |
| r=-0.4 | -0.36 | 239.9 | 1,724 | 64.6% | -0.40 | 232.9 | 1,044 | 2.0% |
| r=-0.3 | -0.36 | 292.9 | 2,296 | 18.0% | -0.61 | 260.8 | 439 | 81.1% |
| r=-0.2 | -0.38 | 318.7 | 2,456 | 22.9% | -0.43 | 488.8 | 3,266 | 0.3% |
| r=-0.1 | -0.40 | 308.5 | 2,480 | 8.2% | -0.40 | 471.1 | 3,252 | 0.1% |
| r=-0.0 | -0.40 | 281.8 | 2,266 | 12.9% | -0.38 | 434.0 | 3,144 | 0.0% |

# Chapter 4

# DISCUSSION AND CONCLUSION

**Interpretation of Results**

To our knowledge, this is the first study of the structure of brain functional networks that employs rigorous statistical analysis. In this study we take into account weighted connections among brain regions. Previous studies have mostly focused on binary networks, thus introducing more noise into the analysis. In addition, we examine positive correlation networks, as well as negative correlation networks. Little attention has been paid to the role of anti-correlation networks in the brain context. Our study shows that the topological structure of anti-correlation networks are consistent with those of positive correlation networks. Anti-correlation networks, therefore, could bear relevance for understanding brain functions.

Our analysis rejects the hypothesis that the brain functional networks follow a power-law, or a power-law with exponential cut-off distribution, as postulated in the literature to date. In addition, an analysis of other popular models of distribution shows that the **generalized Pareto** model is the most plausible one for the distribution of brain functional networks.

The distribution model, especially for the tail, of brain functional networks reflects the topological structure of the brain functional hubs. Power-law, or scale-free distribution, indicates the existence of a fat tail, implying larger number of brain hubs compared to random or other small-world network models, ensuring efficiency of information processing and resilience ( [13] and [14]). As discussed in the first chapter, numerous studies have explored the seemingly ubiquitous presence of scale-free characteristic among biological, technological and social networks. Network dynamics of scale-free networks were linked with scale-free structure through the notion of self-organized criticality, which is a property of dynamical systems which have a critical point as an attractor ( [9], [10], [11], and [12]). Some authors proposed that consciousness as a phenomenon is realized through the scale-free organization of the brain operating at critical state. It has been argued that when a complex system such as the brain operates at the phase

transition of order and chaos, the system exhibits scale-free structure ( [12]). Our results showed that brain functional networks are not at all scale-free, as shown through a consistently very small p-value of power-law distribution tests. The idea that brain functional networks follow a power-law with truncated exponential model is similarly rejected. We now turn to a deeper look at the generalized Pareto distribution model.

Formally, the generalized Pareto distribution can be expressed as:

$$y = f(x|k, \sigma, \theta) = \left(\frac{1}{\sigma}\right)\left(1 + k\frac{(x-\theta)}{\sigma}\right)^{-1-\frac{1}{k}}$$

for $\theta < x$ when $k > 0$ and $\theta < x < -\sigma/k$ when $k > 0$ with $k$ being the shape parameter and $\sigma$ being the scale parameter. The following figure demonstrate the different configurations of the probability density function of the generalized Pareto distribution corresponding to the signs of the shape parameter $k$.

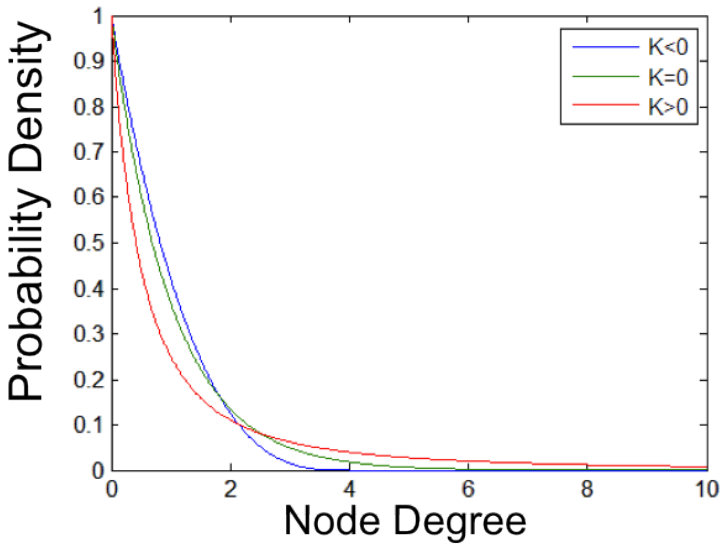



Figure 4.1: *generalized Pareto distribution corresponding to different shape parameter k.*

Note that when $k = 0$, the generalized Pareto distribution becomes the exponential distribution. When $k > 0$, the generalized Pareto distribution is closely related to the normal power-law distribution, exhibiting a fat-tail behavior. When $k < 0$, the generalized Pareto distribution exhibits a short-tail configuration. The results we have across all data sets show that the tails of brain functional networks are topologically approximated by the generalized Pareto model with the shape parameter $k$ consistently negative. What this means is that unlike previous

claims from the literature, brain functional networks do not have fat tails. The case $k < 0$ also corresponds to the q-exponential distribution in statistical physic literature where q¡1. q-exponential distribution was originally proposed to model systems with long-range interactions ( [75]). The link between the q-exponential distribution and brain functional networks should be further explored in the future. Dynamical implications of networks with generalized Pareto distribution and negative shape parameters should also be investigated for future work.

Given that brain functional networks are not scale-free, we wish to examine the structure of brain functional networks under a framework of the generalized Pareto distribution to establish that not only brain networks are efficient, but also are competitive with the scale-free network from the efficiency point of view. Generally this requires the comparison of the our original networks with *null model*, or randomized networks. We employed the methods laid out in section 2.3 for this purpose. Previous studies in the literature have developed a "rewiring" method for binary networks, effectively reshuffle the links in the network in such a way that preserves the degree distribution of the network ( [72]). We developed a similar "rewiring" method, as described in section 2.3.1. When implemented, this method demonstrated that brain networks have high assortativity coefficient and high clustering coefficient, compared to otherwise randomly reshuffled networks, which preserve the degree distribution. Table 4.1 showcases one such comparison between original network of subject 34781 (from 1000 Functional Connectome project) with a rewired, randomized network. The original network is clearly more highly clustered and highly assortative compared with the random network. Note that this holds true across subjects and thresholds.

Table 4.1: *Original Network vs. Rewired Network comparison, subject 34781, threshold r = 0.4*

| Graph Theoretic Measures | Original Functional Network | Rewired, Randomized Network |
|---|---|---|
| Clustering Coefficient | 0.3596 | 0.1442 |
| Assortativity Coefficient | 0.2003 | -0.0082 |

In addition, we compared brain networks from our data sets with two sets of randomized networks created using the bootstrap method described in section 2.3.2.

- *First random bootstrap method:* Original networks were compared to bootstrapped, randomized networks that preserve the same degree distribution as the original networks (see top half of table 4.2)

- *Second random bootstrap method:* Original networks were compared to

bootstrapped, randomized networks that have the scale-free degree distributions, with varying scaling parameter $\alpha$ (see bottom half of table 4.2)

Table 4.2: *Graph-theoretic measures comparison of original vs. randomized, bootstrapped network for subject 34781, threshold r = 0.4*

| Graph Theoretic Measures | Original Functional Network | Bootstraped Random Network with Same Degree Distribution |
|---|---|---|
| Clustering Coefficient | 0.3596 | 0.1709 |
| Characteristic Path Length | 4.7310 | 3.0020 |
| Global Efficiency | 0.2461 | 0.3522 |
| Assortativity | 0.2003 | -0.2718 |
| **Small-world Measure** | **1.3351** | |

| Graph Theoretic Measures | Original Functional Network | Bootstraped Random Scale-Free Network with Alpha =5 |
|---|---|---|
| Clustering Coefficient | 0.3596 | 0.0678 |
| Characteristic Path Length | 4.7310 | 2.7402 |
| Global Efficiency | 0.2461 | 0.3735 |
| Assortativity | 0.2003 | -0.0609 |
| **Small-world Measure** | **3.0720** | |

Results for subject 34781 (one among the 10 subjects) are displayed in table 4.2. Results for other subjects share the same trends and characteristics. As before, we can see that our empirical brain networks display high assortativity and clustering coefficients. Assortativity coefficient measures the tendency of high-degree nodes to be connected to one another. Networks with high assortativity coefficient typically have comparatively resilient cores of mutually inter-connected hubs, effectively allowing for efficient information processing at the global level. This feature of brain functional networks possibly compensates for the relatively less numerous brain hubs compared to scale-free networks. In addition, the presence of densely connected clusters, as indicated through high clustering coefficients, could be another factor that explains the efficiency of brain networks in exchanging information at the local level. Although *characteristic path length* of original networks are higher and thus *global efficiency* of original networks are lower than both versions of randomized networks, *small-worldness* indices in both cases for original networks are both greater than 1, implying brain functional networks possess small-world features in either way that we define randomized network. Interestingly and importantly, the small-world measures of original

networks against randomized networks that preserve degree distribution ($1.3351$ in table 4.2) are less than those of original networks against randomized, scale-free networks ($3.0720$ in table 4.2). This implies the randomized networks with generalized Pareto distribution could outperform the randomized networks with scale-free distribution with regards to the small-worldness attributes. Indeed, the randomized networks with generalized Pareto distributions have relatively similar characteristic path lengths and global efficiency measures, but much higher clustering coefficients than those of randomized networks with scale-free distributions. The take-away from this observation is that scale-free networks are not inherently more efficient than our demonstrated generalized Pareto model. In short, for our brain functional networks of generalized Pareto distribution with negative shape parameters, the combination of the robust local density design (high clustering coefficient) and functionally relevant long-range pathways (likely through assortativity coefficient) provides an economic solution for establishing functionally effective paths across the brain.

**Conclusion**

In summary, we have shown through rigorous statistical analysis that unlike what has been claimed in the literature to date, brain functional networks are not scale-free and also do not follow a power-law with exponential cut-off distribution. Instead, we have demonstrated that the generalized Pareto distribution with negative shape parameter is the most plausible model for brain functional networks. This means brain functional networks do not have fat tails. We propose that brain networks are efficient and competitive with scale-free networks by having high assortativity coefficients, high clustering coefficients and possessing small-world network features. Future research can investigate further into the generalized Pareto distribution to understand its implication both to the structural efficiency of brain networks, as well as to brain network dynamics.

# Bibliography


[1] D. Hao and C. Li, "The dichotomy in degree correlation of biological networks," *PloS one*, vol. 6, no. 12, p. e28322, 2011.

[2] 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.

[3] V. Eguiluz, D. Chialvo, G. Cecchi, M. Baliki, and A. Apkarian, "Scale-free brain functional networks," *Physical Review Letters*, vol. 94, p. 018102, JAN 14 2005. PT: J; NR: 20; TC: 420; J9: PHYS REV LETT; PG: 4; GA: 887LQ; UT: WOS:000226308000087.

[4] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *Journal of Neuroscience*, vol. 26, pp. 63–72, JAN 4 2006. PT: J; NR: 42; TC: 466; J9: J NEUROSCI; PG: 10; GA: 999LM; UT: WOS:000234390800009.

[5] M. P. van den Heuvel, C. J. Stam, M. Boersma, and H. E. H. Pol, "Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain," *NeuroImage*, vol. 43, pp. 528–539, NOV 15 2008. PT: J; NR: 71; TC: 132; J9: NEUROIMAGE; PG: 12; GA: 392JY; UT: WOS:000262300200012.

[6] C. Stam and E. de Bruin, "Scale-free dynamics of global functional connectivity in the human brain," *Human brain mapping*, vol. 22, pp. 97–109, JUN 2004. PT: J; NR: 41; TC: 73; J9: HUM BRAIN MAPP; PG: 13; GA: 824ZZ; UT: WOS:000221727100002.

[7] B. J. He, J. M. Zempel, A. Z. Snyder, and M. E. Raichle, "The temporal structures and functional significance of scale-free brain activity," *Neuron*, vol. 66, pp. 353–369, MAY 13 2010. PT: J; NR: 97; TC: 71; J9: NEURON; PG: 17; GA: 598HE; UT: WOS:000277825200005.

[8] R. Albert, "Scale-free networks in cell biology," *Journal of cell science*, vol. 118, pp. 4947–4957, NOV 1 2005. PT: J; NR: 90; TC: 278; J9: J CELL SCI; PG: 11; GA: 989MS; UT: WOS:000233678700007.

[9] D. Chialvo, "Critical brain networks," *Physica A-Statistical Mechanics and its Applications*, vol. 340, pp. 756–765, SEP 15 2004. PT: J; NR: 36; TC: 65; J9: PHYSICA A; PG: 10; GA: 847LF; UT: WOS:000223393300029.

[10] C.-W. Shin and S. Kim, "Self-organized criticality and scale-free properties in emergent functional neural networks," *Physical Review E*, vol. 74, p. 045101, OCT 2006. PT: J; NR: 21; TC: 27; J9: PHYS REV E; PN: 2; PG: 4; GA: 101EW; UT: WOS:000241723000001.

[11] C. Bedard, H. Kroger, and A. Destexhe, "Does the 1/f frequency scaling of brain signals reflect self-organized critical states?," *Physical Review Letters*, vol. 97, p. 118102, SEP 15 2006. PT: J; NR: 29; TC: 59; J9: PHYS REV LETT; PG: 4; GA: 084PK; UT: WOS:000240545600070.

[12] J. M. Beggs, "The criticality hypothesis: how local cortical networks might optimize information processing," *Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences*, vol. 366, pp. 329–343, FEB 13 2008. PT: J; CT: 9th Experimental Chaos Conference; CY: MAY 29-JUN 01, 2006; CL: San Jose dos Campos, BRAZIL; SP: Natl Inst Space Res; NR: 57; TC: 44; J9: PHILOS T R SOC A; PG: 15; GA: 245IC; UT: WOS:000251927300003.

[13] R. Albert, H. Jeong, and A. Barabasi, "Error and attack tolerance of complex networks," *Nature*, vol. 406, pp. 378–382, JUL 27 2000. PT: J; NR: 24; TC: 2099; J9: NATURE; PG: 6; GA: 337WC; UT: WOS:000088383800038.

[14] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical Review Letters*, vol. 87, p. 198701, NOV 5 2001. PT: J; NR: 25; TC: 588; J9: PHYS REV LETT; PG: 4; GA: 490AT; UT: WOS:000172027200063.

[15] N. L. Bigg, E. K. Lloyd, and R. J. Wilson, *Graph Theory: 1736-1936*. Oxford University Press, 1976.

[16] P. ERDOS and A. RENYI, "On the evolution of random graphs," *Bulletin of the International Statistical Institute*, vol. 38, no. 4, pp. 343–347, 1960. PT: J; NR: 6; TC: 5; J9: B INT STATIST INST; PG: 5; GA: CAW34; UT: WOS:A1960CAW3400027.

[17] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual review of sociology*, pp. 415–444, 2001.

[18] D. Watts and S. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, JUN 4 1998. PT: J; NR: 27; TC: 7302; J9: NATURE; PG: 3; GA: ZR842; UT: WOS:000074020000035.

[19] A. Barabasi, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Physica a*, vol. 272, pp. 173–187, OCT 1 1999. PT: J; NR: 35; TC: 879; J9: PHYSICA A; PG: 15; GA: 244YZ; UT: WOS:000083079500012.

[20] A. Barabasi, A., A. L. Barabasi, and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1015. ID: TNwos000083121200054.

[21] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[22] D. J. de Solla Price, "Networks of scientific papers," *Science*, vol. 149, no. 3683, pp. 510–515, 1965.

[23] S. Redner, "How popular is your paper? an empirical study of the citation distribution," *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 4, no. 2, pp. 131–134, 1998.

[24] P. O. Seglen, "The skewness of science," *Journal of the American Society for Information Science*, vol. 43, no. 9, pp. 628–638, 1992.

[25] R. Albert, H. Jeong, and A.-L. Barabási, "Internet: Diameter of the world-wide web," *Nature*, vol. 401, no. 6749, pp. 130–131, 1999.

[26] A.-L. Barabási, R. Albert, and H. Jeong, "Scale-free characteristics of random networks: the topology of the world-wide web," *Physica A: Statistical Mechanics and its Applications*, vol. 281, no. 1, pp. 69–77, 2000.

[27] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, "Large-scale topological and dynamical properties of the internet," *Physical Review E*, vol. 65, no. 6, p. 066130, 2002.

[28] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, "Graph structure in the web," *Computer networks*, vol. 33, no. 1, pp. 309–320, 2000.

[29] Q. Chen, H. Chang, R. Govindan, and S. Jamin, "The origin of power laws in internet topologies revisited," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, pp. 608–617, IEEE, 2002.

[30] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *ACM SIGCOMM Computer Communication Review*, vol. 29, pp. 251–262, ACM, 1999.

[31] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, "Lethality and centrality in protein networks," *Nature*, vol. 411, no. 6833, pp. 41–42, 2001.

[32] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, "The large-scale organization of metabolic networks," *Nature*, vol. 407, no. 6804, pp. 651–654, 2000.

[33] W. Aiello, F. Chung, and L. Lu, "A random graph model for massive graphs," in *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pp. 171–180, Acm, 2000.

[34] W. Aiello, F. Chung, and L. Lu, "Random evolution in massive graphs," in *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, pp. 510–519, IEEE, 2001.

[35] J. H. Jones and M. S. Handcock, "An assessment of preferential attachment as a mechanism for human sexual network formation," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, no. 1520, pp. 1123–1128, 2003.

[36] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg, "The web of human sexual contacts," *Nature*, vol. 411, no. 6840, pp. 907–908, 2001.

[37] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, "Classes of small-world networks," *Proceedings of the National Academy of Sciences*, vol. 97, no. 21, pp. 11149–11152, 2000.

[38] P. Sen, S. Dasgupta, A. Chatterjee, P. Sreeram, G. Mukherjee, and S. Manna, "Small-world properties of the indian railway network," *Physical Review E*, vol. 67, no. 3, p. 036106, 2003.

[39] M. E. Newman, "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 404–409, 2001.

[40] M. E. Newman, "Assortative mixing in networks," *Physical review letters*, vol. 89, no. 20, p. 208701, 2002.

[41] C. Leung and H. Chau, "Weighted assortative and disassortative networks model," *Physica A: Statistical Mechanics and its Applications*, vol. 378, no. 2, pp. 591–602, 2007.

[42] R. Pastor-Satorras, A. Vazquez, and A. Vespignani, "Dynamical and correlation properties of the internet," *Physical review letters*, vol. 87, no. 25, p. 258701, 2001.

[43] J.-P. Onnela, J. Saramaki, J. Kertesz, and K. Kaski, "Intensity and coherence of motifs in weighted complex networks," *Physical Review E*, vol. 71, no. 6, p. 065103, 2005.

[44] B. Jiang and C. Claramunt, "Topological analysis of urban street networks," *Environment and Planning B*, vol. 31, no. 1, pp. 151–162, 2004.

[45] A. Fronczak, J. A. Hołyst, M. Jedynak, and J. Sienkiewicz, "Higher order clustering coefficients in barabási–albert networks," *Physica A: Statistical Mechanics and its Applications*, vol. 316, no. 1, pp. 688–694, 2002.

[46] G. Caldarelli, R. Pastor-Satorras, and A. Vespignani, "Structure of cycles and local ordering in complex networks," *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 38, no. 2, pp. 183–186, 2004.

[47] G. Bianconi and A. Capocci, "Number of loops of size h in growing scale-free networks," *Physical review letters*, vol. 90, no. 7, p. 078701, 2003.

[48] M. E. Newman, "Ego-centered networks and the ripple effect," *Social Networks*, vol. 25, no. 1, pp. 83–95, 2003.

[49] S. N. Soffer and A. Vazquez, "Clustering coefficient without degree correlations biases," *arXiv preprint cond-mat/0409686*, 2004.

[50] B. V. Cherkassky, A. V. Goldberg, and T. Radzik, "Shortest paths algorithms: theory and experimental evaluation," *Mathematical programming*, vol. 73, no. 2, pp. 129–174, 1996.

[51] S. Milgram, "The small world problem," *Psychology today*, vol. 2, no. 1, pp. 60–67, 1967.

[52] M. D. Humphries and K. Gurney, "Network Ôsmall-world-nessÕ: a quantitative method for determining canonical network equivalence," *PLoS One*, vol. 3, no. 4, p. e0002051, 2008.

[53] 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.

[54] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the structural core of human cerebral cortex," *Plos Biology*, vol. 6, pp. 1479–1493, JUL 2008 2008. PT: J; TC: 520; UT: WOS:000257971100019.

[55] B. B. Biswal, M. Mennes, X.-N. Zuo, S. Gohel, C. Kelly, S. M. Smith, C. F. Beckmann, J. S. Adelstein, R. L. Buckner, S. Colcombe, *et al.*, "Toward discovery science of human brain function," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, pp. 4734–4739, 2010.

[56] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, pp. 1059–1069, SEP 2010. PT: J; NR: 70; TC: 260; J9: NEUROIMAGE; PG: 11; GA: 629FY; UT: WOS:000280181800027.

[57] K. J. Friston, L. Harrison, W. Penny, *et al.*, "Dynamic causal modelling," *Neuroimage*, vol. 19, no. 4, pp. 1273–1302, 2003.

[58] S. Achard and E. Bullmore, "Efficiency and cost of economical brain functional networks," *Plos Computational Biology*, vol. 3, pp. 174–183, FEB 2007. PT: J; NR: 52; TC: 295; J9: PLOS COMPUT BIOL; PG: 10; GA: 143HQ; UT: WOS:000244711500003.

[59] M. E. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.

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

[61] 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.

[62] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social networks*, vol. 1, no. 3, pp. 215–239, 1979.

[63] S. Hayasaka and P. J. Laurienti, "Comparison of characteristics between region-and voxel-based network analyses in resting-state fmri data," *Neuroimage*, vol. 50, no. 2, pp. 499–508, 2010.

[64] K. E. Joyce, P. J. Laurienti, J. H. Burdette, and S. Hayasaka, "A new measure of centrality for brain networks," *PLoS One*, vol. 5, no. 8, p. e12200, 2010.

[65] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *SIAM Review*, vol. 51, pp. 661–703, DEC 2009. PT: J; NR: 69; TC: 505; J9: SIAM REV; PG: 43; GA: 522FO; UT: WOS:000271983500002.

[66] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9673–9678, 2005.

[67] L. Q. Uddin, A. Clare Kelly, B. B. Biswal, F. Xavier Castellanos, and M. P. Milham, "Functional connectivity of default mode network components: correlation, anticorrelation, and causality," *Human brain mapping*, vol. 30, no. 2, pp. 625–637, 2009.

[68] M. D. Fox, D. Zhang, A. Z. Snyder, and M. E. Raichle, "The global signal and observed anticorrelated resting state brain networks," *Journal of neurophysiology*, vol. 101, no. 6, pp. 3270–3283, 2009.

[69] X. J. Chai, A. N. Castañón, D. Öngür, and S. Whitfield-Gabrieli, "Anticorrelations in resting state networks without global signal regression," *Neuroimage*, vol. 59, no. 2, pp. 1420–1428, 2012.

[70] Q. VUONG, "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica*, vol. 57, pp. 307–333, MAR 1989. PT: J; NR: 48; TC: 932; J9: ECONOMETRICA; PG: 27; GA: U7126; UT: WOS:A1989U712600002.

[71] M. E. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Physical Review E*, vol. 64, no. 2, p. 026118, 2001.

[72] S. Maslov and K. Sneppen, "Specificity and stability in topology of protein networks," *Science*, vol. 296, no. 5569, pp. 910–913, 2002.

[73] A. Clauset, C. Moore, and M. E. J. Newman, "Hierarchical structure and the prediction of missing links in networks," *Nature*, vol. 453, pp. 98–101, MAY 1 2008. PT: J; NR: 30; TC: 224; J9: NATURE; PG: 4; GA: 294ID; UT: WOS:000255398800046.

[74] R. P. Brent, *Algorithms for minimization without derivatives*. Courier Dover Publications, 1973.

[75] C. TSALLIS, "Possible generalization of boltzmann-gibbs statistics," *Journal of Statistical Physics*, vol. 52, pp. 479–487, JUL 1988. PT: J; NR: 4; TC: 2946; J9: J STAT PHYS; PG: 9; GA: Q2401; UT: WOS:A1988Q240100029.