RESEARCH

Identifying Brain Network Topology Changes in Task Processes and Psychiatric Disorders

Paria Rezaeinia[1], Kim Fairley[2,3], Piya Pal[1], François G. Meyer[4] & R. McKell Carter[3,5,6]

[1]Department of Electrical and Computer Engineering, University of California San Diego, San Diego, U.S.A.

[2]Department of Economics, Leiden University, Leiden, The Netherlands

[3]Institute of Cognitive Science, University of Colorado Boulder, Boulder, U.S.A.

[4]Department of Applied Mathematics, University of Colorado Boulder, Boulder, U.S.A.

[5]Department of Electrical, Computer and Energy Engineering, University of Colorado Boulder, Boulder, U.S.A.

[6]Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, U.S.A.

Keywords: (fMRI, functional connectivity, random walk, hitting time)

ABSTRACT

A central goal in neuroscience is to understand how dynamic networks of neural activity produce effective representations of the world. Advances in the theory of graph measures raise the possibility of elucidating network topologies central to the construction of these representations. We leverage a result from the description of lollipop graphs to identify an iconic network topology in functional magnetic resonance imaging data and characterize changes to those networks during task performance and in populations diagnosed with psychiatric disorders. During task performance, we find that task-relevant subnetworks change topology, becoming more integrated by increasing connectivity throughout cortex. Analysis of resting-state connectivity in clinical populations shows a similar pattern of subnetwork topology changes; resting-scans becoming less default-like with more integrated sensory paths. The study of brain network topologies and their relationship to cognitive models of information processing raises new opportunities for understanding brain function and its disorders.

AUTHOR SUMMARY

Our mental lives are made up of a series of predictions about the world calculated by our brains. The calculations that produce these predictions are a result of how areas in our brain interact. Measures based on graph representations can make it clear what information can be combined and therefore help us better understand the computations the brain is performing. We make use of cutting-edge techniques that overcome a number of previous limitations to identify specific shapes in the functional brain network. These shapes are similar to hierarchical processing streams which play a fundamental role in cognitive neuroscience. The importance of these structures and the technique is highlighted by how they change under different task constraints and in individuals diagnosed with psychiatric disorders.

INTRODUCTION

How do we link dynamic changes in functional brain structure to the processing of information? Brain activity organizes into stable networks that vary in strength and change with task demands Greicius, Krasnow, Reiss, and Menon (2003); Smith et al. (2009).

Because of its ease of implementation and relatively low cost, the analysis of resting functional magnetic resonance imaging (rfMRI) data Raichle et al. (2001) in particular has had a tremendous impact, leading to several large-scale public initiatives like the Human Connectome Project (HCP) Essen et al. (2013). One of the most promising methods used to study rfMRI activation has been to construct network models of functional connectivity between areas of the brain E. T. Bullmore and Bassett (2011); Goñi et al. (2014); van den Heuvel and Pol (2010). These models are characterized by network measures like efficiency Fornito, Zalesky, and Bullmore (2016) and have been applied to a wide variety of challenges including the study of psychiatric disorders (for review, see Avena-Koenigsberger, Misic, and Sporns (2017)). Improving our ability to interpret the meaning of these measures for brain processing would have tremendous impact.

To improve our ability to interpret network models of brain connectivity, we seek measures of topology that can be related to models of cognitive information processing. The study of the relationship between brain network topology and function has been accelerating and is key to explaining dynamic information processing in health and disease Stiso and Bassett (2018). To better understand how information is processed in a dynamic context, it is necessary to link specific brain-network topologies to cognitively meaningful information processing structures. Network analysis of brain data typically involves

descriptions of an inferred network. Here, we instead describe brain connections as stochastic processes (in our case, using a random walk), avoiding the constraints of a specific network model and instead describing general properties of brain functional connectivity in a given mental state. This improved description of brain connectivity can then be used to link results from graph theory to network topologies common in cognitive models of the brain. As a first step, we utilize a result from the theory of graph measures, which establishes that isolated chains of nodes produce maximally long random walks between points on the graph. In particular, a lollipop graph consists of a set of fully connected nodes attached to a chain of linearly-connected nodes. In a random walk on a lollipop graph, the number of hops required to reach the tail of the lollipop stick is greater than for other topological structures Brightwell and Winkler (1990). We target extremely long random walks between brain areas as a measure of the presence, and relative isolation, of linear chains of nodes. We note that this topology is similar to that found in hierarchical processing streams, a structure important in cognitive models. We hypothesize that those brain areas which take a long time to reach in a random walk are often situated in such an information processing topology.

We focus on the tails of random-walk network connectivity distributions to address the following four key questions. (i) How does the relative isolation of a linear chain of nodes change the distribution of connectivity in a synthetic network? (ii) Are there subnetworks in resting-state cortex that have properties similar to a linear chain of nodes? (iii) How are linear-chain subnetworks changed by task demands? (iv) Does the characterization of network topology have value in understanding and diagnosing psychiatric disorders?

MATERIALS AND METHODS

Hitting-time functional connectivity model

One common approach to find the connectivity matrix of a brain network is to threshold the Pearson correlation matrix to obtain the adjacency matrix for the network. Although this method is very simple, it has some shortcomings that might cause inaccuracy in the results. One challenge is that the Pearson correlation coefficient does not account for latent variables, which might result in a high correlation among two regions that are not directly connected. In addition, the choice of threshold is arbitrary, creating interpretation and generalization issues. To overcome these challenges, we integrated the following changes into a standard network analysis pipeline for neuroimaging. First, to compensate for latent variables, we

use the partial correlation Smith et al. (2011) to find the connectivity matrix. Let ρ_{ij} represent the partial correlation between x_i and x_j (the BOLD time series associated with regions i and j, respectively). Therefore, we use a weighted brain functional network with adjacency matrix $A = [\rho_{ij}]$. The degree of node i is $d_i = \sum_{j=1}^{N} |\rho_{ij}|$. Second, we normalize edge strength using self loops that preserve the overall connectivity of each node relative to others. Third, we characterize the network using the hitting time, a random-walk measure that reflects the expected number of edges that need to be crossed to transition from one node to another. We next describe the edge strength and hitting time approaches in detail.

Edge Strength Normalization

For a random walk, the probability transition matrix is $P = [p_{ij}]$, where p_{ij} is defined as:

$$p_{ij} = \frac{|\rho_{ij}|}{\sum_{j=1}^{N} |\rho_{ij}|} = \frac{|\rho_{ij}|}{d_i} \tag{1}$$

The major drawback of this definition is that it fails to distinguish a strongly connected from a weakly connected node. Consider a network with 5 nodes (a, b, c, d, e) and 6 edges. Suppose that all edges connected to node a have weight 0.9. And, suppose that node b is connected to the same nodes as node a, but with edges with weight 0.1 (Fig. 1A). Applying equation (1), both nodes a and b will have the same transition probabilities, and therefore, the same relative connectivity.

To overcome this problem, we add a self loop to nodes with weaker connections. To implement this, we find the node with maximum degree in the network. For every other node, we subtract the degree of that node from the maximum degree and add that as a self edge to the node. Therefore, the new degree matrix is $D' = d_{max} I$, and the new adjacency matrix is $A' = [\rho'_{ij}]$, where:

$$\begin{aligned} \rho'_{ij} &= \rho_{ij}, \qquad i, j = 1, \ldots, N \; and \; i \neq j, \\ \rho'_{ii} &= d_{max} - d_i, \qquad i = 1, \ldots, N \end{aligned} \tag{2}$$

$d_{max} = \max_i(d_i), i = 1, \ldots, N$ and I is the identity matrix.

Hitting Time

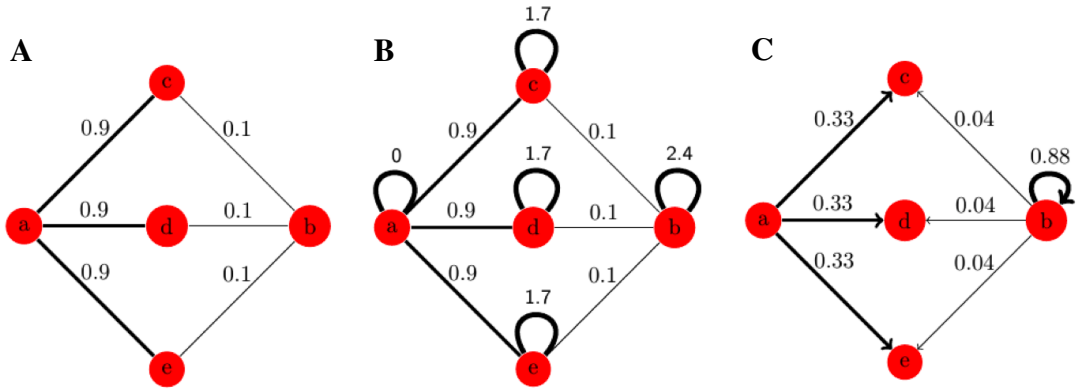

Figure 1: Edge strength normalization to maintain connectivity differences between a strongly connected and a weakly connected node. (A) A weighted graph with 5 nodes and 6 edges. (B) Adding self loops to nodes with weaker connections in order to normalize the probabilities. (C) Transition probabilities from nodes a and b after normalization (the transition probabilities from nodes c, d, and e are not included in this figure).

We run random walk models on the graph with the new transition probability matrix $P' = D'^{-1}A'$ to calculate the hitting time matrix $H = [h_{ij}]$. h_{ij}, the hitting time from node i to node j is the expected number of hops to visit node j for the first time, for a random walk started at node i.

Hitting time is an asymmetric measure, meaning that h_{ij} and h_{ji} might be different. For example for a lollipop graph, the hitting times from the nodes on the complete component to nodes on the chain are much larger than the reverse direction, because a random walker spends more time in the complete component. We compute the hitting times between pairs of nodes using the graph-Laplacian method introduced in spectral graph theory Aldous and Fill (2002). This method is advantageous as it does not require the exact knowledge of the adjacency matrix, instead using a probabilistic approximation of the adjacency matrix of the network. Following Lovász and Simonovits (1993), we calculated the normalized graph Laplacian as:

$$\mathscr{L} = D'^{-1/2}(D' - A')D'^{-1/2} = I - P' \tag{3}$$

where, D' is the degree matrix and A' is the adjacency matrix of the graph after normalization as defined in the main text. We used the eigenvalues and eigenvectors of $\mathscr{L}$ to calculate the hitting time matrix

$H = [h_{ij}]$ Lovász and Simonovits (1993);

$$h_{ij} = \sum_{k>1} \frac{d'}{\lambda_k}\left(\frac{\mu_{kj}^2}{d'_j} - \frac{\mu_{ki}\mu_{kj}}{\sqrt{d'_i d'_j}}\right), \quad i, j = 1, \ldots, N \; and \; i \neq j \tag{4}$$
$$h_{ii} = 0, \qquad i = 1, \ldots, N$$

where, d'_i is the degree of node $i, i = 1, ..., N$, and d' is the sum of all degrees (after normalization, see main text). $0 = \lambda_1 < \lambda_2 < \cdots < \lambda_n$ are the n eigenvalues of $\mathscr{L}$, and μ_{kj} is the j^{th} element of k^{th} eigenvector of $\mathscr{L}$ Lovász and Simonovits (1993).

Adding the self loops in the normalization step does not make the graphs reducible or periodic, meeting the requirements of the hitting time calculation we use here Norris (1997). Code for analysis in this project can be found under the first author's name on github: https://github.com/SNaGLab.

RESULTS

To detect and characterize linear chains of nodes, we focus on the random-walk measure of connectivity hitting time Lovász and Simonovits (1993) defined above. In synthetic graphs and estimated networks, a node is a point in the graph (or area of the brain) and an edge is a connection between two nodes. Hitting time between nodes i and j is a random variable describing the number of steps to get from node i to node j for the first time (represented as h_{ij}) during a random walk, a measure equivalent to mean first-passage time Avena-Koenigsberger et al. (2017). Diffusion measures of networks, like hitting time, are becoming more commonly used and are the focus of active research Goñi et al. (2013); Lambiotte, Delvenne, and Barahona (2014); Shen and Meyer (2008). Diffusion-based measures carry significant methodological advantages. First, they overcome common issues caused when thresholding is used to define binary connections Goulas, Schaefer, and Margulies (2015); Reijneveld, Ponten, Berendse, and Stam (2007); Rubinov and Sporns (2011); Zalesky, Fornito, and Bullmore (2010). Second, they do not require perfect knowledge of the network to make a robust estimation of connectivity (i.e. it does not require the exact adjacency matrix Lovász and Simonovits (1993)). Third, measures like hitting time are asymmetric, meaning hitting time from one node to another may be different from the return trip, giving the best opportunity to identify extremeness in connectivity Lovász and Simonovits (1993). Here, we will use "hitting time" to refer to the expected number of edges to be traversed rather than the variable itself and

"hitting-time distribution" to be the subject-average distribution of the expected number of edges to be traversed when moving between combinations of nodes. We begin by looking at the relationship between extreme hitting times and synthetic graph structure and then extend those findings to a publicly available functional magnetic resonance imaging (fMRI) dataset.

How does the relative isolation of a linear chain of nodes change the distribution of connectivity in a synthetic network? We consider a chain of sequentially connected nodes as a model for a hierarchical processing stream. Formally, a chain of sequentially connected nodes can be described as N nodes arranged in a line, so that there is an edge between nodes i and $i+1$ for $i = 1, \ldots, N-1$, and no edges between nodes i and j where, $j \neq i-1, i+1$. Theoretical results have found that a chain of sequentially-connected nodes attached to a fully-connected network (i.e. a lollipop graph) results in maximal hitting times when the chain is a third of the network Brightwell and Winkler (1990). We now compare the distribution of hitting times over nodes in a lollipop graph to small-world Watts and Strogatz (1998), random (Erdős-Rényi) Erdős and Rényi (1959), and complete synthetic graphs in Fig. 2. Each graph consists of 100 nodes. Random and small-world graphs are an average of 100 configurations. Linear chains of nodes result in larger hitting times which produce increased skewness in the hitting-time distribution. For example, in Fig. 2, panels A and D represent two extreme examples of hitting-time distribution in a network. Due to the presence of a path in A, the hitting time has a long tail (non-zero probability) that extends to large values (above 120,000). While in D each node is fully connected to every other node and hence, hitting time is the same across all pairs (in this case, 100) and the distribution is a single value with no tail. We focus on Kelley skewness Kelley (1923), as our measure of skewness because it directly compares the tails of the distribution. Kelley skewness (hereafter just skewness) therefore provides a more robust separation of extreme cases from changes in the interior of the distribution.

Although perfectly isolated linear chains of nodes produce extreme hitting times, it is possible that even weak connections to the chain might significantly reduce hitting times to nodes on the chain. To characterize changes in the hitting-time distribution when a linear chain of nodes is not perfectly isolated, we begin with a random graph and alter its connectivity to isolate a linear chain of nodes (Fig. 3). Beginning with 50 nodes, edges of weight 1 were added between each pair of nodes with a probability of 0.6. We then randomly chose 10 connected nodes in the graph ($1/5$ of the graph) and reduced the weight of edges between those connected nodes and the rest of the network by 0.05 for 19 iterations. This process

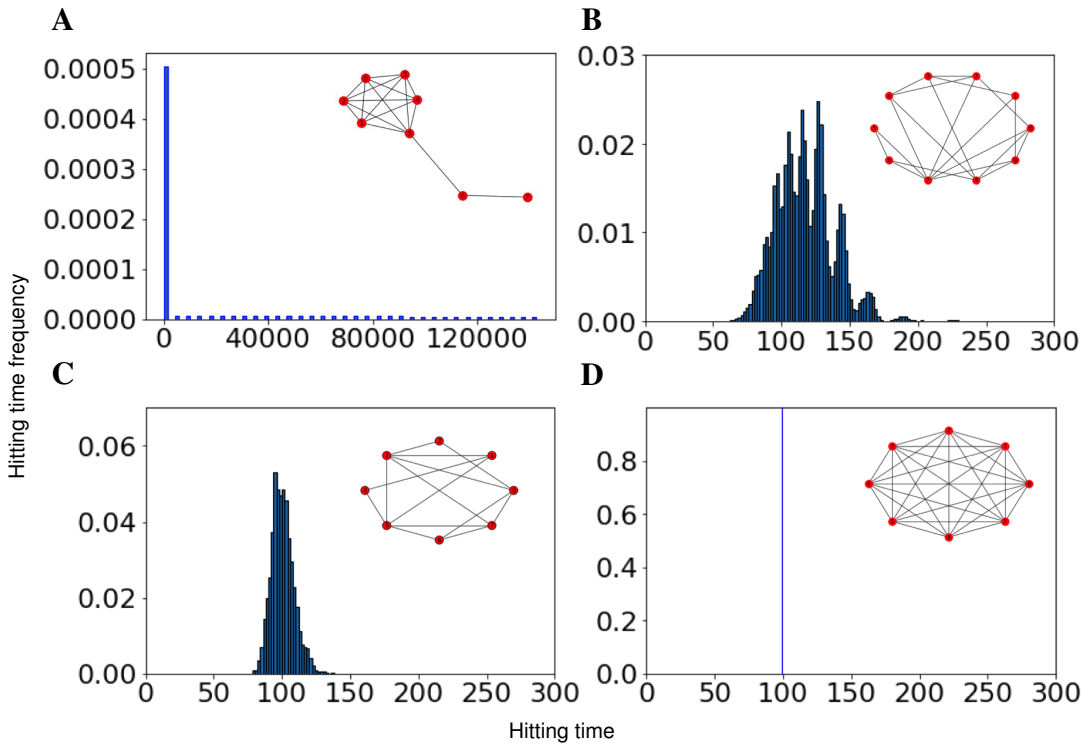

Figure 2: Hierarchical processing streams in lollipop graphs produce extremely long hitting times. Hitting-time distributions for (A) lollipop, (B) small world, (C) random and (D) complete graphs with 100 nodes (averaged over 100 runs for small world and random networks). The graphs on top of the distributions are smaller representations of the graphs used to generate the distributions. Axis scales change significantly with graph type (e.g. lollipop hitting time is several orders of magnitude larger than the other graphs).

created a linear subgraph that becomes progressively more isolated until it resembles the stick of a lollipop graph. As a control for reduced connectivity across the network as a whole, we took the same graph we started with above and reduced all existing edge weights by 0.05 for 19 iterations. To preserve the reduction in overall connectivity (edge weights are typically normalized by the total connectivity of a node) edge weight reductions were added back as self loops (see Materials and Methods). These self loops are required as part of the random-walk to preserve the physiological principle that reduced neuronal activity would result in a reduction of connectivity (not just a shift between connections).

Average hitting time increases as the chain of nodes becomes more isolated but also when the graph becomes more disconnected as a whole. Hence, mean hitting time does not distinguish between these two

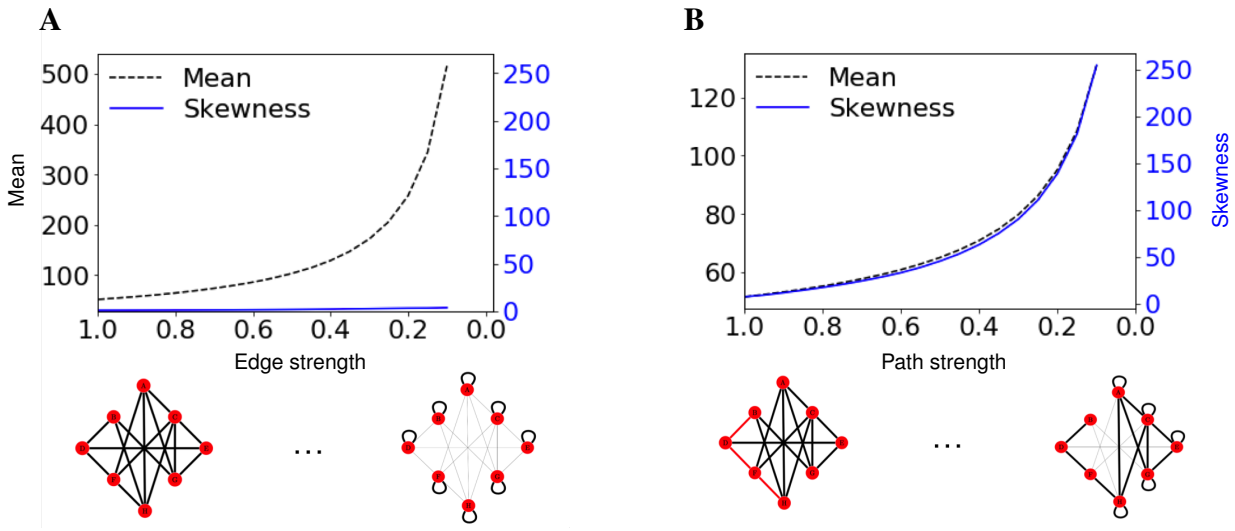

Figure 3: The skewness of the hitting-time distribution distinguishes a reduction of overall connectivity from a subgraph that becomes more linear. (A) Mean and skewness of hitting-time distribution as the strength of all connections is reduced by 0.05 for 19 iterations. Toy networks for this transition are represented on the x-axis. Reductions in connectivity are added as self loops. (B) Mean and skewness of hitting-time distribution as the strength of connections between linear component and the rest of the graph is reduced by 0.05 for 19 iterations. Toy networks on the x-axis represent synthetic graphs with a linear subgraph as the path (red) is made more linear.

scenarios. However, in our simulations, skewness changed significantly as the linear subgraph becomes more isolated but only minimally when the average connectivity of the whole graph decreased. Skewness also increased with the relative isolation of the chain of nodes but was present even when each node in the chain was somewhat connected to the rest of the graph (Fig. 3).

We have shown that a lollipop component present in a graph results in significant increase of (Kelley) skewness of hitting-time distribution but is this relationship true in heterogeneous topologies? It is important to note that other changes in graph structure may also result in extreme hitting times. One commonly employed graph measure is modularity, the extent to which the graph can be easily separated into different communities.

To evaluate the effect of modularity on hitting-time distribution, we have tested a large number of networks with different levels of Louvain modularity and numbers of chain motifs (3 node linear

components). We generated random networks each with 100 nodes varying the number of edges. To allow a comparison across a given number of edges, we generated multiple graphs with the same number of edges by choosing k edges uniformly from the full possible set of edges. k was varied from 200 to 1000 in intervals of 50 and 1000 to 2500 in intervals of 100. The range of the average degree is $[4, 50]$, with a range of Louvain modularity of $[.1, .6]$. Keeping only those graphs that were connected resulted in 15,243 graphs for comparison. Using a linear model, we sought to explain skewness as a function of modularity, number of edges, and number of chain motifs. Number of edges ($p < .001, t(15241) = 4.16, \beta = -.0052$), modularity ($p < .001, t(15241) = 37.4, \beta = 239$), and the number of chain motifs ($p < .001, t(15241) = 13.8, \beta = 9.5$) all independently explain some variance in skewness. In the remaining analysis of brain networks, we therefore test whether or not nodes with extreme hitting times also become less chain like.

Are there subnetworks in resting-state cortex that have properties similar to a linear chain of nodes? Motivated by the above simulations, we utilized the skewness of the hitting-time distribution to identify potential linear chains of nodes in cortical connectivity data. The brain is made up of a large number of highly interconnected regions Cherniak (1990) evolved to efficiently integrate a variety of sources of information Friston (2010) that can be represented as a network. Graph-theoretic models of the brain have been used to effectively segment commonly associated regions of the brain into large-scale networks and describe the properties of brain information processing (for review see E. Bullmore and Sporns (2009)) in health and disorder Bassett and Sporns (2017); Fox and Greicius (2010). Characterizations of brain network changes in development or psychiatric disorder often utilize graph measures like efficiency Latora and Marchiori (2001) and small-worldness Watts and Strogatz (1998), which typically include the average path length in their definition (for common measurement descriptions, see Achard and Bullmore (2007)). Even measures that may not directly utilize the average path length (e.g. modularity Newman and Girvan (2004); Stiso and Bassett (2018)) sometimes rely on community detection methods that incorporate the average path length. The use of an average path-length rests on the assumption that path lengths in that network are normally distributed and so can lead to the mischaracterization of the topology of the network. The concern arises because of the use of an average and is present in both traditional and diffusion-based graph measures. Overcoming this assumption requires the use of specific subnetwork models (for example, see Khambhati, Medaglia, Karuza, Thompson-Schill,

and Bassett (2018)) or the capture of deviations from normality in the path-length distribution. Here, we use Kelley skewness of the hitting-time distribution to distinguish changes in the network as a whole from the presence of network topologies resembling hierarchical processing streams.

To test for the presence of skewness in cortical connectivity, we generated a hitting-time measure of connectivity (see materials and methods) for resting-state functional data from neurotypical participants who were part of a large open-source dataset (LA5c, UCLA Consortium for Neuropsychiatric Phenomics Poldrack et al. (2016), see supporting information). Network nodes were 180 anatomical regions from the multi-modal parcellation of Glasser et al. Glasser et al. (2016). The average hitting-time distribution of neurotypical resting-state functional connectivity is positively skewed (Kelley skewness of 15.04 and Pearson's coefficient of skewness of 2.3), Fig.4A. A D'Agostino-Pearson test Trujillo-Ortiz and Hernandez-Walls (2003) showed that as a whole the hitting times were not normally distributed ($Z(skew) = 110,3496, \chi^2(2) = 17864.8071, p < 0.001$).

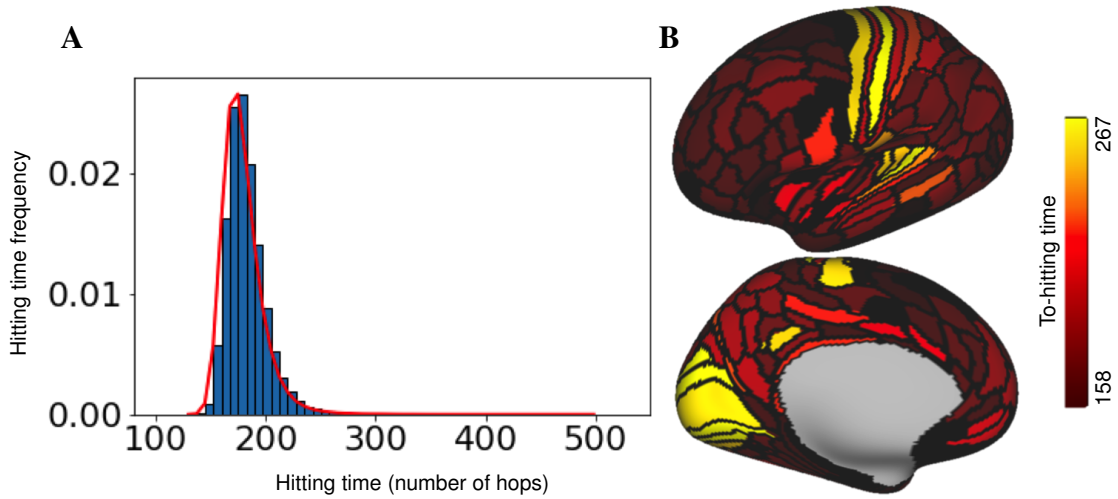

Figure 4: Hitting-time measures of resting-state functional connectivity in neurotypical participants are positively skewed. (A) Average normalized hitting-time distribution for control subjects during resting-state from the publicly available LA5c study. (B) Average to-hitting time from all other regions of cortex for lateral (top) and medial (bottom) left maps thresholded to $[158(10\%), 267(95\%)]$. The range of data is $[129, 310]$. Primary auditory, visual, and somatosensory cortices have the largest to-hitting times in the cortex.

Although skewness of a distribution can originate from many sources, ranging from a smooth shift of the distribution as a whole to far ranging outliers, the particular skewness measure used here (Kelley skewness) directly compares the extremes of the distribution (90% compared to 10%), limiting the potential causes of the skewness. Limiting our test for skewness to the tails of the distribution is consistent with our aim of identifying changes in linear-chain topologies, which have been shown to produce maximal hitting times in lollipop networks (see above). The primary auditory, visual, and somatosensory hierarchies, show the largest average to-hitting times (Fig. 4B), and are therefore possibly related to chain-like network topologies. It is important to note that even the use of a Kelley skewness metric does not guarantee the presence of chain-like network topologies. In fact, random graphs generated from a stochastic block model that precluded chain-like topologies exhibited Kelley skewness explained by modularity and node degree. We generated 100 graphs with 180 nodes from a stochastic block model. To define groups and mixing structure, we fixed the probability of connections within communities to be $(p = 0.7)$ and between communities to be $(q = 0.1)$. To expand the range of possible modularity, we randomly picked the number of nodes in each community until we reach 180 (if the total number of nodes passes 180, we reduce the size of the last community to have a total of 180 nodes in the graph). After removing the null values, we ended up with 89 graphs with number of communities from 2 to 6 and Louvain modularity levels of 0.03 to 0.42. These SBM models contain Kelley skewness which is explained by both modularity $(\beta = -1.591e+03, t(86) = -10.76, p < 0.001)$ and degree $(\beta = -8.617e-02, t(86) = -7.39, p < 0.001)$. Because Kelley skewness may arise from multiple sources, we next look for changes in the skewness of the hitting-time distribution during a task and ask whether those changes are related to regions of the brain with the largest hitting times during resting scans, and then, whether those areas with the largest rest hitting times also become less chain-like.

How are linear-chain subnetworks changed by task demands? To better interpret the skewness of cortical networks during resting-state fMRI, we sought to test whether hitting times become more or less skewed during task performance and which connectivity changes underlie those shifts. We compared resting-state and balloon analogue risk task (BART) functional connectivity from the LA5c study. The BART is a paradigm designed to study risk taking in an experimental setting. Participants in the BART decide whether or not to pump a balloon that is at risk of popping. The BART is defined by visual input and motor responses without structured auditory stimulation.

Hitting times between cortical areas were calculated for fMRI data collected during the performance of the BART task using the same processing pipeline as for the resting-state scans. To test for differences in skewness, we then ran a linear mixed effect model (lme in R) of skewness of the hitting-time distributions (dependent variable) modeling task (with resting-state as a reference), gender, and age as independent variables. The task variable was treated as a random effect (BART and resting-state points were paired by participant), which characterizes idiosyncratic variation that is due to individual differences. In our first model, we found significant difference in skewness for control subjects between BART and rest ($\beta = -10.25, t(118) = 0.79, p < 0.001$), see Fig. 5. The skewness of hitting-time distribution is significantly reduced in the BART ($\mu = 5.48, \sigma = 3.96$) compared to rest ($\mu = 15.73, \sigma = 8.49$). Age and gender did not significantly explain variance in this model. We next sought to test whether this skewness could be related to nodes with extreme hitting times and whether those nodes define network topologies that become less chain-like.

To identify those nodes related to differences between rest and task, we first ask which nodes had the largest hitting-time changes. The ten regions with the largest hitting-time changes (paired t-tests comparing task and rest hitting times, significantly different with $p < 0.05$, Bonferroni corrected) are 'V2', 'V3', 'V4', 'V3A' and 'PGs' within the visual cortex, '1', '4', '3b' and 'OP4' within somatosensory cortex, and the area 'PF', Fig. 5B-D. Regions are labeled according to Glasser et al. (2016). These nodes, which show decreased hitting times during task performance, overlap heavily with the visual and motor processing streams and correspond to many of the nodes with the largest hitting times during rest scans. This reduction in hitting time in the visual and motor pathways during the BART provides support for the role of these pathways in skewness but does not address whether differences in the chain-like topology is responsible for the change in hitting times.

To add support for the role of chain-like topologies in large hitting times in brain data, we calculated a chain-index for each node that provides a measure of how similar to an isolated chain the local connectivity of the node is. For every node in the network located on a three node chain motif, we define a chain index by focusing on its two strongest connections. A node is located on a chain motif if its neighbors with the two strongest connections are stronger than the remaining connections and if the two strongest connections have a significantly weaker connection with each other. Assuming that node i is on a chain motif and has

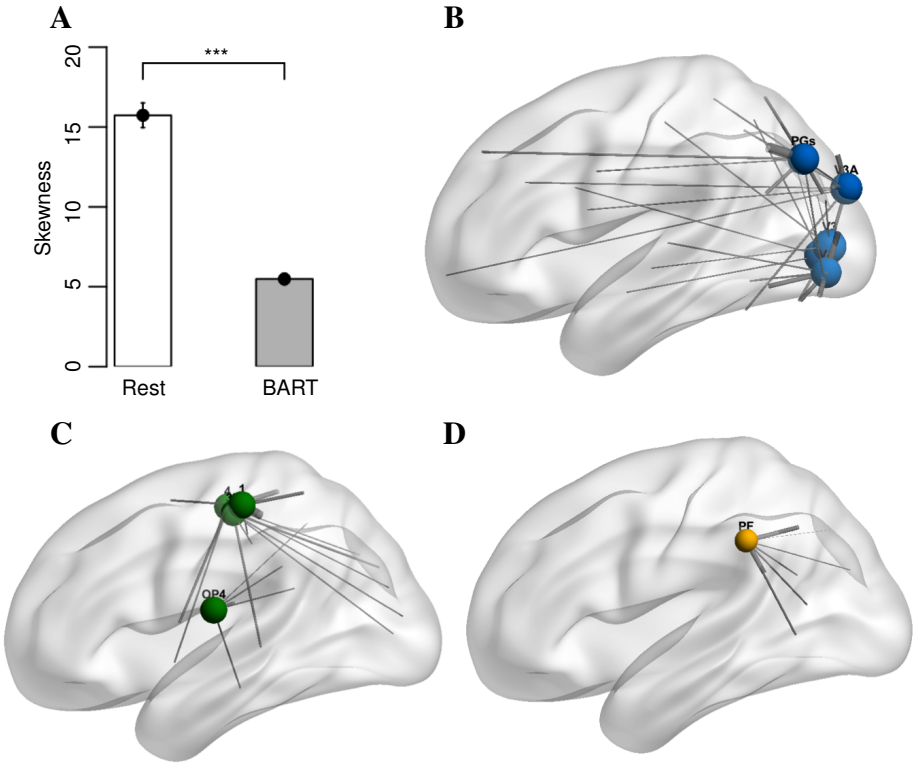

Figure 5: The distribution of skewness vs. task for control subjects. p-values significance codes: $. = 0.1, * =< 0.05, ** =< 0.01, *** =< 0.001$. The skewness of hitting-time distribution for control subjects is significantly smaller when the subjects are engaged in BART task compared to rest. The 10 nodes with largest hitting time changes are (B) V2, V3, V4, V3A and PGs (visual), (C) 1, 4, 3b and OP4 (motor) and (D) PF. The size of each node represents the magnitude of difference of average to-hitting times (range from 19 to 30.2) and thickness of each edge represents the magnitude of difference of partial correlation in BART compared to rest.

N_i neighbors and nodes i_1 and i_2 have the strongest connections to node i, we define chain index for node i as:

$$\zeta_i = \rho_{ii_1} + \rho_{ii_2} - \sum_{j=3}^{N_i} \rho_{ii_j}$$

If a node is located on a perfect chain, the chain index will be at its maximum. If a node has equal connections to every other node in the network (the least chain-like network), the chain index will be a large negative number that depends on the number of nodes in the network.

Comparing connectivity during the BART to rest, the largest reductions in the chain index were for a set of nodes that overlapped heavily with those nodes with the largest hitting time (8 of the 10 nodes described above), including 'V2', 'V3', 'V3A', 'V6' and 'V6A' within the visual cortex, '1', '4', '3b' and 'OP4' within somatosensory cortex, and the area 'PF'. Labels are according to Glasser et al. (2016). In accordance, the nodes with the largest hitting-time changes also show increased connectivity during the BART. Connections from nodes with the largest hitting times that have changed significantly (paired t-tests comparing connections from each node during task and rest, $p < 0.05$ Bonferroni corrected) are indicated in Fig. 5 by gray lines, with the thickness of the line indicating the size of the change. During task performance, nodes in task-related sensory streams (which have large hitting times during rest) have smaller hitting times, becoming less like isolated chains and instead more widely integrated.

Does the characterization of network topology have value in understanding and diagnosing psychiatric disorders? Brain network efficiency, which is commonly defined by the mean distance between nodes, has been shown to change in disorders such as Alzheimer's Dennis and Thompson (2014), schizophrenia Besnard et al. (2018); Li et al. (2017), and others Cheng et al. (2016). Reductions in measures of efficiency that utilize mean distance could be due to: 1. Reduced overall connectivity; or, 2. Subnetwork changes that lead to skewed hitting-time distributions. As described in Fig. 3, we can distinguish these possibilities by focusing on skewness. If skewness changes, the differences between psychiatric populations and controls is more likely to be due to subnetwork changes than a change in over all connectivity. We ran an ordinary least squares regression model with skewness of the resting-state hitting-time distribution as the dependent variable, and group (dummy coded, reference controls), gender (dummy coded, reference females), and age (mean centered, linear) as independent variables. We analyzed resting-state functional data from four patient groups, control, schizophrenia, bipolar and attention deficit hyperactivity disorder (ADHD). We found significant differences in skewness between schizophrenia and control populations ($\beta = -5.130, t(252) = 1.268, p < 0.001$), bipolar and control populations ($\beta = -4.060, t(252) = 1.324, p < 0.001$), and a trend-level difference between ADHD patients and control populations ($\beta = -2.445, t(252) = 1.324, p = 0.066$), see Fig. 6. Gender and age did not significantly explain variance in this model.

To identify those nodes related to skewness differences between schizophrenia, bipolar disorder, and controls, we first ask which nodes had the largest hitting-time changes. The ten regions with the largest

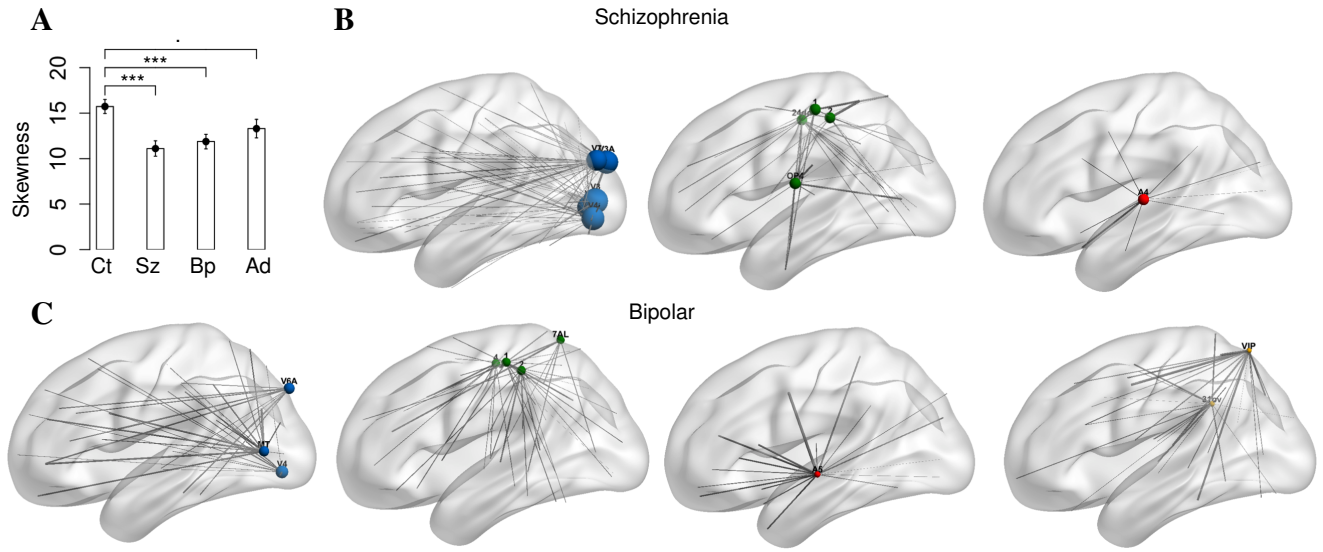

Figure 6: Skewness of the hitting-time distribution is significantly different across patient groups. (A) Distribution of skewness of the hitting-time distributions in patient and control groups during resting-state scans. Ct, Sz, Bp and Ad stand for control, schizophrenia, bipolar and ADHD, respectively. Significance codes: $. = 0.1, * =< 0.05, ** =< 0.01, *** =< 0.001$. The skewness of hitting-time distribution is significantly smaller for subjects with schizophrenia and bipolar disorders compared to neurotypicals. The 10 nodes with the largest change in hitting time for subjects with schizophrenia include V2, V3, V4, V3A and V7 (visual, top blue), 1, 2, 24dd and 4 (motor, top green) and A4 (auditory, top red). The 10 nodes with largest change of hitting time for subjects with bipolar include MT, V4, V3A and V6A (visual, bottom blue), 1, 2, 4 and 7AL (motor, bottom green), A5 (auditory, bottom red) and VIP and 31pv (bottom yellow). The size of each node represents the magnitude of difference of average to-hitting times (range from 7.1 to 33.4) and thickness of each edge represents the magnitude of difference of partial correlation in psychiatric disorder compared to control subjects during rest.

hitting-time changes between (t-tests comparing schizophrenia and control groups for each region, significantly different with $p < 0.05$ Bonferroni corrected) were 'V2', 'V3', 'V4', 'V3A' and 'V7' within the visual cortex, '1', '2', '4' and '24dd' within somatosensory cortex and 'A4' within the auditory cortex for subjects with schizophrenia. For subjects with bipolar disorder, the ten regions with the largest hitting-time changes (significantly different in t-tests comparing bipolar and control groups with $p < 0.05$ Bonferroni corrected) were 'MT', 'V4', 'V3A' and 'V6A' within visual cortex, '1', '2', '4' and '7AL' within somatosensory cortex, 'A5' within auditory cortex and 'VIP' and '31pv'. Regions are labeled

according to Glasser et al. (2016). Connections that have changed significantly (t-tests comparing connections from each node between groups, $p < 0.05$ Bonferroni corrected) from these nodes are indicated by gray lines with the thickness of the line indicating the size of the change.

We find evidence that individuals with schizophrenia and bipolar disorder have less skewed hitting-time distributions than controls during resting-state fMRI. The regions of cortex with the largest hitting-time reductions between patient and control populations are in sensory/motor cortex and overlap with many of the same regions that have extremely large to-hitting times during rest in controls. It is possible that the large hitting-time values found in these regions of the cortex are related to the theoretical finding that linear-chains of nodes produce maximal hitting times and that a reduction in hitting time in these regions occurs when the nodes are part of a topology that is less chain like. We therefore also compared the chain index for schizophrenia, bipolar and control groups. The ten regions with the largest changes of chain index (t-tests comparing schizophrenia and control groups for each region, significantly different with $p < 0.05$ Bonferroni corrected) are 'V2', 'V3', 'V4', 'V3A' and 'V7' within the visual cortex, '1', 'OP4' and '24dd' within somatosensory cortex and 'A4' and 'A5' within the auditory cortex for subjects with schizophrenia. For subjects with bipolar disorder, the ten regions with the largest changes of chain index (significantly different in t-tests comparing bipolar and control groups with $p < 0.05$ Bonferroni corrected) are 'V4', 'V3A' and 'V6A' within visual cortex, '1', '2', '4' and '7AL' within somatosensory cortex, 'A5' within auditory cortex and 'VIP' and 'PFcm'. The convergence of evidence of changes in extreme hitting-time values in sensory areas of cortex and those same areas being connected in a less chain-like topology in schizophrenia and bipolar disorder is consistent with our hypothesis that path-length changes in these populations are likely to be related to sub-network topology shifts and not changes the network on average.

DISCUSSION

We have presented evidence that the skewness of connectivity in cortical brain networks can be used to infer likely network topology changes that improve our understanding of information processing in the brain. Using random graphs we showed that the isolation of linear motifs is one prominent cause of skewness in hitting time even in the presence of mixed topologies. We then showed that skewness, but not average of brain-connectivity distributions is related to psychiatric diagnosis. We confirmed that these differences in skewness were related to a linear-chain topology by testing for changes in a chain index. In

networks for which connectivity is positively skewed, a change in subnetwork topology (and possibly a different brain state) is a more parsimonious explanation than changes in average connectivity. These topology changes are focused in sensory areas of the brain and when compared to changes brought about during task performance, provide an initial mechanistic link between resting-state connectivity changes and clinical diagnosis.

Extremely large hitting times can be linked to linear chain topologies through theoretical work showing that lollipop networks result in maximal hitting times Brightwell and Winkler (1990). We have shown using a toy problem that the extremeness of hitting-time values scales with how isolated a linear chain of nodes is and that the presence of chain motifs is related to extreme hitting-time values even in random networks with mixed topologies above and beyond modularity of the network. Resting brain networks also have extreme hitting times that are likely related to hierarchical processing in sensory cortex. When they are most isolated from the rest of the network, hierarchical processing streams, resemble the chain of nodes on the linear component of a lollipop graph. This parallel motivates the use of extremely long random walks between brain areas as a measure of the presence, and relative isolation, of hierarchical processing streams. In particular, hitting time of a region of interest can be utilized to detect presence of linear components likely to be hierarchical processing streams. A central dogma of neuroscience is that sensory representations are constructed hierarchically Hubel and Wiese (1962); Kikuchi, Horwitz, and Mishkin (2010); Van Essen and Maunsell (1983). Hierarchical processing streams have also been a focal component of computer vision models since 1971 Giebel (1971) and a significant contributor to the success of modern convolutional neural networks LeCun, Bengio, and Hinton (2015). Their foundational nature has made the study of hierarchical processing streams the focus of targeted analyses (e.g. Sepulcre, Sabuncu, Yeo, Liu, and Johnson (2012)). Here, we showed that nodes from sensory and motor areas of the brain have extreme hitting times which contribute to Kelley skewness. By comparing resting to task-based network topologies, we can show that decreases in hitting time are also associated with sensory hierarchies becoming less chain like.

Task changes in functional connectivity can be observed as changes in functional brain network topology. Previous work has highlighted the central role of path length and the integration of isolated paths in function. In Goni et al. Goñi et al. (2014), a notion of path transitivity – which accounted not only for the shortest path, but also local detours along that path – was the best predictor of functional connectivity.

Similar notions of distributed communicability (related to the diffusion of information over the network) were used to quantify the disruption of the global communication in the cortex that was triggered by the pharmacogenetic inactivation of the amygdala Grayson et al. (2016), and to detect changes in functional connectivity after a stroke Crofts et al. (2011). A thorough review of these concepts and their relationships to the notion of mean first-passage time, which is equivalent to the mean hitting time, is provided in Avena-Koenigsberger et al. Avena-Koenigsberger et al. (2017). Because the mean hitting time conflates overall changes in connectivity with changes to a subnetwork, these changes may be better explained using Kelley skewness, which specifically focuses on extreme values and so provides a mechanism to identify potential subnetwork changes. Extreme values contributing to connectivity skewness were associated with sensory areas of the brain specifically associated with the task. Hitting times in these sensory areas of the brain became shorter and less extreme during BART task performance. Sensory areas became more strongly connected to distant areas throughout the brain. The nodes with the largest reductions in hitting time were found in brain sensory areas related to the BART task (somatosensory and visual areas). One additional area also showed a decrease in hitting time, PF. PF is located in the inferior parietal lobule and is thought to be related to risk processing Weber and Huettel (2008). In line with its role in the processing of visual magnitude, it showed increased connectivity with visual inputs. Broadly, the introduction of a task caused sensory processing streams to become better connected to other task relevant areas, and less chain-like. This result is somewhat counterintuitive since hierarchical processing pipelines are often described as most distinct when active. We do find evidence of increased connectivity within sensory networks but these increases in strength within the sensory processing stream are offset by wider integration making their network topology less chain like. In addition, whether the processing pipeline becomes more or less integrated depends on the calculation underlying the transition probability between areas. When similar models are constructed using the raw correlation values, group and task differences in the same data set are consistent but in the opposite direction Rezaeinia and Carter (2017). This is likely due to the redundant connections and task event correlations Cole et al. (2018) included in raw correlation models. Here, we focused on the partial correlation of brain-region time-series which minimizes redundant connections, following the state-of-the-art in the field Smith et al. (2011). In spite of the complexities raised, the distribution of task relevant information throughout the brain is consistent with what would be most likely to improve BART task performance and supports the interpretation of sensory hierarchies as existing as relatively isolated linear networks topologies during rest. In future work it would be helpful to

incorporate additional multifaceted tasks to generalize these findings to the incorporation of sensory information under other constraints.

Skewness of the connectivity distribution also explains cortical-network differences between psychiatric diagnoses. Resting functional connectivity in a large neurotypical population is significantly positively skewed. In such a case, average efficiency for such a network would be biased and less representative of the network as a whole. An important finding is that changes in connectivity between clinical and control populations are due to changes in skewness rather than average differences. In fact, the median of connectivity measures changed in the opposite direction with respect to the average. Network connectivity changes between clinical and control populations are therefore due to a subset of connections rather than the network as a whole. The identification of specific cortical regions involved in topological changes between neurotypical and clinical populations provides an opportunity to better understand functional changes that occur in those populations as well as opportunities for improving diagnosis. Task performance reduced hitting-time skewness by increasing the connectivity between sensory areas and the rest of the cortical network. Qualitatively similar changes are seen in clinical populations, implying further work exploring network topology changes to specific tasks may help characterize the atypical resting connectivity for individuals with a schizophrenia or bipolar diagnosis. A testable prediction from this implied mechanistic difference would be that individuals with these diagnoses spend less time in activities typically associated with resting fMRI (e.g. future planning).

The results presented were based on theoretical predictions and applied to a publicly available dataset in a rigorous manner. We would like to document the following caveats and qualifications. First, although linear components in networks produce maximal hitting times in theory, it is possible that other network topologies could also produce some degree of skewness. To answer this concern, we showed that in toy examples skewness was related to short linear graphs or linear graphs that were still connected to the rest of the graph. We also found that the hitting-time distribution of both small-world, lollipop, and random graphs are skewed but that skewness in both cases is dominated by linear paths, even with a linear-path that comprises significantly less than a third of the graph. It is, however, important to note that the number of linear topologies does not explain all of the variance in Kelley skewness and so there could be other contributing topologies, perhaps related to modularity and degree. In our human fMRI analyses, the network as a whole, did not become less connected (see supplementary materials) and those areas with

extreme values become less chain-like when they have smaller hitting times. Thus, the relationship between extreme hitting times and linear paths is robust. In addition, past work on sensory hierarchies Hubel and Wiese (1962); Kikuchi et al. (2010); Van Essen and Maunsell (1983), recent work showing parallels between convolutional neural networks and sensory networks Güçlü and van Gerven (2015); Kell, Yamins, Shook, Norman-Haignere, and McDermott (2018); Khaligh-Razavi and Kriegeskorte (2014), and the integration of sensory networks during task performance (see above) are all consistent with the presence of linear components in the cortical network. Second, we focused on a cortical model of brain function and the absence of subcortical nodes could have affected the topology of the network model. However, the inclusion of subcortical connections to cortical network endpoints should not change connectivity measures since the path would still produce larger hitting times (the largest times may then be shifted to the middle of the sensory hierarchy).

CONCLUSION

In conclusion, establishing a link between network topologies, hierarchical processing pipelines, task engagement, and psychiatric disorders provides an opportunity to interpret cortical network changes in the light of cognitive models of brain function. The interpretation of network connectivity and information processing topologies is an area of significant focus for neuroscience Fornito and Bullmore (2015). The widespread collection of rfMRI in particular provides a unique opportunity to extend this work to numerous psychiatric disorders and compare these findings with the growing body of open-source fMRI task data.

SUPPORTIVE INFORMATION

Data.

We used the functional magnetic resonance imaging (fMRI) data from the LA5c Study Poldrack et al. (2016), collected by the UCLA Consortium for Neuropsychiatric Phenomics (CNP), which is funded by the NIH Roadmap Initiative. This data was obtained from the OpenfMRI database. Its accession number is ds000030. The dataset is formatted according to the Brain Imaging Data Structure K. J. Gorgolewski et al. (2016) (BIDS) standard. This study contains neuroimaging data from 290 participants. We ended up with a sample number of 255 subjects after removing subjects with missing functional measurements. In this sample, there are 119 healthy individuals (labeled as control), 49 individuals diagnosed with schizophrenia,

450 48 individuals diagnosed with bipolar disorder and lastly 39 individuals diagnosed with attention deficit
451 hyperactivity disorder (ADHD). The focus of the LA5c Study is to understand memory and cognitive
452 functional structures across patient groups. Therefore, the data set includes resting-state fMRI data as well
453 as fMRI data collected during several different tasks. In this paper we focus on the analysis of resting-state
454 and balloon analogue risk task (BART) fMRI data for the four specified groups. We focused on the BART
455 because it has reliable visual input which has been shown to produce large-scale changes in sensory
456 cortical hierarchies, which could be compared to the internally-driven resting-state network. In addition,
457 we develop variants of the BART and so we have expertise with the task and any findings would have local
458 applications. We obtained the preprocessed data which is de-identified, motion corrected and coregistered
459 to Montreal Neurological Institute (MNI) standard space K. Gorgolewski, Durnez, and Poldrack (2017).
460 We used FSL Jenkinson, Beckmann, Behrens, Woolrich, and Smith (2012); Smith et al. (2004) to correct
461 for motion and apply a high-pass filter to remove low-frequency noise (cut-off frequency of 100 seconds).
462 The data also includes potential confound regressors K. Gorgolewski et al. (2017). In order to remove the
463 effect of motion artifacts, we used the 36-parameter motion regression technique introduced in Satter et al.
464 Satterthwaite et al. (2013), which has been shown to be most effective in decoupling modular structure
465 from subject motion Ciric et al. (2017).

466 ***Parcellation.***

467 The fMRI data has the following parameters; the matrix consists of 64×64 voxels for 34 slices recorded
468 with a TR of 2 seconds. In total, there are 147 time samples for resting-state data and 267 time samples for
469 BART data. To extract a reliable cortical network for each participant, we reduced the number of nodes by
470 averaging over voxels within an anatomical region. We used the multi-modal parcellation developed by
471 Glasser, et al. Glasser et al. (2016) to map the fMRI data into a more sparse decomposition framework.
472 There are 180 regions in each hemisphere, which increase the neuroanatomical precision for studying the
473 structural and functional organization. This parcellation is based on multiple neurobiological properties,
474 connectivity, functional and architecture, which improves the consistency across subjects.

HITTING TIME VS. DEGREE DISTRIBUTION

475 To compare the ability of hitting time and degree distributions to distinguish the presence of linear
476 components, we generated a random graph with $N = 50$ nodes in which a connection between each node

randomly occurred with $p = 0.6$. We then attached a linear component of length 1 to the graph and kept increasing its length by 1 for 20 iterations (Fig. 7A). Skewness of the hitting-time distribution increases as we increase the length of linear component (Fig. 7B, blue/solid). The skewness of degree distribution remains close to constant (Fig. 7B, black/dashed). The change in other features of degree distribution such as mean or median do not reliably reflect the linear component either. Depending on the relative size of network and linear component, we might observe change in the degree distribution, but it is not consistent and is not most closely related to the presence of linear components.

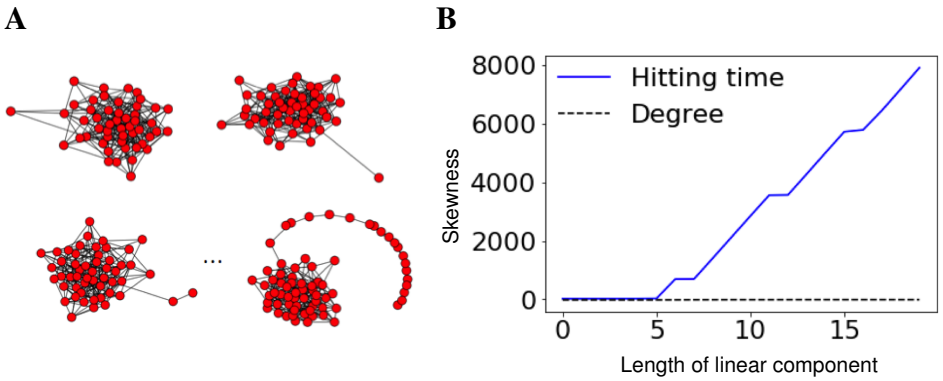

Figure 7: The hitting-time distribution of a graph becomes more skewed as the length of a linear subgraph increases. (A) We started with a random graph with 50 nodes, added a linear component of length 1 and increased the length of the linear component by 1 for 20 iterations. (B) Skewness of the hitting-time distribution increased significantly as the length of linear component increased. The degree distribution does not demonstrate a consistent relationship with linear path length.

MEAN AND MEDIAN OF HITTING-TIME DISTRIBUTION

Effect of task on mean and median of hitting-time distribution.

Aligned with the tests for skewness, we seek the effect of task on mean and median of hitting-time distribution. We ran a linear mixed effect model by adding a random effect for participant (BART and resting-state points were paired by participant). We found that mean of hitting-time distribution is significantly smaller for BART compared to rest ($\beta = -1.6, t(118) = 0.13, p < 0.001$).

Finally, a linear mixed effects model on the median of hitting times reveals a significant positive effect of BART vs. rest, ($\beta = 1.3, t(118) = 0.16, p < 0.001$), Fig. 8.

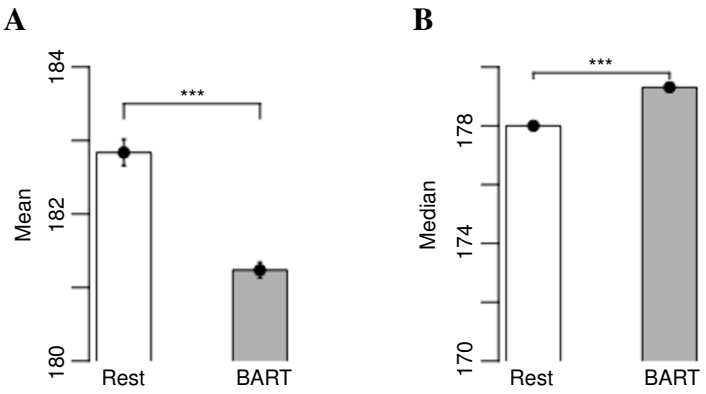

Figure 8: The distribution of (A) mean and (B) median vs. task for control subjects. p-values significance codes: $. = 0.1, * =< 0.05, ** =< 0.01, *** =< 0.001$.

Effect of psychiatric disorders on mean and median of hitting-time distribution.

To test the effect of psychiatric disorders on the average efficiency, we run an ordinary least squares regression model with mean of hitting-time distribution as the dependent variable. The independent variables are group (coded as a dummy variable with the control group as the reference group), gender (coded as a dummy variable with the female gender acted as the reference category) and age (mean centered, linear). We found significant differences between schizophrenia versus control ($\beta = -1.113, t(252) = 0.255, p < 0.001$), bipolar versus control ($\beta = -1.064, t(252) = 0.252, p < 0.001$) and ADHD patients versus control ($\beta = -0.835, t(252) = 0.266, p = 0.002$). The qualitative results are similar to skewness, but the coefficients are approximately one quarter the magnitude of those in the model explaining skewness. Finally, if we run a similar model with the median of the resting-state hitting-time distribution as the dependent variable, we only find a trend toward significance between schizophrenia and control in the opposite direction ($\beta = 0.473, t(252) = 0.244, p = 0.054$). Hierarchical sensory processing streams identified above as those with the longest hitting times show the largest changes between control and diagnosed groups, Fig. (9).

ACKNOWLEDGMENTS

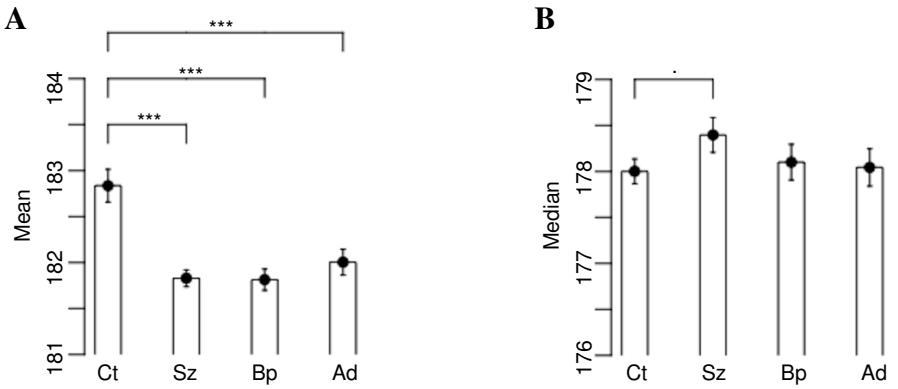

Figure 9: Mean of hitting-time distribution is significantly different across patient groups. Distribution (A) mean and (B) median of the hitting-time distribution in patient and control groups during resting-state scans. p-values significance codes: $. = 0.1, * =< 0.05, ** =< 0.01, *** =< 0.001$.

Publication of this article was funded by the University of Colorado Boulder Libraries Open Access Fund. The research was supported by the Department of Electrical and Computer Engineering, University of California San Diego, and Department of Electrical, Computer and Energy Engineering and Institute of Cognitive Science of the University of Colorado Boulder. Piya Pal was supported in parts by NSF NCS-FO 1734940 and UC San Diego. We would also like to acknowledge members of Social Neuroscience and Games (SNaG) Lab, and Psychology and Neuroscience Department at the University of Colorado Boulder whose comments helped us improve our work. We would like to especially thank Heejung Jung, John Pearson, Terry Sejnowski, and David Smith for manuscript review during its preparation.

514

515 **REFERENCES**

516

517 Achard, S., & Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS computational biology*,
518 *3*(2), e17.

519 Aldous, D., & Fill, J. (2002). *Reversible markov chains and random walks on graphs.*
520 https://www.stat.berkeley.edu/~aldous/RWG/book.html. Berkeley.

521 Avena-Koenigsberger, A., Misic, B., & Sporns, O. (2017). Communication dynamics in complex brain networks. *Nature*
522 *Reviews Neuroscience*, *19*, 17 EP -. Retrieved from http://dx.doi.org/10.1038/nrn.2017.149

523 Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*, 353 EP -. Retrieved from
524 http://dx.doi.org/10.1038/nn.4502

525 Besnard, J. L., Bassett, D. S., Smallwood, J., Margulies, D. S., Derntl, B., Gruber, O., ... Bzdok, D. (2018). Different shades
526 of default mode disturbance in schizophrenia: Subnodal covariance estimation in structure and function. *Human Brain*
527 *Mapping*, *39*(2), 644-661. Retrieved from
528 https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.23870 doi: 10.1002/hbm.23870

529 Brightwell, G., & Winkler, P. (1990). Maximum hitting time for random walks on graphs. *Random Structures & Algorithms*,
530 *1*(3), 263-276. Retrieved from
531 https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.3240010303 doi:
532 10.1002/rsa.3240010303

533 Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems.
534 *Nature reviews. Neuroscience*, *10*(3), 186.

535 Bullmore, E. T., & Bassett, D. S. (2011). Brain graphs: Graphical models of the human brain connectome. *Annual Review of*
536 *Clinical Psychology*, *7*(1), 113-140. Retrieved from
537 https://doi.org/10.1146/annurev-clinpsy-040510-143934 (PMID: 21128784) doi:
538 10.1146/annurev-clinpsy-040510-143934

539 Cheng, W., Rolls, E., Zhang, J., Sheng, W., Ma, L., Wan, L., ... Feng, J. (2016). Functional connectivity decreases in autism
540 in emotion, self, and face circuits identified by knowledge-based enrichment analysis. , *148*.

541 Cherniak, C. (1990). The bounded brain: Toward quantitative neuroanatomy. *Journal of cognitive neuroscience*, *2*(1), 58–68.

542 Ciric, R., Wolf, D. H., Power, J. D., Roalf, D. R., Baum, G. L., Ruparel, K., ... Satterthwaite, T. D. (2017). Benchmarking of
543 participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity.
544 *Neuroimage*, *154*, 174–187. doi: 10.1016/j.neuroimage.2017.03.020

545 Cole, M. W., Ito, T., Schultz, D., Mill, R., Chen, R., & Cocuzza, C. (2018). Task activations produce spurious but systematic
546 inflation of task functional connectivity estimates. *bioRxiv*. Retrieved from
547 https://www.biorxiv.org/content/early/2018/10/10/292045 doi: 10.1101/292045

548 Crofts, J. J., Higham, D. J., Bosnell, R., Jbabdi, S., Matthews, P. M., Behrens, T. E. J., & Johansen-Berg, H. (2011). Network
549 analysis detects changes in the contralesional hemisphere following stroke. *Neuroimage*, *54*(1), 161–169. doi:
550 10.1016/j.neuroimage.2010.08.032

551 Dennis, E., & Thompson, P. (2014). Functional brain connectivity using fmri in aging and alzheimer's disease. , *24*.

552 Erdős, P., & Rényi, A. (1959). On random graphs i. *Publicationes Mathematicae (Debrecen)*, *6*, 290-297.

553 Essen, D. C. V., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., & Ugurbil, K. (2013). The wu-minn human
554 connectome project: An overview. *NeuroImage*, *80*, 62 - 79. Retrieved from
555 http://www.sciencedirect.com/science/article/pii/S1053811913005351 (Mapping the
556 Connectome) doi: https://doi.org/10.1016/j.neuroimage.2013.05.041

557 Fornito, A., & Bullmore, E. T. (2015). Reconciling abnormalities of brain network structure and function in schizophrenia.
558 *Current opinion in neurobiology*, *30*, 44–50.

559 Fornito, A., Zalesky, A., & Bullmore, E. T. (Eds.). (2016). Chapter 1 - an introduction to brain networks. , 1 - 35. Retrieved
560 from http://www.sciencedirect.com/science/article/pii/B9780124079083000017 doi:
561 https://doi.org/10.1016/B978-0-12-407908-3.00001-7

562 Fox, M. D., & Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in systems*
563 *neuroscience*, *4*, 19.

564 Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature reviews. Neuroscience*, *11*(2), 127–138.

565 Giebel, H. (1971). Feature extraction and recognition of handwritten characters by homogeneous layers. In O.-J. Grüsser &
566 R. Klinke (Eds.), *Zeichenerkennung durch biologische und technische systeme / pattern recognition in biological and*
567 *technical systems* (pp. 162–169). Berlin, Heidelberg: Springer Berlin Heidelberg.

568 Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., … Van Essen, D. C. (2016). A
569 multi-modal parcellation of human cerebral cortex. *Nature*, *536*, 171 EP -. Retrieved from
570 http://dx.doi.org/10.1038/nature18933

571 Goñi, J., van den Heuvel, M. P., Avena-Koenigsberger, A., Velez de Mendizabal, N., Betzel, R. F., Griffa, A., … Sporns, O.
572 (2014). Resting-brain functional connectivity predicted by analytic measures of network communication. *Proceedings of*
573 *the National Academy of Sciences*, *111*(2), 833–838. Retrieved from
574 https://www.pnas.org/content/111/2/833 doi: 10.1073/pnas.1315529111

575 Gorgolewski, K., Durnez, J., & Poldrack, R. (2017). Preprocessed consortium for neuropsychiatric phenomics dataset [version
576 1; referees: 2 approved with reservations]. *F1000Research*, *6*(1262). doi: 10.12688/f1000research.11964.1

577 Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., … Poldrack, R. A. (2016). The brain
578 imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci Data*, *3*, 160044.
579 doi: 10.1038/sdata.2016.44

580 Goulas, A., Schaefer, A., & Margulies, D. S. (2015). The strength of weak connections in the macaque cortico-cortical
581 network. *Brain Struct Funct*, *220*(5), 2939–2951. doi: 10.1007/s00429-014-0836-3

582 Goñi, J., Avena-Koenigsberger, A., Velez de Mendizabal, N., van den Heuvel, M. P., Betzel, R. F., & Sporns, O. (2013).
583 Exploring the morphospace of communication efficiency in complex networks. *PLOS ONE*, *8*(3), 1-10. Retrieved from
584 https://doi.org/10.1371/journal.pone.0058070 doi: 10.1371/journal.pone.0058070

585 Grayson, D. S., Bliss-Moreau, E., Machado, C. J., Bennett, J., Shen, K., Grant, K. A., … Amaral, D. G. (2016). The rhesus
586 monkey connectome predicts disrupted functional networks resulting from pharmacogenetic inactivation of the amygdala.
587 *Neuron*, *91*(2), 453–466. doi: 10.1016/j.neuron.2016.06.005

588 Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network
589 analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, *100*(1), 253–258. Retrieved

590 from http://www.pnas.org/content/100/1/253 doi: 10.1073/pnas.0135058100

591 Güçlü, U., & van Gerven, M. A. J. (2015). Deep neural networks reveal a gradient in the complexity of neural representations
592 across the ventral stream. *Journal of Neuroscience*, *35*(27), 10005–10014. Retrieved from
593 http://www.jneurosci.org/content/35/27/10005 doi: 10.1523/JNEUROSCI.5023-14.2015

594 Hubel, D. H., & Wiese, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual
595 cortex. *J Physiol*, *160*, 106–154.

596 Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *NeuroImage*, *62*(2), 782–790.
597

598 Kell, A. J. E., Yamins, D. L. K., Shook, E. N., Norman-Haignere, S. V., & McDermott, J. H. (2018). A task-optimized neural
599 network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy. *Neuron*,
600 *98*(3), 630–644.e16. Retrieved from https://doi.org/10.1016/j.neuron.2018.03.044 doi:
601 10.1016/j.neuron.2018.03.044

602 Kelley, T. L. (1923). *Statistical method / by truman l. kelley* [Book]. The Macmillan Company New York.

603 Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain it cortical
604 representation. *PLOS Computational Biology*, *10*(11), 1-29. Retrieved from
605 https://doi.org/10.1371/journal.pcbi.1003915 doi: 10.1371/journal.pcbi.1003915

606 Khambhati, A. N., Medaglia, J. D., Karuza, E. A., Thompson-Schill, S. L., & Bassett, D. S. (2018). Subgraphs of functional
607 brain networks identify dynamical constraints of cognitive control. *PLOS Computational Biology*, *14*(7), 1-33. Retrieved
608 from https://doi.org/10.1371/journal.pcbi.1006234 doi: 10.1371/journal.pcbi.1006234

609 Kikuchi, Y., Horwitz, B., & Mishkin, M. (2010). Hierarchical auditory processing directed rostrally along the monkey's
610 supratemporal plane. *J Neurosci*, *30*(39), 13021–13030. doi: 10.1523/JNEUROSCI.2267-10.2010

611 Lambiotte, R., Delvenne, J., & Barahona, M. (2014). Random walks, markov processes and the multiscale modular
612 organization of complex networks. *IEEE Transactions on Network Science and Engineering*, *1*(2), 76-90. doi:
613 10.1109/TNSE.2015.2391998

614 Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Phys. Rev. Lett.*, *87*, 198701. Retrieved from
615 https://link.aps.org/doi/10.1103/PhysRevLett.87.198701 doi: 10.1103/PhysRevLett.87.198701

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*, 436 EP -. Retrieved from http://dx.doi.org/10.1038/nature14539

Li, P., Fan, T.-T., Zhao, R.-J., Han, Y., Shi, L., Sun, H.-Q., ... Lu, L. (2017). Altered brain network connectivity as a potential endophenotype of schizophrenia. *Scientific Reports*, *7*(1), 5483. Retrieved from https://doi.org/10.1038/s41598-017-05774-3 doi: 10.1038/s41598-017-05774-3

Lovász, L., & Simonovits, M. (1993). Random walks in a convex body and an improved volume algorithm. *Random Structures & Algorithms*, *4*(4), 359-412. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.3240040402 doi: 10.1002/rsa.3240040402

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev. E*, *69*, 026113. Retrieved from https://link.aps.org/doi/10.1103/PhysRevE.69.026113 doi: 10.1103/PhysRevE.69.026113

Norris, J. R. (1997). *Markov chains*. Cambridge University Press. doi: 10.1017/CBO9780511810633

Poldrack, R., Congdon, E., Triplett, W., Gorgolewski, K., Karlsgodt, K., Mumford, J., ... Bilder, R. (2016). A phenome-wide examination of neural and cognitive function. *bioRxiv*. Retrieved from http://biorxiv.org/content/early/2016/06/19/059733 doi: 10.1101/059733

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676–682. Retrieved from http://www.pnas.org/content/98/2/676 doi: 10.1073/pnas.98.2.676

Reijneveld, J. C., Ponten, S. C., Berendse, H. W., & Stam, C. J. (2007). The application of graph theoretical analysis to complex networks in the brain. *Clinical Neurophysiology*, *118*(11), 2317 - 2331. Retrieved from http://www.sciencedirect.com/science/article/pii/S1388245707004257 doi: https://doi.org/10.1016/j.clinph.2007.08.010

Rezaeinia, P., & Carter, R. M. (2017). Using hitting-time interdecile differences to identify brain networks with path-like features.

Rubinov, M., & Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *NeuroImage*,

642 *56*(4), 2068 - 2079. Retrieved from
643 http://www.sciencedirect.com/science/article/pii/S105381191100348X doi:
644 https://doi.org/10.1016/j.neuroimage.2011.03.069

645 Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., ... Wolf, D. H. (2013). An
646 improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state
647 functional connectivity data. *Neuroimage*, *64*, 240–256. doi: 10.1016/j.neuroimage.2012.08.052

648 Sepulcre, J., Sabuncu, M. R., Yeo, T. B., Liu, H., & Johnson, K. A. (2012). Stepwise connectivity of the modal cortex reveals
649 the multimodal organization of the human brain. *Journal of Neuroscience*, *32*(31), 10649–10661. Retrieved from
650 http://www.jneurosci.org/content/32/31/10649 doi: 10.1523/JNEUROSCI.0759-12.2012

651 Shen, X., & Meyer, F. (2008). Low dimensional embedding of fmri datasets. , *41*, 886-902.

652 Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... Beckmann, C. F. (2009). Correspondence
653 of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A*, *106*(31), 13040–13045. doi:
654 10.1073/pnas.0905267106

655 Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., ... Matthews, P. M.
656 (2004). Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage*, *23 Suppl 1*,
657 S208-19. doi: 10.1016/j.neuroimage.2004.07.051

658 Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., ... Woolrich, M. W. (2011).
659 Network modelling methods for fmri. *NeuroImage*, *54*(2), 875 - 891. Retrieved from
660 http://www.sciencedirect.com/science/article/pii/S1053811910011602 doi:
661 https://doi.org/10.1016/j.neuroimage.2010.08.063

662 Stiso, J., & Bassett, D. S. (2018). Spatial embedding imposes constraints on neuronal network architectures. *Trends in*
663 *Cognitive Sciences*, *22*(12), 1127 - 1142. Retrieved from
664 http://www.sciencedirect.com/science/article/pii/S1364661318302250 doi:
665 https://doi.org/10.1016/j.tics.2018.09.007

666 Trujillo-Ortiz, A., & Hernandez-Walls, R. (2003). D'agostino-pearson's k2 test for assessing normality of data using skewness
667 and kurtosis.

668 van den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: A review on resting-state fmri functional
669 connectivity. *European Neuropsychopharmacology*, *20*(8), 519 - 534. Retrieved from
670 http://www.sciencedirect.com/science/article/pii/S0924977X10000684 doi:
671 https://doi.org/10.1016/j.euroneuro.2010.03.008

672 Van Essen, D. C., & Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in*
673 *neurosciences*, *6*, 370–375.

674 Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *Nature*, *393*, 440 EP -. Retrieved from
675 http://dx.doi.org/10.1038/30918

676 Weber, B. J., & Huettel, S. A. (2008). The neural substrates of probabilistic and intertemporal decision making. *Brain Res*,
677 *1234*, 104–115. doi: 10.1016/j.brainres.2008.07.105

678 Zalesky, A., Fornito, A., & Bullmore, E. T. (2010). Network-based statistic: Identifying differences in brain networks.
679 *NeuroImage*, *53*(4), 1197 - 1207. Retrieved from
680 http://www.sciencedirect.com/science/article/pii/S1053811910008852 doi:
681 https://doi.org/10.1016/j.neuroimage.2010.06.041