Brain Struct Funct (2017) 222:603–618
DOI 10.1007/s00429-016-1238-5

ORIGINAL ARTICLE

Path ensembles and a tradeoff between communication efficiency and resilience in the human connectome

Andrea Avena-Koenigsberger[1] · Bratislav Mišić[1] · Robert X. D. Hawkins[2] · Alessandra Griffa[3,4] · Patric Hagmann[3,4] · Joaquín Goñi[5] · Olaf Sporns[1,6]

Received: 1 February 2016 / Accepted: 18 May 2016 / Published online: 22 June 2016

Abstract Computational analysis of communication efficiency of brain networks often relies on graph-theoretic measures based on the shortest paths between network nodes. Here, we explore a communication scheme that relaxes the assumption that information travels exclusively through optimally short paths. The scheme assumes that communication between a pair of brain regions may take place through a path ensemble comprising the k-shortest paths between those regions. To explore this approach, we map path ensembles in a set of anatomical brain networks derived from diffusion imaging and tractography. We show that while considering optimally short paths excludes a significant fraction of network connections from participating in communication, considering k-shortest path ensembles allows all connections in the network to contribute. Path ensembles enable us to assess the resilience of communication pathways between brain regions, by measuring the number of alternative, disjoint paths within the ensemble, and to compare generalized measures of path length and betweenness centrality to those that result when considering only the single shortest path between node pairs. Furthermore, we find a significant correlation, indicative of a trade-off, between communication efficiency and resilience of communication pathways in structural brain networks. Finally, we use k-shortest path ensembles to demonstrate hemispherical lateralization of efficiency and resilience.

Keywords Connectomics · Communication efficiency · Communication resilience

Electronic supplementary material The online version of this article (doi:10.1007/s00429-016-1238-5) contains supplementary material, which is available to authorized users.

✉ Joaquín Goñi
jgonicor@purdue.edu

✉ Olaf Sporns
osporns@indiana.edu

1 Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

2 Department of Psychology, Stanford University, Stanford, CA, USA

3 Signal Processing Lab., Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland

4 Department of Radiology, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland

5 School of Industrial Engineering and Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

6 IU Network Science Institute, Indiana University, Bloomington, IN, USA

Introduction

Human connectomics delivers network maps recording the patterns of structural connections among brain regions and systems (Sporns et al. 2005). Empirical studies and computational models suggest that the topology of structural connections constrains the flow of neural signals across the network (Passingham et al. 2002; Galán 2008; Honey et al. 2009; Park and Friston 2013; Hermundstad et al. 2013; Goñi et al. 2014; Mišić et al. 2015) and shapes the statistical dependencies among regional time courses of neuronal responses, generally captured in functional brain networks (Friston 2011). Furthermore, several studies have pointed out that the structural organization of the connectome optimizes the trade-off between network cost and competing functional demands, including efficient

communication (Bullmore and Sporns 2012; Vértes et al. 2012; Betzel et al. 2016).

Most extant approaches towards characterizing communication in brain networks rely on mapping the shortest paths between each pair of nodes (Bassett and Bullmore 2006; Achard and Bullmore 2007; van den Heuvel et al. 2009; Rubinov and Sporns 2010; van den Heuvel et al. 2012; Goñi et al. 2014). Measures of (average) shortest path length and its inverse, the global efficiency (Latora and Marchiori 2001), have been widely used to characterize differences in structural networks across individuals or groups (Gong et al. 2009; Yan et al. 2011; Lynall et al. 2010; Betzel et al. 2014). In addition, path-based measures of centrality have been used to identify nodes and edges that are thought to be most vital for global communication (Sporns et al. 2007; Joyce et al. 2010; Zuo et al. 2012). These measures assume that optimally short paths are highly privileged and are exclusively selected for signaling among remote node pairs. However, this presupposes that neural signals have access to information or "knowledge" about the global network topology (Boguña et al. 2009; Goñi et al. 2013; Abdelnour et al. 2014) which is unlikely to be the case. Furthermore, it excludes from consideration numerous alternative paths that, while not optimally short, may represent near-optimal alternative routes. Importantly, such paths could offer additional robustness and redundancy to network communication (Achard et al. 2006; Kaiser et al. 2007; Wook Yoo et al. 2015).

Indeed, in most real-world complex networks communication processes do unfold on numerous alternative paths (da Fontura Costa and Travieso 2007; Estrada and Hatano 2008). Computational analyses suggest that neuronal signals have the ability to recruit different alternative pathways in order to optimize performance, or to compensate for network impairment in brain damage (Kaiser et al. 2007; Wook Yoo et al. 2015). Various metrics have been developed with the aim of characterizing aspects of network communication that do not exclusively rely on shortest paths. Among these are approaches that take into account all possible paths (or walks) between nodes, for example the *mean first passage time* (Grinstead and Snell 2012), various spectral methods that reveal stationary distributions of stochastic processes (Chung 1997), random-walk betweenness centrality (Newman 2005), or communicability (Estrada and Hatano 2008). Each of these measures implies a specific model for network communication, and these models can be viewed as forming a continuous spectrum. On one end of the spectrum is the shortest-path framework, where the system has full "knowledge" of the global topology and information travels exclusively through the shortest possible path. The other end of the spectrum is represented by communication schemes in which elements of the system have no information about the global topology, and hence information travels randomly through the system. The latter case includes communication models that take into account all possible paths (or walks) between nodes.

In this paper, we explore a communication scheme that relaxes the assumption that information travels exclusively through optimally short paths without treating all possible paths as equally likely. We propose a model of communication between nodes that utilizes ensembles of paths composed of the k-shortest paths between nodes, with $k = 1$ corresponding to the classic shortest-path model. Furthermore, while communication through a given path ensemble generally privileges shorter over longer paths, we weigh the contributions of different paths within the ensemble according to the amount of information (in a local statistical sense) needed to travel such paths. We propose the extension of two topological measures, namely the shortest path length and the betweenness centrality, and re-define them in order to take into account the topology of k-shortest path ensembles between any given pair of nodes.

Using a set of high-resolution structural connectivity (SC) networks obtained from a cohort of 40 healthy individuals, we first study statistical features that characterize the connections and nodes associated with high values of edge betweenness centrality. Second, we study topological properties of subgraphs formed by the nodes and edges comprising path ensembles, focusing on the trade-offs between communication efficiency and resilience that result from considering multiple, non-optimally short paths as viable communication pathways between nodes. Finally, we quantify the extent to which path ensembles between anatomical regions promote integration and/or segregation of information flow and we examine differences that emerge between the two cortical hemispheres.

Materials and methods

Data set. Informed written consent in accordance with the Institutional guidelines (protocol approved by the Ethics Committee of Clinical Research of the Faculty of Biology and Medicine, University of Lausanne, Switzerland) was obtained for all subjects. Forty healthy subjects (16 females; 25.3 ± 4.9 years old) underwent an MRI session on a 3T Siemens Trio scanner with a 32-channel head coil. Magnetization prepared rapid acquisition with gradient echo (MPRAGE) sequence was 1-mm in-plane resolution and 1.2-mm slice thickness. DSI sequence included 128 diffusion weighted volumes + 1 reference b_0 volume, maximum b value 8000 s/mm^2, and 2.2 × 2.2 × 3.0 mm voxel size. EPI sequence was 3.3-mm in-plane resolution and 3.3-mm slice thickness with TR 1920 ms. DSI, resting-state fMRI, and MPRAGE data were processed using the

Connectome Mapping Toolkit (Daducci et al. 2012). Each participant's gray and white matter compartments were segmented from the MPRAGE volume. The grey matter volume was subdivided into 68 cortical and 15 subcortical anatomical regions, according to the Desikan-Killiany atlas (Desikan et al. 2006), defining 83 anatomical regions. These regions were hierarchically subdivided to obtain five parcellations, corresponding to five different scales (Cammoun et al. 2012). The present study uses a parcellation comprising 234 regions of interest (ROI); furthermore, we focus on cortical structures only, discarding all subcortical regions including the bilateral thalamus, caudate, putamen, pallidum, nucleus accumbens, hippocampus, and amygdala, as well as the brainstem, resulting in 219 remaining ROI. Whole brain streamline tractography was performed on reconstructed DSI data (Wedeen et al. 2008), and connectivity matrices were estimated from the streamlines connecting each pair of cortical ROI. We quantify the connection strength between each pair of regions as a fiber density (Hagmann et al. 2008) instead of fiber count. Thus, the connection weight between the pair of brain regions $\{u,v\}$ captures the average number of connections per unit surface between u and v, corrected by the length of the fibers connecting such brain regions. The aim of these corrections is to control for the variability in cortical region size and the linear bias toward longer fibers introduced by the tractography algorithm. Fiber densities were used to construct subject-wise structural connectivity (SC) matrices. Each SC matrix can be modeled as the adjacency matrix $A = \{a_{ij}\}$ of a graph $G \equiv \{V,E\}$, with nodes $V = \{v_1, \ldots, v_n\}$ representing ROI, and weighted, undirected edges $E = \{e_{v_i,v_j}, \ldots, e_{v_r,v_q}\}$ representing anatomical connections.

Intrinsic connectivity networks. In order to ground some of our structural connectivity-based findings in the context of the resting-state literature, we mapped the Desikan-Killiany anatomical parcels (Desikan et al. 2006) used to construct our SC networks, onto the *seven intrinsic connectivity networks* (ICN) defined by Yeo et al. (2011). The so-called Yeo parcellation was derived by using a clustering algorithm to partition the cerebral cortex of 1000 healthy subjects into networks of functionally coupled regions. This procedure resulted in the definition of seven clusters comprising networks previously described in the literature including the visual (VIS) and somatomotor (SM) regions, dorsal (DA) and ventral (VA) attention networks, frontoparietal control (FP), limbic (LIM) and default mode network (DMN). The mapping between the Desikan-Killiany anatomical parcels and the seven ICNs from the Yeo parcellation was obtained by extracting the vertices of the brain surface corresponding to each anatomical region in the Desikan-Killiany atlas, and then evaluating the mode of such vertices' assignment in the Yeo parcellation. Figure S1 in Online Resource shows the re-labeling of the Desikan-Killiany regions according to their corresponding label in the Yeo parcellation; the mapping from the 219 anatomical regions (obtained by randomly subdividing the Desikan-Killiany atlas) onto the 7-ICN from the Yeo parcellation is shown in Fig. S2 in Online Resource.

Finally, Fig. S3 in Online Resource shows an alternative ICN partition derived from our own data, by running a community detection algorithm on a group Functional Connectivity network. Analyses and results derived from this data-driven approach to extracting ICNs are shown in the Online Resource.

Wiring cost. The wiring cost of the network attempts to capture the metabolic cost of creating and maintaining the network connections by assuming that such cost is proportional to the volume of white matter in the brain (Bullmore and Sporns 2012). Hence, we express the cost of a single connection $\{u,v\}$ as the product between the number of streamlines and the average length of the streamlines between u and v. Then, the cost of a node is defined as the sum over the cost of all its connections.

k-Shortest paths. To rank a set of communication pathways between any pair of brain ROIs according to path length, a topological distance measure must first be defined over the set of edges in the network. Here, we use the function $l_{u,v} = -\log(w_{u,v})$ to transform connectivity strengths into connection lengths (as introduced by Goñi et al. 2014). This transformation ensures that if $w_{u,v} > w_{p,q}$ then $l_{u,v} < l_{p,q}$, that is, strong connection weights get mapped onto short distances, while weak connection strengths get mapped onto long distances and, $l_{u,v} = \infty$ if u and v are not connected. Then, the topological length $d(\pi_{u,v})$ of a path $\pi_{u,v}$ formed by a sequence of edges $\{e_{u,q_1}, \ldots, e_{q_r,v}\}$ is defined as the sum of the connection distances of the edge set forming the path. It is worth mentioning that the transformation from connection strengths to connection distances can be defined by different functions; it has been shown that different functions impose different distortions of the topology of the network (Simas and Rocha 2015). A widely used transform for brain networks has been $l_{u,v} = 1/w_{u,v}$, however, this transform yields extremely skewed distributions of the values $l_{u,v}$, with a few connections having extremely small values of $l_{u,v}$ compared to all other connections. As a consequence, these connections are highly prioritized to participate in the overwhelming majority of shortest paths. Here, we control for this skewed effect by selecting the transform $l_{u,v} = -\log(w_{u,v})$, which produces log-normal distributions of the values $l_{u,v}$ and avoids the extreme prioritization of a handful of connections.

Yen's algorithm (Yen 1971) was implemented to find the k-shortest (loop-less) paths between pairs of nodes in the network. This algorithm uses Dijkstra's algorithm to compute the first shortest path, and then finds candidates for the second shortest path by looking for detours around each edge in the original path. This is continued until the desired number of paths are found, giving an overall complexity of $O(kN^3)$, where N is the number of nodes in the network. Given the computational time required to compute k-shortest paths for a network of size $N = 219$, for this study we computed up to $k = 100$ shortest paths between every pair of nodes in the SC networks of 40 subjects.

k-Shortest path length. We define a composite measure to characterize the topological distance between a pair of ROIs, while taking into account the length $d(\pi^k_{s,t})$ of all k-shortest paths, and the embedding of such paths within the network. The measure is computed as follows:

$$D_k(s,t) = \sum \hat{P}\left(\pi^k_{s,t}\right) d(\pi^k_{s,t})$$

where $P\left(\pi^k_{s,t}\right) = \sum_{e_{u,v}\in\pi^k_{s,t}} \frac{w_{u,v}}{\sum_{w_{u,j}}}$ is the probability of following the kth shortest path from a source to a target node under random-walk dynamics;

$\hat{P}\left(\pi^k_{s,t}\right) = P\left(\pi^k_{s,t}\right) / \sum_{k=1}^{K} P(\pi^k_{s,t})$ is the normalized probability, such that $\sum_{k=1}^{K} \hat{P}\left(\pi^k_{s,t}\right) = 1$. Therefore, $D_k(s,t)$ is a weighted average of the lengths of the k-shortest paths from s to t, where the weights $\hat{P}\left(\pi^k_{s,t}\right)$ represent the ease with which a given path $\pi^k_{s,t}$ is traveled, relative to the other $k-1$ paths. A path is easy to travel when there are few potential detours from the path, that is, when nodes within the path have small degrees. Conversely, paths are difficult to travel if intermediate nodes are hubs because there is a higher probability of deviating from reaching the target node.

As an example, consider the unweighted graph shown in Fig. 1a. Notice that, because this graph is unweighted, there is degeneracy in the second shortest path, that is, the second shortest path is not unique but instead there are two paths ($\pi^2_{s,t}$ and $\pi^3_{s,t}$) with equal length. Nonetheless, $\pi^2_{s,t}$ is easier to travel because all nodes in the path have low degrees, whereas $\pi^3_{s,t}$ contains two high-degree nodes (nodes $\mathbf{v_2}$ and $\mathbf{v_3}$), increasing the difficulty of traveling the path. As a consequence, if we compute the distance $D_2(s,t)$ using $\pi^1_{s,t}$ and $\pi^2_{s,t}$ as the first and second shortest paths, we find that the contributions of both paths are comparable, resulting in $D_2(s,t) \approx 2.6$. On the other hand, if we use $\pi^3_{s,t}$ as the second shortest path, we get $D_2(s,t) \approx 2.1$ because the contribution of $\pi^3_{s,t}$ is significantly smaller. The interpretation of the degree sequence of a path as an indicator of how easily paths can be traveled is related to the work by Rosvall et al. (2005), where a measure called *search information* is proposed to quantify the information cost of traveling a specific path. Search information derived for weighted networks has been shown to be a significant predictor of functional connectivity (Goñi et al. 2014).

Finally, it is important to point out that for an undirected network ($w_{s,t} = w_{t,s}$), the kth shortest path between a pair $\{s,t\}$ is the same as the kth shortest path between $\{t,s\}$ (that is, $\pi^k_{s,t} = \pi^k_{t,s}$); however, the degree sequence of $\pi^k_{s,t}$ may

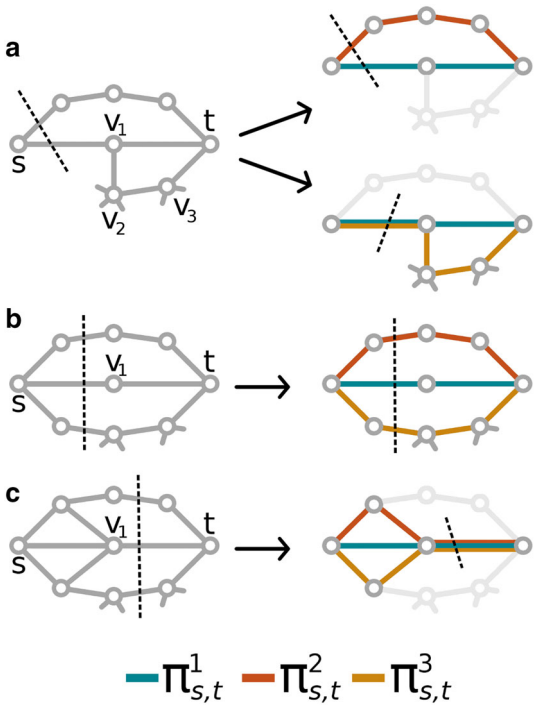

Fig. 1 Toy models illustrating subgrahs $G_{\pi^k}(s,t)$ formed by the k-shortest paths ($k = 2$ and $k = 3$) between nodes s and t. **a** Unweighted graph with two equally long second shortest paths: $\pi^2_{s,t}$ (*orange path*) and $\pi^3_{s,t}$ (*yellow path*) have path length 4. Nonetheless, $\pi^2_{s,t}$ is easier to travel because all nodes in the path have low degrees; conversely, $\pi^3_{s,t}$ contains two high-degree nodes (nodes $\mathbf{v_2}$ and $\mathbf{v_3}$) that increase the difficulty of traveling the path, given that the probability of a message getting "lost" or "detouring" from the path is higher. *Dashed lines* indicate partitions of the graph (*left panel*) and subgraphs (*right panel*) that would result in two disjoint subsets, cutting the communication between nodes s and t. **b** Unweighted graph and a path ensemble with 3-shortest paths between nodes s and t. None of the three shortest paths have edges in common, therefore, they are edge-disjoint paths and $F_3(s,t) = 3$. *Dashed lines* indicate partitions of the graph (*left panel*) and the subgraph $G_{\pi^3}(s,t)$ (*right panel*) that would result in two disjoint subsets, cutting the communication between nodes s and t. Notice that $F_3(s,t) = F_{\max}(s,t) = 3$. **c** Unweighted graph in which the three shortest paths between nodes s and t contain the edge $\{\mathbf{v_1}, t\}$, hence, the subgraph $G_{\pi^3}(s,t)$ contains a single edge-disjoint path ($F_3(s,t) = 1$), as indicated by the *dashed line* intersecting a single edge used by all three shortest paths. The *dashed line* crossing three edges in the graph (*left panel*) indicates that $F_{\max}(s,t) = 3$, and therefore $|F_3(s,t)| = F_3(s,t)/F_{\max}(s,t) = 1/3$

differ from the degree sequence of $\pi^k_{t,s}$, and then $\hat{P}\left(\pi^k_{s,t}\right)$ may be different from $\hat{P}\left(\pi^k_{t,s}\right)$. Hence, we symmetrize D_k as follows: $\hat{D}_k(s,t) = (D_k(s,t) + D_k(t,s))/2$.

Number of edge-disjoint paths. Given a graph G, a set of edge-disjoint paths between a pair of nodes $\{s,t\}$ consists of a set of paths connecting s and t, such that no path shares an edge with another path in the set (Biggs et al. 1976). Furthermore, the maximum number of edge-disjoint paths between a pair of nodes $\{s, t\}$ is equal to the minimum number of edges that must be removed from the network in order to disconnect nodes s and t. To calculate $F_{\max}(s,t)$, the maximum number of edge-disjoint paths between any pair of nodes $\{s,t\}$, we implement the Ford-Fulkerson (Ford and Fulkerson 1987; Cormen et al. 2001) method on the binarized graph G. For this work, we also use this method to compute the number of edge-disjoint paths F_k between a pair of nodes $\{s,t\}$ with respect to a subgraph $G_{\pi^k}(s,t)$, which is obtained from G, by selecting all nodes and edges that participate in any one of the k-shortest paths between $\{s,t\}$. Therefore, given a subgraph $G_{\pi^k}(s,t) \subset G$ the number of edge-disjoint paths $F_k(s,t)$ depends on k (the number of k-shortest paths used to construct $G_{\pi^k}(s,t)$) and is bounded by the maximum number of edge-disjoint paths between $\{s, t\}$ when the entire graph G is considered ($F_k(s,t) \leq F_{\max}(s,t)$).

The unweighted graph in Fig. 1b shows that there are at most three edge-disjoint paths between nodes s and t, therefore at least three edges must be removed in order to cut the communication between s and t. However, the subgraph $G_{\pi^2}(s,t)$ formed by the set of paths $\{\pi^1, \pi^2\}$ has only two edge-disjoint paths, whereas the subgraph $G_{\pi^3}(s,t)$ formed by the set of paths $\{\pi^1, \pi^2, \pi^3\}$ has three edge-disjoint paths. Finally, the graph in Fig. 1c has at most three edge-disjoint paths between nodes s and t, nonetheless, both subgraphs, $G_{\pi^2}(s,t)$ and $G_{\pi^3}(s,t)$ have one edge-disjoint path because all three paths $\{\pi^1, \pi^2, \pi^3\}$ share the edge $\{\mathbf{v_1}, t\}$, and removing such an edge would disconnect s and t within any one of the subgraphs.

Directed vs. undirected networks. All of the measures presented here can be applied for both directed and undirected networks (the measure D_k should not be symmetrized for undirected networks). In this study, all networks are undirected and therefore, all measures are symmetric: for any given pair of nodes $\{s,t\}$, and some topological measure m, we have $m(s,t) = m(t,s)$. Thus, the notion of *source* and *target* nodes is not meaningful as it is in the context of directed networks. For the results reported in this paper, we use the more generic terms $\{i, j\}$ to refer to pairs of nodes in the network, with the understanding that both nodes are the source and target of any communication pathway between them.

Results

Path ensembles widen the participation of edges in communication processes

We study white matter communication pathways between pairs of cortical regions of interest (ROI) of human structural connectivity (SC) networks obtained from a cohort of 40 healthy subjects (see "Materials and methods"). For each subject's SC network we use Yen's algorithm (Yen 1971) to compute the k-shortest paths between all ROI pairs, where $k = 1$ corresponds to the shortest path and as k increases, paths get progressively longer. For each ROI pair $\{i, j\}$ we define ensembles $KSP(i,j;k) = \{\pi^1_{i,j}, \pi^2_{i,j}, \ldots, \pi^k_{i,j}\}$ where $\pi^k_{i,j}$ is the kth shortest path between i and j, and therefore, it is the longest path in the ensemble. For all results presented in this paper, we use $k = 100$ as the maximum ensemble size. Given a pair of ROI $\{i, j\}$, and the corresponding $KSP(i,j;k)$ ensemble, we define the subgraph $G_{\pi^k}(i,j)$ by selecting, from the SC network, the set of nodes and connections that participate in at least one path contained in the ensemble of size k. Notice that nodes and edges in the subgraph $G_{\pi^k}(i,j)$ may participate in more than one path.

It is important to note that under an optimal communication framework where information transmission between nodes takes place through shortest paths only, there is a significant fraction of structural edges that do not participate in any communication pathway. Figure 2 shows the

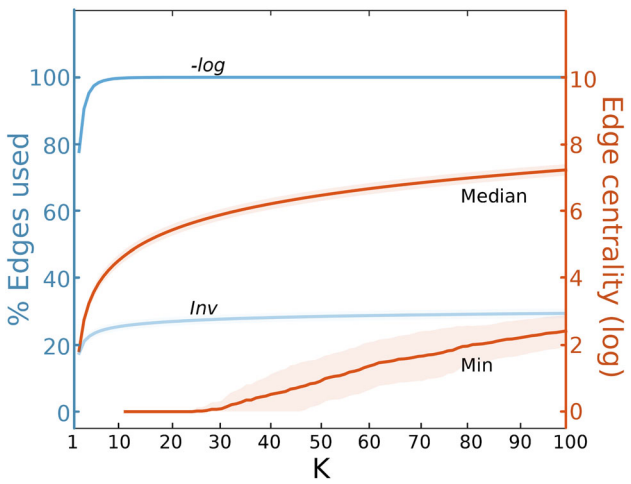

Fig. 2 Percentage of the connectome edges appearing in any ensemble (*blue lines*) as a function of the number of k-shortest paths contained in the path ensembles between all node pairs. The two *blue lines* correspond to two different functions mapping connection strengths to connection lengths: $l_{u,v} = -\log(w_{u,v})$ and $l_{u,v} = 1/w_{u,v}$, where $w_{u,v}$ are the connection strengths quantified as fiber densities. *Orange lines* show minimum and median values of k-shortest path edge betweenness centrality as a function of k. All values correspond to averages across all subjects

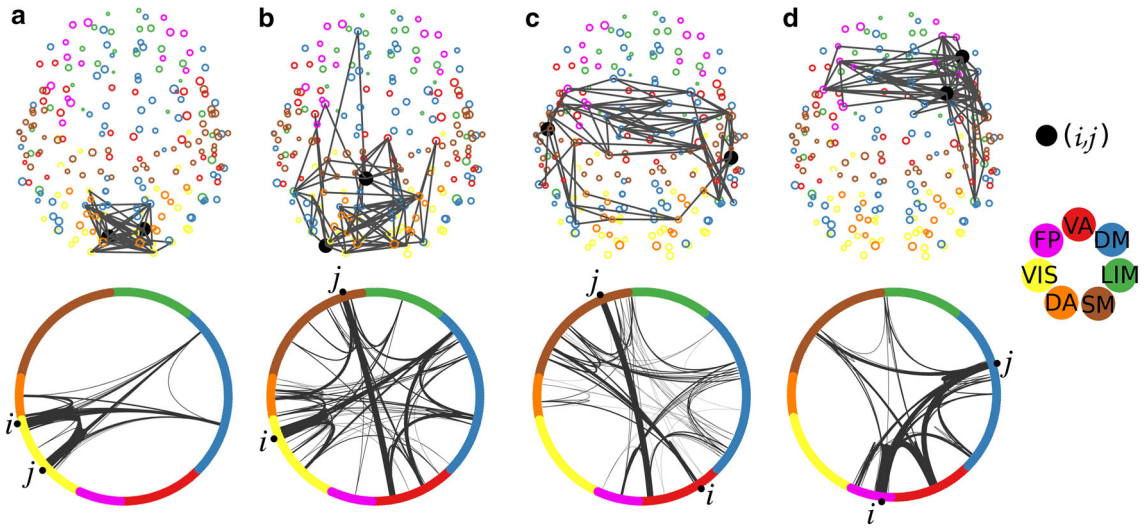

Fig. 3 Examples of subgraphs G_{π^k} obtained from four different *KSP* ensembles ($k = 100$) corresponding to pairs of nodes located within the following ICNs: visual–visual (**a**); visual–somatomotor (**b**); somatomotor–ventral attention (**c**); frontoparietal–default mode network (**d**). Figures in *upper row* display spatially embedded layouts of the subgraphs, illustrating the extent to which subgraphs may span extended portions of the network. Nodes are *colored* according to ICN membership. Figures in the *lower row* display a circular layout where nodes are arranging according to ICN membership, highlighting how connections within path ensembles are distributed within and between ICNs

percentage of the connectome edges used for communication (blue lines) as a function of the number of k-shortest paths contained in the $KSP(i,j;k)$ ensembles (these results show the average across all subjects). The two blue lines correspond to two different definitions of connection lengths: $l_{u,v} = -\log(w_{u,v})$ and $l_{u,v} = 1/w_{u,v}$, where $w_{u,v}$ are the connection weights (see "Materials and methods"). Notice how the $-Log$ line rapidly converges to 100 % of edges used, whereas the *Inv* line seems to asymptote towards a value between 20 and 30 % of edges used. Most important is the counterintuitive fact that the set of edges that do not participate in at least one k-shortest path (22 and 82 % of the edges for the $-Log$ and *Inv* functions, respectively, both for a value of $k = 1$) can be removed from the network without affecting classic graph measures such as the *characteristic path length*, *global efficiency*, and the node and edge *betweenness centrality*. In contrast, when considering path ensembles at increasing values of k, not only do all edges participate in at least one communication pathway, but the median number of times that edges get used in a path grows exponentially, while the total number of k-shortest paths increases linearly, as $kN(N-1)/2$. Hence, path ensembles ensure that communication load is more evenly distributed across the entire structural network and that individual edges participate in communication processes among multiple node pairs.

While communication pathways are invariably widened as the number of k-shortest paths increases, the structural patterns of the *KSP* ensembles vary across ROI pairs. Figure 3 shows some examples of subgraphs G_{π^k} that are obtained from four *KSP* ensembles corresponding to four different ROI pairs within the SC network of a single subject. A subgraph G_{π^k} only includes the nodes and edges that are contained in the corresponding *KSP* ensemble; nonetheless, to provide a more comprehensive picture of how subgraphs are embedded in the network, we show spatially embedded (Fig. 3, upper row) and circular (Fig. 3, lower row) layouts that display all the nodes of the network. The spatially embedded layouts illustrate the extent to which subgraphs may span extended portions of the network. Furthermore, by arranging nodes according to ICN membership, the circular layouts highlight how connections are distributed within and between ICNs.

The emergence of alternative high-centrality nodes and edges

Figure 2 suggests that the level of centrality of an edge or node may vary as a function of the number of k-shortest paths considered between every pair of ROI. Thus, we extend the definitions of edge and node betweenness centrality (from now on, referred to as edge centrality and node centrality, respectively) to quantify the number of k-shortest paths that traverse each edge and each node of the network, taking into account all k-shortest paths from any node to all others in the network. We normalize centrality values by computing the number of times an edge (node) is traversed by a k-shortest path, and dividing by $kN(N-1)/2$, where N is the number of nodes in the network and k is the number of k-shortest paths between any pair of

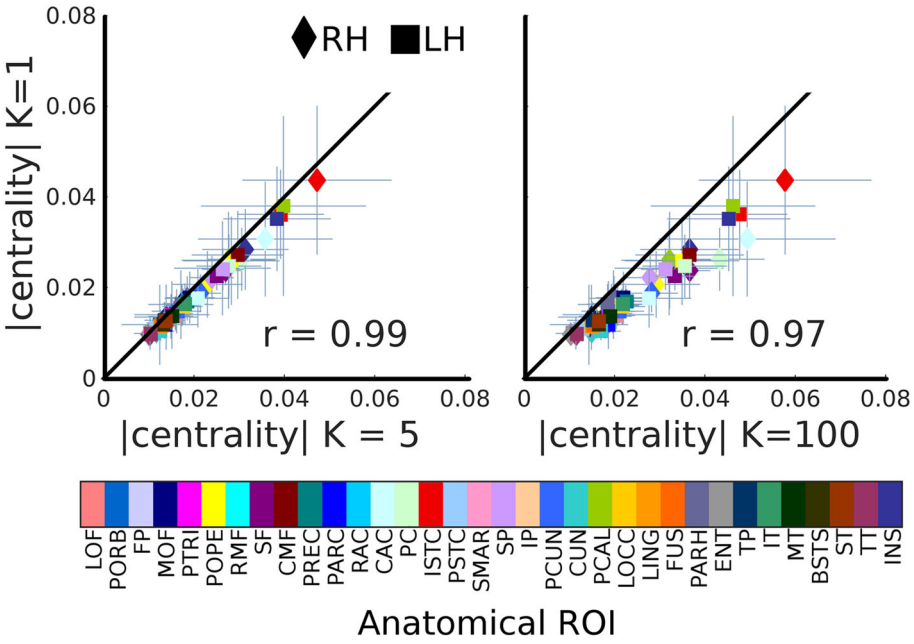

Fig. 4 Normalized shortest-path centrality ($k = 1$) versus the normalized k-shortest path centrality ($k = 5$ and $k = 100$) of 68 anatomical regions, covering the entire cortical surface. Diamond-shaped markers indicate right hemisphere anatomical regions, and squared-shaped markers indicate left hemispheric anatomical regions. All data points represent averages across all subjects; *vertical* and *horizontal lines* at each point indicate the standard deviation over all subjects

nodes. In addition, we assess the average centrality of anatomical regions by averaging across the centrality of all nodes belonging to a given anatomical region.

Figure 4 shows scatter plots of the normalized shortest-path centrality ($k = 1$) versus the normalized k-shortest path centrality (for $k = 5$ and 100) of 68 anatomical regions, covering the entire cortical surface (equivalent scatter plots at the individual node level are shown in Fig. S4 in Online Resource). While normalized centrality values are highly correlated with classic ($k = 1$) centrality (Pearson correlation coefficients are $r = 0.991$ for k = 5; $r = 0.972$ for $k = 100$; both $p < 0.01$; r values correspond to averages across all subjects), there are anatomical regions whose centrality rank changes as k increases. We find that the top five central regions for all values of k are right and left hemisphere isthmus of the cingulate cortex, left hemisphere pericalcarine cortex, right hemisphere caudal anterior cingulate cortex, and left hemisphere insula. While these are the most central regions across all values of k, their ranking within the top five does vary as a function of k. In addition, we identify other cortical regions that are among the ten lower ranked central regions when $k = 1$, and gain up to 16 positions within the ranking as k increases; the cuneus and paracentral lobule regions of both hemispheres are examples of regions that become significantly more central at higher values of k. A table with the ranking of all anatomical regions for different values of k is shown in Online Resource.

Changes in centrality ranking as a function of k are also shown in Fig. 5, which depicts the top 10 % central edges (Fig. 5a) and top 10 % central nodes (Fig. 5c) of a single subject for different values of k. Purple colored edges (nodes) ($b_1 \cap b_k$) indicate the set of edges (nodes) that are highly central according to both measures, the shortest-path ($k = 1$) centrality and the k-shortest path centrality; turquoise colored edges (nodes) (b_1–b_k) indicate the set of edges (nodes) that are highly central only for the shortest-path centrality; orange colored edges (nodes) (b_k–b_1) indicate the set of edges (nodes) that are highly central only for the k-shortest path centrality ($k > 1$). As k increases, centrality ranks of edges and nodes continue to deviate from their ranks at $k = 1$, resulting in progressively fewer edges in the intersection ($b_1 \cap b_k$).

In order to determine whether there are characteristic differences between edges and nodes belonging to the different sets ($b_1 \cap b_k$, b_1–b_k, b_k–b_1) we compare the wiring cost of nodes and edges, and the average degrees of the participating nodes. The degree associated with an edge $\{i, j\}$ is computed as the sum of the degrees of the two connected nodes i and j. Both, cost and degree are then averaged across all edges (nodes) within each set.

Figure 5b and 5d show distributions of average cost and average degree across all subjects, expressed as z scores relative to null distributions obtained for each subject, by randomly sampling sets of M_k edges (nodes), where M_k is equal to the number of edges (nodes) in the corresponding set. For all values of k, edges and nodes belonging to the set $b_1 \cap b_k$ have average cost and average degree that are significantly higher than chance, with a tendency for cost to increase for higher values of k. It is noteworthy that there is a significant difference between the wiring cost and degree distributions corresponding to the sets b_1–b_k and b_k–b_1

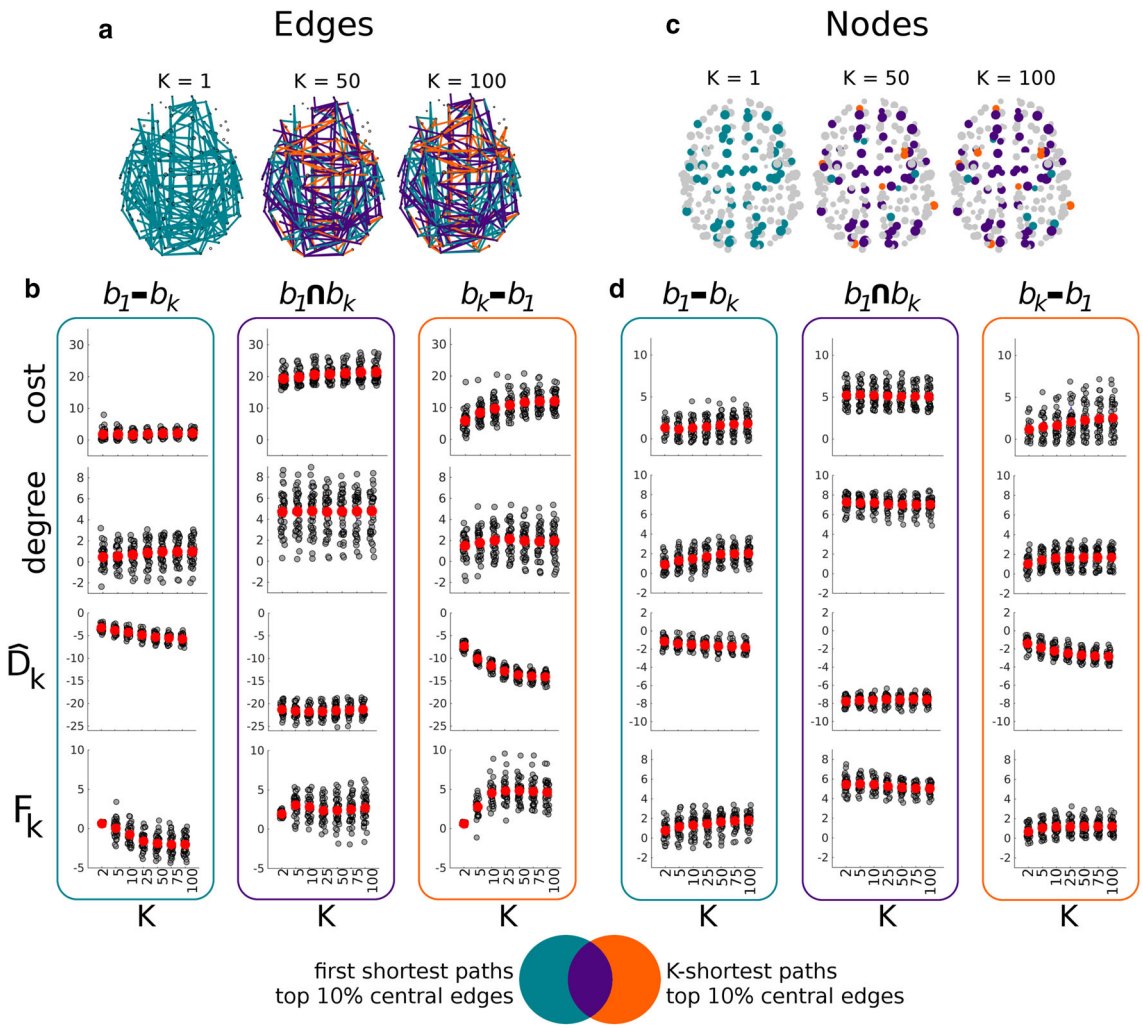

Fig. 5 Top 10 % central nodes (*right*) and edges (*left*) differ as a function of k. Edges can belong to one of the following three sets: $b_1 \cap b_k$ (*purple*), b_1–b_k (*turquoise*), b_k–b_1 (*orange*) where b_1 is the set of top 10 % central edges (nodes) only for $k = 1$, and b_k is the set of top 10 % central edges (nodes) for $k > 1$. **a** Top 10 % central edges for $k = 1$, 50 and 100. **b** Distributions of cost, degree, $\hat{D}_k$ and F_k expressed as z scores relative to null distributions obtained from randomly sampling edges from the empirical networks. **c** Top 10 % central nodes for $k = 1$, 50 and 100. **b** Distributions of cost, degree, $\hat{D}_k$ and F_k expressed as z scores relative to null distributions obtained from randomly sampling nodes from the empirical networks

($p < 0.01$, Kolmogorov–Smirnov test), for both edges and nodes, and across all values of k. These differences suggest that, when we consider alternative paths for communication between nodes by increasing k, we uncover additional high-cost edges and high-degree nodes, which do not appear as central if only shortest paths are considered.

We study additional properties of the top 10 % central edges and nodes by computing the composite k-shortest path length measure, $\hat{D}_k$ (see "Materials and methods"), and the number of edge-disjoint paths, F_k (see "Materials and methods"), corresponding to the $KSP(i,j:k)$ ensembles of all edges$\{i, j\}$ in the SC networks of all subjects. The bottom two rows of Fig. 5b show the distributions of $\hat{D}_k$ and F_k of the $KSP(i,j;k)$ ensembles where i and j correspond to all node pairs directly connected by the edges contained in the sets $b_1 \cap b_k$, b_1–b_k, b_k–b_1, respectively. $\hat{D}_k$ and F_k are expressed as z scores relative to null distributions obtained for each subject, by randomly sampling sets of edges from the networks (analogously to the null distributions built for cost and degree). The negative values in the distributions of $\hat{D}_k$ indicate that the *KSP* ensembles associated to highly central edges have significantly shorter k-shortest path lengths than expected by chance. The distributions of F_k corresponding to the set $b_1 \cap b_k$ and b_k–b_1 in Fig. 5b demonstrate that for $k \geq 5$, path ensembles have more edge-disjoint paths than expected by chance. These results indicate that nodes that are directly connected by highly central edges also have significantly more alternative (edge-disjoint) pathways to communicate, which we interpret as higher resilience. Interestingly, path ensembles

associated with edges that are highly central only at $k = 1$ have significantly fewer edge-disjoint paths, suggesting that the communication between the corresponding node pairs is less resilient.

A similar analysis is performed at the level of nodes (Fig. 5d). We define the k-shortest path closeness of a node i as $\hat{D}_k(i) = \frac{1}{N}\sum_{j=1}^{N} \hat{D}_k(i,j)$, and similarly, we can define the average number of edge-disjoint paths between a node i and all other nodes in the network as $F_k(i) = \frac{1}{N}\sum_{j=1}^{N} F_k(i,j)$. The distributions of $\hat{D}_k$ at the node level demonstrate that highly central nodes have significantly greater k-shortest path closeness than expected by chance. Furthermore, central nodes belonging to the set $b_1 \cap b_k$ also exhibit significantly higher resilience compared to other nodes.

A trade-off between efficiency and resiliency in path ensembles

We study additional topological properties of path ensembles that can give us some insight about the degree of communication efficiency between distinct ROI, given the admissibility of sub-optimal paths. To this end, for every pair of ROI in the network, we first examine how the lengths $d(\pi^k)$ of the kth shortest path in a *KSP* ensemble increase as a function of k; then, we compute $\hat{D}_k$ and F_k for all possible node pairs in a network; finally, we derive normalized versions of these measures as follows: $|\hat{D}_k|$ is defined as the ratio of $\hat{D}_k$ at values $k > 1$ with respect to the shortest-path length ($k = 1$); $|F_k|$ is defined as the ratio of F_k with respect to $F_{\max}$, where $F_{\max}$ is the maximum number of edge-disjoint paths between a pair of nodes, when the entire network is considered (see "Materials and methods").

Figure 6a shows distributions of the ratio of the kth shortest-path lengths and the ($k = 1$) shortest-path length (left panel). These distributions reveal two salient features. First, the median (averaged across all subjects, indicated by red points) rises gradually with increasing k to a value of 1.46 for $k = 100$. Second, while there are kth shortest paths whose lengths have increased significantly (in some cases the increase is a tenfold increase with respect to the optimal path length), we find that the composite distance measure $\hat{D}_k$ is not strongly affected by the inclusion of such long paths, given that the maximum value of the ratio between $\hat{D}_k$ and the shortest path length is 1.67 (average across all subjects) when $k = 100$ (for a single subject, and $k = 250$, the maximum value of the ratio between $\hat{D}_k$ and the shortest path length is 1.84). In other words, $\hat{D}_k$ deviates at most from the optimal path length by 67 %. This confirms the intuition that longer paths will make smaller contributions towards the composite distance measure $\hat{D}_k$, given that long paths tend to be more difficult to travel.

The number of edge-disjoint paths within a subgraph $G_{\pi^k}(i,j)$ is a non-decreasing function of k, however, this measure is also bounded by the minimum between the degrees of node i and node j. Interestingly, for many node pairs, including high-degree node pairs, the majority of their edge-disjoint paths are not ranked among the 100 shortest paths. The maximum value of F_k achieved at $k = 100$, is 34 (see Fig. 6a, right panel), whereas the total number of edge-disjoint paths available when the entire network is considered ($F_{\max}$), has a maximum value of 56. The fact that a large fraction of edge-disjoint paths have longer lengths that rank them below $k = 100$ is confirmed by the quartile values of $|F_k|$ evaluated over all node pairs: $q_1 = 0.50$, $q_2 = 0.56$, $q_3 = 0.65$, $q_4 = 0.98$ (for $k = 100$, average across all subjects).

Next, we investigate the relationship between the number of edge-disjoint paths, F_k, and the extent to which the composite distance measure $\hat{D}_k$ deviates from the optimal path length, as measured by $|\hat{D}_k|$. The scatter plots in Fig. 6b show both of these measures evaluated over all node pairs, averaged across all subjects. Interestingly, we find that both measures are significantly correlated (Pearson correlation coefficients are $r = 0.154$, 0.499, 0.506, and 0.470, for $k = 2$, 10, 50 and 100, respectively; $p < 0.01$). Minimizing the distance of a communication pathway is desirable because it allows faster and less noisy communication while favoring low wiring cost. However, if distance minimization constrains multiple pathways to traverse the same set of short structural edges, then the path ensemble becomes vulnerable to edge failure or damage. Hence, path ensembles that "branch-out" gain robustness and resilience but at the expense of becoming less efficient. The relationship between these two measures may thus be interpreted as a trade-off. Finally, it is worth noting that for values of $k < 3$, path ensembles can only have a maximum of 2 edge-disjoint paths; this constraint over the values of F_k yields a smaller correlation between F_k and $|\hat{D}_k|$.

We now turn to study these topological measures of the *KSP* ensembles at the ICN level by computing, for each measure, the median across all node pairs between and within the seven ICNs defined in Yeo et al. (2011). The scatter plot in Fig. 6c shows F_k and $|\hat{D}_k|$ median values for $k = 100$, averaged across all subjects. Once again, we find that the two measures are highly correlated (average Pearson correlation coefficients across subjects are r = 0.535, 0.637, 0.697, and 0.651, for $k = 2$, 10, 50 and 100, respectively; $p < 0.01$). Furthermore, we find that data points corresponding to pathways within the visual (VIS), frontoparietal (FP) and dorsal attention (DA) networks have higher values of F_k and/or $|\hat{D}_k|$, compared to other ICNs. This finding suggests topological differences among the communication pathways within these functional

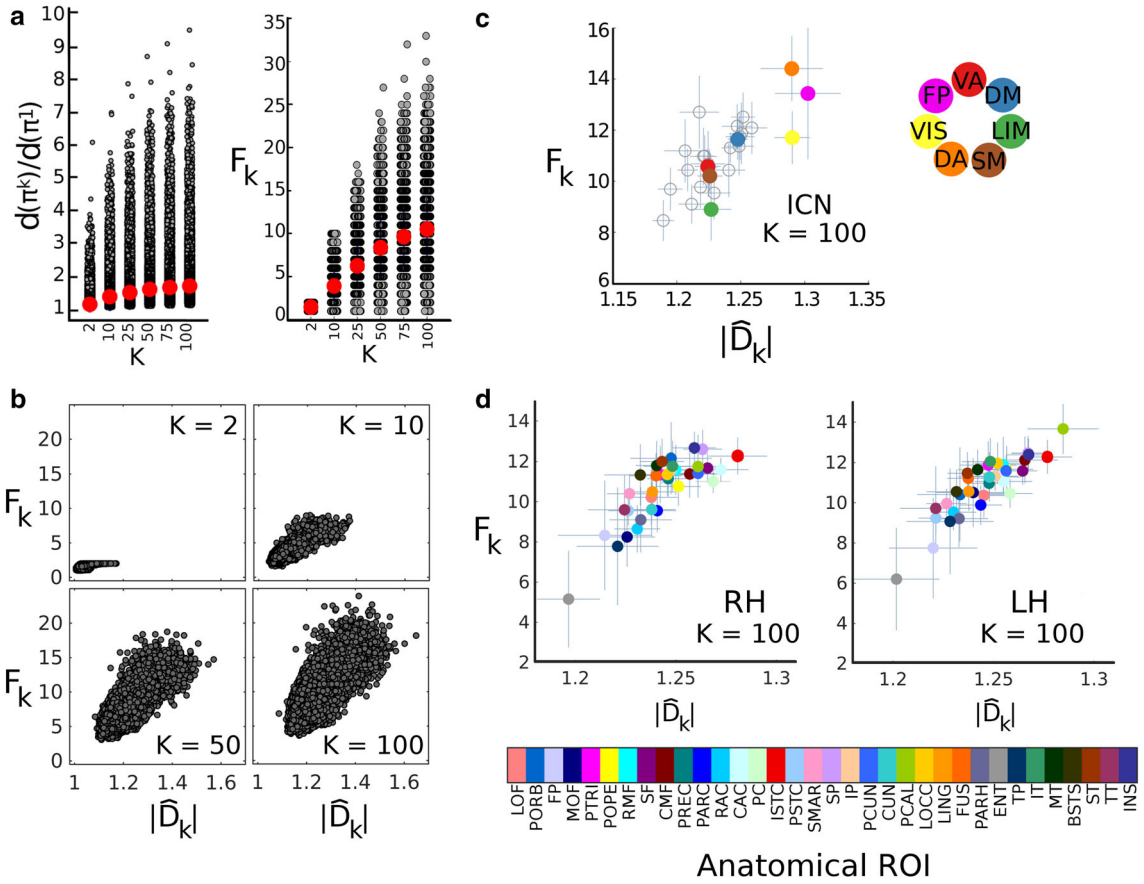

Fig. 6 Scatter plots of k-shortest path length and number of edge-disjoint paths at different spatial scales. **a** Distributions of the ratio of the *k*th shortest-path lengths and the ($k = 1$) shortest-path length (*left panel*), and distribution of the number of edge-disjoint paths (*right panel*). *Red points* indicate the average median across all subjects. **b** *Scatter plot* of $|\hat{D}_k|$ and F_k for all node pairs, for $k = 2$, 10, 50, and 100; points represent averages across all subjects. **c** *Scatter plot* of $|\hat{D}_k|$ and F_k corresponding to path ensembles within (*color-filled points*) and between (*white points*) functional communities, for $k = 100$. *Vertical* and *horizontal lines* in each point represent standard deviations over all subjects. **d** *Scatter plot* of $|\hat{D}_k|$ and F_k corresponding to 34 anatomical regions within the right hemisphere (*left panel*) and 34 anatomical regions within the left hemisphere (*right panel*), both for $k = 100$. *Vertical* and *horizontal lines* in each point represent standard deviations over all subjects

communities (analysis from data-driven ICN parcellation is shown in Fig. S5 in the Online Resource).

It is well known that in spatially embedded networks, the spatial distribution of nodes has significant effects on the topology of the network (Barthélemy 2011) and as a consequence, it is often the case that a significant portion of the variability of topological measures across node pairs may be accounted for by the physical and/or topological distance between nodes. In order to determine whether the relationship between F_k and $|\hat{D}_k|$ is driven by the spatial distance between nodes pairs, we performed a regression analysis for each measure with a log-normal model (average $R^2 = 0.189$ for $|\hat{D}_k|$, and $R^2 = 0.113$ for F_k; $p < 0.01$). After regressing out the effects of Euclidean distance, the measures remain strongly and significantly correlated for values of $k > 2$ ($r = 0.01$, 0.395, 0.455, and 0.437, for $k = 2$, 10, 50 and 100, respectively; $p < 0.01$). These findings suggest that the relationship between these topological measures, evaluated on the SC networks, cannot be fully accounted for by the spatial separation of the nodes.

Regarding the median values of the measures at the ICN level, we examine whether the spatial locations of the nodes within each functional community can explain the segregation of the three ICNs that exhibit higher mean values of F_k and/or $|\hat{D}_k|$ (see Fig. 6c). To do so, we perform permutation tests (1000 repetitions) in which ICN assignments are randomly permuted within each hemisphere. In addition, we constraint the permuted node assignments to ensure that the average intra-hemispheric Euclidean distance between nodes assigned to each ICN is preserved to be equal (within a 5 % error) to the average intra-hemispheric Euclidean distance measured over the empirical ICNs. On average, while high correlations between both topological measures persist across permutations, we find

that the mean values of F_k and $|\hat{D}_k|$ corresponding to path ensembles between and within permuted ICNs are restricted to significantly smaller intervals ([1.204,1.268] and [9.61,12.482] for F_k and $|\hat{D}_k|$, respectively), compared to the values exhibited by empirical ICNs (Fig. S6 and Fig. S7 in Online Resource show scatter plots with average F_k and $|\hat{D}_k|$ values corresponding to permuted ICNs). Finally, relative to the permutation tests, all empirically observed values of F_k and $|\hat{D}_k|$ corresponding to within ICN path ensembles differed significantly ($p < 0.01$), confirming that the path ensembles among nodes belonging to the different functional communities exhibit structural organization patterns that are not solely driven by patterns of spatial location.

We also study the relationship between F_k and $|\hat{D}_k|$ at the level of 68 cortical anatomical regions. For every anatomical region ar_i we compute F_k (ar_i) and $|\hat{D}_k(ar_i)|$, first computing the measures corresponding to every node (by averaging values across all edges adjacent to each node); then averaging across all nodes belonging to each anatomical region ar_i. The scatter plot in Fig. 6d shows F_k and $|\hat{D}_k|$ of 34 right and 34 left hemispheric cortical anatomical regions, for $k = 100$; values are averages across all subjects. Once again, we find that the measures are highly correlated ($r = 0.97$ and 0.98 for right and left hemisphere, respectively; $p < 0.01$). These results show that some anatomical regions that are classified as being highly central are also regions with a large number of edge-disjoint paths, and a k-shortest path length that deviates the most from the optimal path length. Among these regions we find the pericalcarine, the isthmus of the cingulate, and the insula (all bilateral), as well as the right caudal anterior cingulate gyrus, right superior frontal gyrus, right posterior cingulate cortex and left caudal middle frontal gyrus. Likewise, some of the regions that were identified as the least central appear to be regions with a small number of edge-disjoint paths and a smaller deviation from the optimal path length. Among these low-centrality, low resilience regions we find right and left hemispheric entorhinal cortex, and right and left hemispheric frontal pole.

Lateralization of efficiency and resiliency

The spread of right and left hemispheric data points in Fig. 6d suggests that there are structural differences in the connectivity patterns of homotopic cortical regions. We examine hemispheric lateralization by comparing nodal F_k and $\hat{D}_k$ of homotopic anatomical regions in the left and right hemispheres, assessing what regions have an underlying structural connectivity that enables them to communicate more efficiently and/or more resiliently with all other ipsilateral or contralateral anatomical regions. Furthermore, for this analysis we use the normalized version of the number of edge-disjoint paths, $|F_k|$, to remove possible effects of nodal degree and connection density differences across hemispheres.

We quantify lateralization using the measures of *integration* and *segregation* suggested in (Gotts et al. 2013), but we use structural-based measures ($|F_k|$ and $\hat{D}_k$) instead of functional-based measures. For each anatomical region, we compute its average $|F_k|$ and D_k with respect to all other anatomical regions within the ipsilateral and the contralateral hemisphere. The level of integration (with respect to $|F_k|$ or $\hat{D}_k$) associated with an anatomical region is computed as the sum of ipsilateral and contralateral structural-based measures; the level of segregation of an anatomical region is computed as the difference between ipsilateral and contralateral structural-based measures.

We perform separate Wilcoxon signed-rank test for the integration and segregation measures, evaluated over $|F_k|$ and $\hat{D}_k$ across all participants ($p < 0.005$, corrected for false discovery rate (FDR) to $q < 0.01$ for all measures). Figure 7 shows left and right lateralized brain regions across all values of k. These results demonstrate that there are broad ranges of k that exhibit significant left-lateralization for the segregation measures, with a stronger effect for the segregation measure evaluated by $|F_k|$. We also find significant left-lateralization for the integration measure, with more regions being significant for the segregation measure evaluated by $\hat{D}_k$.

Among the left hemispheric anatomical regions that exhibit significantly higher segregation measures with respect to $\hat{D}_k$ and $|F_k|$ are motor-related areas such as precentral and paracentral cortex, language areas such as pars triangularis, pars opercularis, and temporal cortex regions. Among the left hemisphere anatomical regions that exhibit significantly higher integration measures with respect to $\hat{D}_k$ and $|F_k|$ across different values of k are frontal cortex regions, cuneus, temporal cortex regions, and insula for the $\hat{D}_k$ integration measure only.

Discussion

In this study, we introduce the path ensembles comprising a set of k-shortest paths between network nodes and we apply this concept in an analysis of structural paths and communication patterns in human brain networks. In doing so, we build on three interwoven conceptual ideas that underpin our analysis and a set of findings regarding the topology of the human connectome.

The first conceptual idea involves relaxing the assumption that communication takes place only through minimally short paths. While shortest paths represent an optimally short and direct solution for communication between distinct brain regions, relying on them exclusively

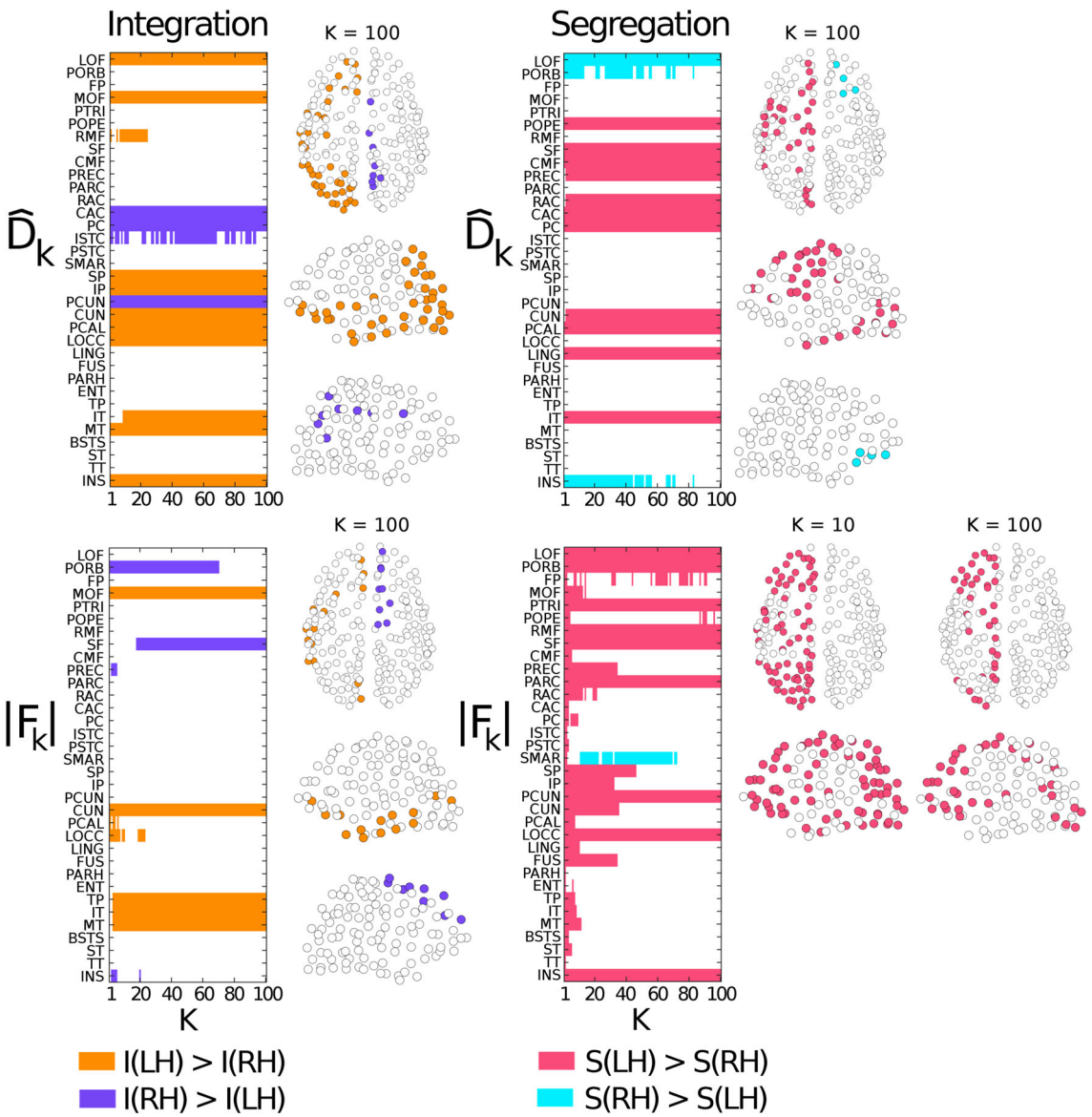

Fig. 7 Right and left hemispheric lateralization of integration and segregation with respect to $|F_k|$ and $\hat{D}_k$. $|F_k|$ and $\hat{D}_k$ segregation and integration measures were compared across all subjects at homotopic anatomical regions in the left and right hemispheres (LH and RH, respectively). Anatomical regions in the left hemisphere found to perform better at the integration measures (I(LH) > I(RH)) are colored with *yellow*, whereas anatomical regions in the right hemisphere found to perform better at the integration measures (I(RH) > I(LH)) are colored with *purple* (*top* and *bottom left panels*). Anatomical regions in the left hemisphere found to perform better at the segregation measures (S(LH) > S(RH)) are colored with *pink*; anatomical regions in the right hemisphere found to perform better at the integration measures (S(RH) > S(LH)) are colored with *cyan* (*top* and *bottom right panels*)

does not allow for the possibility that communication may also take place through alternative or multiple paths; moreover, the shortest paths do not provide any information about the resilience of communication pathways between brain regions. The existence of a short path between a pair of brain regions does not imply that information will uniquely and deterministically flow along that path. The shortest path may be "hidden" and difficult to travel if there are many possible detours from the path; the shortest path also may be a bottleneck. Conversely, there may be many paths that are identical in length (in the case of un-weighted networks), or nearly identical in length to the shortest path (especially in the case of weighted networks). Furthermore, k-shortest path-based measures may be less susceptible to potential noise and artifacts that limit the accuracy of estimating the edges in the data. While the technologies for acquisition and processing of diffusion imaging data are continually improving, the relationship between connection strengths derived from diffusion imaging data (such as fractional anisotropy, streamline

count, or fiber density) and the physiological or anatomical traits determining information flow across white matter pathways remains unclear (Jones et al. 2013). Hence, there is no principled rationale to prioritize optimally short paths in weighted brain networks.

Second, building on elemental concepts of network science, we extend two basic graph-theoretic measures, specifically, the shortest path betweenness centrality of nodes and edges, and the path length between node pairs. As previously noted in (Borgatti 2005), the dynamics of information flow within a network affects the appropriateness of centrality measures. Here we claim that the same is true for measures of communication efficiency and resilience. Basic assumptions about the manner in which information flow takes place through the network define which communication paths are considered and how they contribute to the overall topological measures. The most widely used centrality and distance measures only consider optimal information transmission, with the implicit assumption that the system has full knowledge of the network's topology. On the other hand, measures such as *mean first passage times* (Grinstead and Snell 2012), *diffusion efficiency* (Goñi et al. 2013), *communicability* (Estrada and Hatano 2008), and random-walk betweenness (Newman 2005) take into account all paths (or walks) between node pairs, making no assumptions about optimality or a need for "knowing" about the system's topology. In this paper, we consider a scenario that lies between these two extreme communication schemes.

Third, our approach allows us to explore additional aspects of communication efficiency and resilience of human brain SC networks. There is mounting evidence showing that minimizing wiring cost and the length of communication pathways simultaneously is an important driver for the organization of spatially embedded networks (Barthélemy 2011) such as brain networks (Kaiser and Hilgetag 2006; Bassett et al. 2010; Vertes et al. 2012; Betzel et al. 2016). Nonetheless, empirical and computational studies have also shown that there are additional aspects of communication efficiency that are fundamental for brain function, for example: limiting the occurrence of bottlenecks that cause delays of information loss (Tombu et al. 2011; Mišić et al. 2014); minimizing the number of intermediate steps between a signaling source and its target (Kaiser and Hilgetag 2006); recruiting different pathways to optimize performance, robustness and resilience (Kaiser et al. 2007; Masel and Trotter 2010; Stern et al. 2005; Wook Yoo et al. 2015); and promoting the capacity to integrate and segregate information processing at various spatial–temporal scales (Tononi et al. 1994; Gallos et al. 2012; Sporns 2013). In this study, we address some of these aspects, and furthermore, we uncover a competitive interaction, or trade-off, between communication efficiency and resilience, which, to our knowledge, has not been explored in the context of SC brain networks.

Building on these ideas, our first main finding is that k-shortest path ensembles broaden the participation of all connections in communication processes, including faint but direct connections often spanning long physical distances. As a result of broadening the participation of connections, path ensembles linking many pairs of nodes are diffusely laid out, in some cases crossing multiple different ICN and anatomical boundaries. Interestingly, the participation of connections in communication pathways as a function of k is also dependent on the function used to transform connection strengths into connection lengths. For the results presented in this work, we use a *Log* transform that yields a log-normal distribution of the resulting connection lengths. There are two main arguments supporting this choice: First, this distribution is consistent with evidence pointing towards log-normal distributions of synaptic strengths between cortical cells (Buzsáki and Mizuseki 2014) as well as strengths of cortico-cortical projections (Markov et al. 2012). Second, the *Log* transform (unlike the often used inverse transform) results in edge lengths that allow all edges to contribute in efficient communication paths (see Fig. 2). We note that other transforms resulting in heavy-tailed distributions of edge lengths may also be compatible with empirical data and recruit all edges for communication processes and should be explored in future work.

Our next result shows that the centrality ranking of nodes and edges varies as a function of k, the number of k-shortest paths contained in the path ensembles. We also find that, as the value of k increases, there is an increase in the average degree and wiring cost of the top 10 % central nodes and edges. While the relationship between high shortest-path centrality, high degree (hubs) and high network cost has been elucidated before as a characteristic of the rich club (van den Heuvel et al. 2011; Collin et al. 2014), our findings suggest that these rich club properties might be accentuated when we consider k-shortest path ensembles. An additional interesting finding regarding the properties of highly central nodes and edges is that they are significantly more resilient. However, these features are only found in edges that are highly central for values of $k > 1$ (Fig. 5b). Because we measure resilience in terms of the number of edge-disjoint paths, this result suggests that the removal of a high $(k > 1)$-centrality edge may not compromise the communication between the pair of brain regions connected by such edge. Moreover, the larger values of F_k exhibited by high-centrality edges suggest that they are less prone to becoming bottlenecks for communication.

A main component of our work is the finding of a competitive relationship between communication efficiency, and

resilience. Such a relationship is present at multiple spatial scales, from the level of ROI pairs, to anatomically based regions, and functional communities. It suggests that the k-shortest paths comprising path ensembles between brain regions cannot be both optimally short (as short as the shortest path) and maximally resilient. As a result, we observe a strong positive linear relationship between F_k and $|\hat{D}_k|$, a characteristic feature of traits that engage in mutual competition, or a trade-off (Shoval et al. 2012); hence, some path ensembles are organized to favor short communication pathways, while other path ensembles "branch out" and thus favor more resilient communication. An important question to consider is whether the relationship between F_k and $|\hat{D}_k|$ is the product of geometrical or mathematical constraints, or conversely, it is the result (or byproduct) of selective pressures that work towards optimizing the functionality of the system. In other words, are path ensembles with high values of F_k and low values of $|\hat{D}_k|$ geometrically impossible to create, or are they non-functional, and thus, they are not observed in neuronal systems (McGhee 2006)? The work presented here does not allow us to answer this question. To address the question of whether F_k and $|\hat{D}_k|$ are geometrically or functionally constrained, a multi-objective evolutionary approach could be adopted, such as the one carried out in (Goñi et al. 2013; Avena-Koenigsberger et al. 2014). This approach would allow us to determine whether or not it is possible to generate brain-like topologies that break the linear relationship between F_k and $|\hat{D}_k|$.

Our final result demonstrates a hemispheric lateralization of integration and segregation, which we evaluate by comparing $\hat{D}_k$ and $|F_k|$ of homotopic anatomical regions. While hemispheric lateralization has been previously demonstrated in neuroimaging studies, it has been mostly studied from a functional perspective, quantified by examining differences in the overall magnitude of brain activity (van Essen et al. 2012; Seghier et al. 2011), or differences in the magnitude of BOLD time-series correlations (Powell et al. 2012; Liu et al. 2009). More recently, Gotts et al. (2013) presented two functionally based measures that aim to quantify distinct forms of hemispheric lateralization, specifically, integration and segregation. Despite that the findings in (Gotts et al. 2013) have advanced the understanding of how the two hemispheres interact, the underlying structural substrates of the observed dynamical differences are still unknown. The work by Iturria-Media et al. (2011) takes a first step using shortest path-based measures to detect differences in the connectivity patterns of the right and left hemispheres. Moving forward, our work uses path ensembles to demonstrate differences in the patterns of structural connectivity that yield hemispheric lateralization of integration and segregation. Importantly, our results are in line with the findings presented in (Iturria-Medina et al. 2011), and several of the regions we find to exhibit high integration/segregation measures match with regions found in (Gotts et al. 2013), particularly for the case of left-lateralization of the functional-based and $|F_k|$-based segregation measures.

There are a number of methodological aspects of the present work that require further considerations. First, computational limitations precluded us from computing k-shortest path ensembles for $k > 100$, for all 40 subjects. A limited exploration in which we computed 250 paths for a single subject shows that the median of the distributions of $\hat{D}_k$ and F_k continue to increase gradually, but we are unable to identify an asymptotic value. Moreover, we are unable to perform the hemispheric lateralization analysis within single subjects, so we cannot determine whether our results would vary greatly when $k > 100$.

Second, while all the networks used for this work were derived through the same processing pipeline, individual differences did not allow us to identify sets of nodes and edges with consistent properties across all subjects. For example, the top 10 % central edges differed for all subjects. This also raises the questions of whether our results are qualitatively consistent across different data sets, processed with different pipelines and different parcellations. Moreover, as higher quality data becomes more available, it may become more common to use fine-grained parcellations that include subcortical regions. This kind of model would certainly require an analysis with larger values of k, which call for much larger computational resources. It is worth noting that, despite the increasing quality of diffusion imaging data and data-processing algorithms, these techniques are still prone to false positives and negatives (Thomas et al. 2014) that may affect our results. One of the most important shortcomings of diffusion imaging is the lack of information about the directionality of anatomical connections in human networks. The findings presented in this study could differ significantly if we were able to perform our analysis on directed anatomical networks. Tangentially related is the definition of ICNs based on Pearson correlations, which in the future should be supplemented by definitions that draw on more sophisticated measures of effective connectivity. We note that all the concepts and measures presented in this study apply for both undirected and directed networks.

Finally, a number of aspects of this work will benefit from future extensions. These include studying the effects resulting from perturbations over anatomical connections, more specifically, the directed removal of highly central edges, where centrality is assessed according to k-shortest paths, and studying the effects of setting limits on the length of viable alternative paths by defining the value of k as a local property of a node pair, instead of defining k as a global variable. More generally, a number of network-based measures that have been used to characterize the

organization of brain networks can be extended to consider path ensembles. These generalizations are not limited to study and characterization of structural connectivity, but may provide a fuller understanding of the relationship between brain structure and function.

Compliance with ethical standards

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments of comparable ethical standards.

Funding statement B.M. was supported by a Natural Science and Engineering Research Council of Canada postdoctoral fellowship. A.G. and P.H. were supported by the Swiss National Science Foundation (#310030-156874, NCCR-Synapsy 51AU40_125759), the Center for Biomedical Imaging (CIBM) of the Geneva-Lausanne Universities and EPFL, Leenaards Foundation, and Louis-Jeantet Foundation. J.G. was supported by the National Institute of Health (1R01 MH108467-01). O.S. was supported by the J.S. McDonnell Foundation (220020387), the National Science Foundation (1212778), and the National Institutes of Health (R01 AT009036-01).

References

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

Achard S, Bullmore E (2007) Efficiency and cost of economical brain functional networks. PLoS Comput Biol 3(2):e17–e17

Achard S, Salvador R, Whitcher B, Suckling J, Bullmore ED (2006) A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J Neurosci 26(1):63–72

Avena-Koenigsberger A, Goñi J, Betzel RF, van den Heuvel MP, Griffa A, Hagmann P, Sporns O (2014) Using Pareto optimality to explore the topology and dynamics of the human connectome. Philos Trans Royal Soc B Biol Sci 369(1653):20130530

Barthélemy M (2011) Spatial networks. Phys Rep 499(1):1–101

Bassett DS, Bullmore ED (2006) Small-world brain networks. Neuroscientist 12(6):512–523

Bassett DS, Greenfield DL, Meyer-Lindenberg A, Weinberger DR, Moore SW, Bullmore ET (2010) Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. PLoS Comput Biol 6(4):e1000748

Betzel RF, Byrge L, He Y, Goñi J, Zuo XN, Sporns O (2014) Changes in structural and functional connectivity among resting-state networks across the human lifespan. NeuroImage 102:345–357

Betzel RF, Avena-Koenigsberger A, Goñi J, He Y, de Reus MA, Griffa A, van den Heuvel M (2016) Generative models of the human connectome. Neuroimage 124:1054–1064

Biggs N, Lloyd EK, Wilson RJ (1976) Graph theory 1736–1936. Oxford University Press, Oxford

Boguña M, Krioukov D, Claffy KC (2009) Navigability of complex networks. Nat Phys 5(1):74–80

Borgatti SP (2005) Centrality and network flow. Soc Netw 27(1):55–71

Bullmore E, Sporns O (2012) The economy of brain network organization. Nat Rev Neurosci 13(5):336–349

Buzsáki G, Mizuseki K (2014) The log-dynamic brain: how skewed distributions affect network operations. Nat Rev Neurosci 15(4):264–278

Cammoun L, Gigandet X, Meskaldji D, Thiran JP, Sporns O, Do KQ, Maeder P, Meuli R, Hagmann P (2012) Mapping the human connectome at multiple scales with diffusion spectrum MRI. J Neurosci Methods 203:386–397

Chung FR (1997) Spectral graph theory, vol 92. American Mathematical Society

Collin G, Sporns O, Mandl RC, van den Heuvel MP (2014) Structural and functional aspects relating to cost and benefit of rich club organization in the human cerebral cortex. Cereb Cortex 24(9):2258–2267

Cormen TH, Leiserson CE, Rivest RL, Stein C (2001) Introduction to algorithms, vol 6. MIT press, Cambridge

da Fontoura Costa L, Travieso G (2007) Exploring complex networks through random walks. Phys Rev E 75(1):016102

Daducci A, Gerhard S, Griffa A, Lemkaddem A, Cammoun L, Gigandet X, Meuli R, Hagmann P, Thiran JP (2012) The connectome mapper: an open-source processing pipeline to map connectomes with MRI. PLoS One 7:e48121

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Albert MS (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31(3):968–980

Estrada E, Hatano N (2008) Communicability in complex networks. Phys Rev E 77(3):036111

Ford LR Jr, Fulkerson DR (1987) Maximal flow through a network. Classic papers in combinatorics. Birkhäuser, Boston, pp 243–248

Friston KJ (2011) Functional and effective connectivity: a review. Brain Connect 1(1):13–36

Galán RF (2008) On how network architecture determines the dominant patterns of spontaneous neural activity. PLoS One 3(5):e2148

Gallos LK, Makse HA, Sigman M (2012) A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. Proc Natl Acad Sci 109(8):2825–2830

Gong G, Rosa-Neto P, Carbonell F, Chen ZJ, He Y, Evans AC (2009) Age-and gender-related differences in the cortical anatomical network. J Neurosci 29(50):15684–15693

Goñi J, Avena-Koenigsberger A, de Mendizabal NV, van den Heuvel M, Betzel R, Sporns O (2013) Exploring the morphospace of communication efficiency in complex networks. PLoS One 8(3):e58070

Goñi J, van den Heuvel MP, Avena-Koenigsberger A, de Mendizabal NV, Betzel RF, Griffa A, Sporns O (2014) Resting-brain functional connectivity predicted by analytic measures of network communication. Proc Nat Acad Sci 111(2):833–838

Gotts SJ, Jo HJ, Wallace GL, Saad ZS, Cox RW, Martin A (2013) Two distinct forms of functional lateralization in the human brain. Proc Natl Acad Sci 110(36):E3435–E3444

Grinstead CM, Snell JL (2012) Introduction to probability. American Mathematical Society

Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O (2008) Mapping the structural core of human cerebral cortex. PLoS Biol 6:e159

Hermundstad AM, Bassett DS, Brown KS, Aminoff EM, Clewett D, Freeman S, Frithsen A, Johnson A, Tipper CM, Miller MB, Grafton ST, Carlson JM (2013) Structural foundations of resting-state and task-based functional connectivity in the human brain. Proc Natl Acad Sci USA 110(15):6169–6174

Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, Hagmann P (2009) Predicting human resting-state functional connectivity from structural connectivity. Proc Natl Acad Sci 106(6):2035–2040

Iturria-Medina Y, Fernández AP, Morris DM, Canales-Rodríguez EJ, Haroon HA, Pentón LG, Melie-García L (2011) Brain hemispheric structural efficiency and interconnectivity rightward asymmetry in human and nonhuman primates. Cereb Cortex 21(1):56–67

Jones DK, Knösche TR, Turner R (2013) White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. Neuroimage 73:239–254

Joyce KE, Laurienti PJ, Burdette JH, Hayasaka S (2010) A new measure of centrality for brain networks. PLoS One 5(8):e12200

Kaiser M, Hilgetag CC (2006) Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. PLoS Comput Biol 2(7):e95–e95

Kaiser M, Martin R, Andras P, Young MP (2007) Simulation of robustness against lesions of cortical networks. Eur J Neurosci 25(10):3185–3192

Latora V, Marchiori M (2001) Efficient behavior of small-world networks. Phys Rev Lett 87(19):198701

Liu H, Stufflebeam SM, Sepulcre J, Hedden T, Buckner RL (2009) Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors. Proc Natl Acad Sci USA 106(48):20499–20503

Lynall ME, Bassett DS, Kerwin R, McKenna PJ, Kitzbichler M, Muller U, Bullmore E (2010) Functional connectivity and brain networks in schizophrenia. J Neurosci 30(28):9477–9487

Markov NT, Ercsey-Ravasz MM, Gomes AR, Lamy C, Magrou L, Vezoli J, Sallet J (2012) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cereb Cortex. bhs270

Masel J, Trotter MV (2010) Robustness and evolvability. Trends Genet 26(9):406–414

McGhee GR (2006) The geometry of evolution: adaptive landscapes and theoretical morphospaces. Cambridge University Press, England

Mišić B, Sporns O, McIntosh AR (2014) Communication efficiency and congestion of signal traffic in large-scale brain networks. PLoS Comput Biol 10(1):e1003427

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

Newman ME (2005) A measure of betweenness centrality based on random walks. Soc Netw 27(1):39–54

Park HJ, Friston K (2013) Structural and functional brain networks: from connections to cognition. Science 342(6158):1238411

Passingham RE, Stephan KE, Kötter R (2002) The anatomical basis of functional localization in the cortex. Nat Rev Neurosci 3(8):606–616

Powell JL, Kemp GJ, García-Finaña M (2012) Association between language and spatial laterality and cognitive ability: an fMRI study. Neuroimage 59(2):1818–1829

Rosvall M, Grönlund A, Minnhagen P, Sneppen K (2005) Searchability of networks. Phys Rev E 72(4):046117

Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52(3):1059–1069

Seghier ML, Josse G, Leff AP, Price CJ (2011) Lateralization is predicted by reduced coupling from the left to right prefrontal cortex during semantic decisions on written words. Cereb Cortex 21(7):1519–1531

Shoval O, Sheftel H, Shinar G, Hart Y, Ramote O, Mayo A, Alon U (2012) Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. Science 336(6085):1157–1160

Simas T, Rocha LM (2015) Distance closures on complex networks. Netw Sci 3(02):227–268

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF (2009) Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci 106(31):13040–13045

Sporns O (2013) Network attributes for segregation and integration in the human brain. Curr Opin Neurobiol 23(2):162–171

Sporns O, Tononi G, Kötter R (2005) The human connectome: a structural description of the human brain. PLoS Comput Biol 1(4):e42

Sporns O, Honey CJ, Kötter R (2007) Identification and classification of hubs in brain networks. PLoS One 2(10):e1049–e1049

Stern Y, Habeck C, Moeller J, Scarmeas N, Anderson KE, Hilton HJ, van Heertum R (2005) Brain networks associated with cognitive reserve in healthy young and old adults. Cereb Cortex 15(4):394–402

Thomas C, Frank QY, Irfanoglu MO, Modi P, Saleem KS, Leopold DA, Pierpaoli C (2014) Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proc Natl Acad Sci 111(46):16574–16579

Tombu MN, Asplund CL, Dux PE, Godwin D, Martin JW, Marois R (2011) A unified attentional bottleneck in the human brain. Proc Natl Acad Sci 108(33):13426–13431

Tononi G, Sporns O, Edelman GM (1994) A measure for brain complexity: relating functional segregation and integration in the nervous system. Proc Natl Acad Sci 91(11):5033–5037

van den Heuvel MP, Sporns O (2011) Rich-club organization of the human connectome. J Neurosci 31(44):15775–15786

van den Heuvel MP, Stam CJ, Kahn RS, Pol HEH (2009) Efficiency of functional brain networks and intellectual performance. J Neurosci 29(23):7619–7624

van den Heuvel MP, Kahn RS, Goñi J, Sporns O (2012) High-cost, high-capacity backbone for global brain communication. Proc Natl Acad Sci 109(28):11372–11377

Van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T (2012) Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. Cereb Cortex 22(10):2241–2262

Vértes PE, Alexander-Bloch AF, Gogtay N, Giedd JN, Rapoport JL, Bullmore ET (2012) Simple models of human brain functional networks. Proc Natl Acad Sci 109(15):5868–5873

Wedeen VJ, Wang RP, Schmahmann JD, Benner T, Tseng WYI, Dai G, Pandya DN, Hagmann P, D'Arceuil H, de Crespigny AJ (2008) Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. Neuroimage 41:1267–1277

Yan C, Gong G, Wang J, Wang D, Liu D, Zhu C, He Y (2011) Sex- and brain size–related small-world structural cortical networks in young adults: a DTI tractography study. Cereb Cortex 21(2):449–458

Yen JY (1971) Finding the k shortest loopless paths in a network. Manage Sci 17(11):712–716

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106(3):1125–1165

Yoo SW, Han CE, Shin JS, Seo SW, Na DL, Kaiser M, Seong JK (2015) A network flow-based analysis of cognitive reserve in normal ageing and Alzheimer's Disease. Scientific reports 5:10057

Zuo XN, Ehmke R, Mennes M, Imperati D, Castellanos FX, Sporns O, Milham MP (2012) Network centrality in the human functional connectome. Cereb Cortex 22(8):1862–1875