Supplementary Materials: The specificity and robustness of long-distance connections in weighted, interareal connectomes

Richard F. Betzel[1*] and Danielle S. Bassett[1,2,3,4†]
[1]*Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, 19104 USA*
[2]*Department of Physics & Astronomy, University of Pennsylvania, Philadelphia, PA, 19104 USA*
[3]*Department of Neurology, University of Pennsylvania, Philadelphia, PA, 19104 USA and*
[4]*Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA, 19104 USA*
(Dated: April 10, 2018)

OVERVIEW

This supplement includes methodological details about the provenance and processing pipelines for each of the five interareal connectome datasets. It also includes derivations of the network metrics reported in the main text and an overview of the process used to define structural and functional network modules, an explanation of alternative communication schemes (i.e. communicability and navigation routing), and demonstrations that our results are robust with respect to alternative weight-to-length mappings and thresholding of weak connections.

MATERIALS AND METHODS

We analyzed mouse, *Drosophila*, macaque, and human weighted, interareal network datasets. Each dataset was distinct in terms of imaging modality, reconstruction technique, and connection weighting scheme. This variability in processing strategy was unintentional, though we exploit this feature in order to demonstrate the universality of our findings and their robustness to acquisition and processing schemes. In this section, we describe the methods used to reconstruct and analyze the networks.

Network datasets

Mouse. The mouse connectivity matrix was reconstructed based on freely available tract-tracing data from the Allen Institute Mouse Brain Connectivity Atlas (`http://connectivity.brain-map.org`; see [1] for more details of tract-tracing experiments). Anterograde recombinant adeno-associated viral tracer was injected into target areas in the right hemisphere of mouse brains, which was extracted three weeks post-injection at which time viral tracer projection patterns were reconstructed. Reconstructions were then smoothed and aligned to a common coordinate space of the Allen Reference Atlas.

Network nodes were defined according to a custom parcellation based on the Allen Developing Mouse Brain Atlas [2]. This parcellation contains 65 areas in each hemisphere, 9 of which were removed because they were not involved in any tract-tracing experiment. The resulting weighted and directed network contained $N = 112$ areas of interest linked by edges corresponding to interareal axonal projections and weighted as normalized connection densities: the number of connections from unit volume of a source area to unit volume of a target area (Fig. 1a). This network has been analyzed elsewhere [2–4].

Drosophila. The connectivity matrix for *Drosophila* was reconstructed from the FlyCircuit 1.1 database (`http://www.flycircuit.tw`), a repository of images of 12,995 projections in the female *Drosophila* brain [5]. Neurons were labeled with green fluorescent protein (GFP) using genetic mosaic analysis with a repressible cell marker. GFP-labeled neurons were delineated from whole-brain three-dimensional images and co-registered to a female template brain using a rigid linear transform. Individual neurons were partitioned into $N = 49$ local processing units (LPUs; network nodes) with distinct morphological and functional characteristics. LPUs were defined so as to contain their own population of local interneurons whose fibers were limited to that LPU. The result was a weighted and directed network comprised of projections among LPUs. This network has been analyzed elsewhere [3, 4, 6, 7] (Fig. 1b).

Macaque. The macaque connectivity matrix was based on retrograde tract-tracing experiments and originally reported in [8]. Injections of fluorescent tracers were made in 28 macaque monkeys. Reconstructed projections were localized with respect to a parcellation comprised of $N = 91$ cortical areas based on histological and atlas-based landmarks. For each tract-tracing experiment, the number of labeled neurons in each of the 91 areas was counted. This number was then expressed relative to the number of labeled neurons minus the number of neurons intrinsic to the injection site. The result is a $[29 \times 91]$ matrix of connection weights from each injection site to the rest of the brain. We focused on the $[29 \times 29]$ weighted and directed connectivity matrix [9] (Fig. 1c).

Human structural networks. Human brain networks were reconstructed from diffusion weighted MRI

* rbetzel @ seas.upenn.edu
† dsb @ seas.upenn.edu

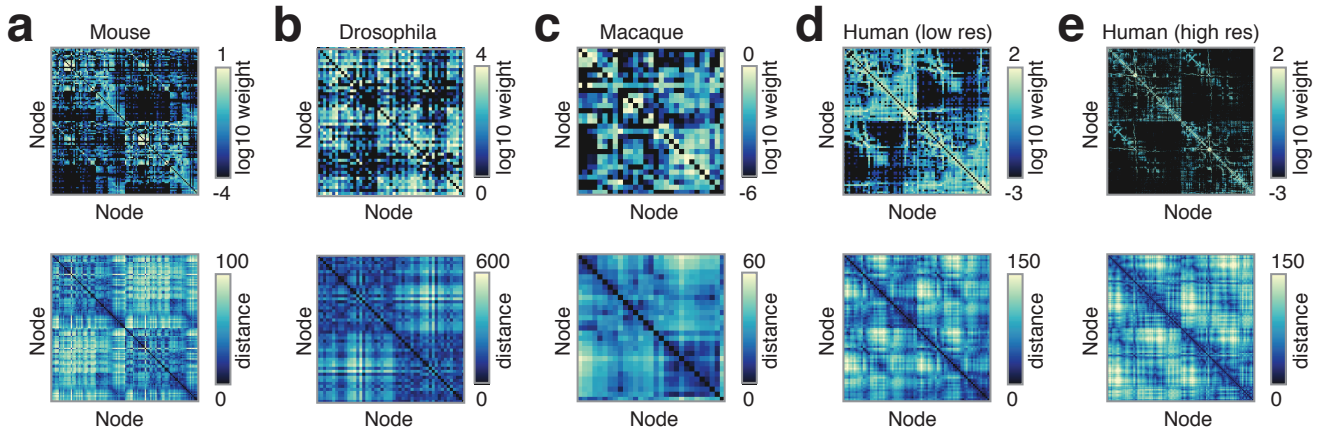

FIG. 1. **Network matrices.** (*top row*): Matrix representations of connectomes reconstructed from macaque, *Drosophila*, mouse, and human imaging data. (*bottom row*): Euclidean distance matrices for the same species.

using deterministic tractography algorithms. The networks we analyzed were group-representative composites of subject-level networks (30 subjects). This network construction process entailed acquiring diffusion spectrum and T1-weighted anatomical images for each individual. DSI scans sampled 257 directions using a Q5 half-shell acquisition with a maximum b-value of 5000, an isotropic voxel size of 2.4 mm, and an axial acquisition with repetition time $TR = 5$ seconds, echo time $TE = 138$ ms, 52 slices, and field of view of $[231, 231, 125]$ mm. All procedures were approved by the Institutional Review Board of the University of Pennsylvania and all participants gave informed consent.

DSI data were reconstructed in DSI Studio (`www.dsi-studio.labsolver.org`)), using q-space diffeomorphic reconstruction (QSDR) [10]. QSDR reconstructs diffusion-weighted images in native space, computes the quantitative anisotropy (QA) of each voxel, warps the image to a template QA volume in Montreal Neurological Institute (MNI) space using the statistical parametric mapping nonlinear registration algorithm, and reconstructs spin-density functions with mean diffusion distance of 1.25 mm with three fiber orientations per voxel. Fiber tracking was performed using a modified FACT algorithm with an angular cutoff of 55°, step size of 1.0 mm, minimum length of 10 mm, spin density function smoothing of 0.00, maximum length of 400 mm, and a QA threshold determined by DWI signal in the colony-stimulating factor. For each individual, the algorithm terminated when 1,000,000 streamlines were reconstructed.

In parallel, T1 anatomical scans were segmented using FreeSurfer and parcellated using the Connectome Mapping Toolkit (`http://www.connectomics.org`) according to low- and high-resolution atlases ($N_{low} = 82$ and $N_{high} = 1000$ areas) [11]. The low-resolution atlas comprised 68 cortical areas and 14 subcortical structures. The high-resolution atlas comprised 1000 cortical areas, representing subdivisions of cortical areas delineated in the low-resolution atlas, and 14 subcortical structures. Note that the upsampling procedure applied to the cortical areas was not applied to the subcortical structures. As a result, the volumes and surface areas of subcortical structures in the high-resolution atlas were many times greater than that of the high-resolution cortical areas. Because large morphometric disparities can induce unwanted biases in network analysis, we elected to exclude sub-cortical structures from our analysis of networks constructed using the high-resolution atlas. Each parcellation was registered to the B0 volume of subjects' DSI data, and a B0-to-MNI voxel mapping was used to map area labels from native space to MNI coordinates. Streamlines were aggregated by the areas in which their starting and termination points were located. The connection weight between any pair of areas was defined as their streamline count normalized by the geometric means of their volumes.

Group-representative matrices were generated using a distance-dependent, consistency-based thresholding procedure. This procedure was applied separately to inter- and intra-hemispheric connections. The resulting networks had a binary density equal to the average across subjects, approximately the same distribution of inter- and intra-hemispheric edge lengths, and approximately the same edge weight distribution as every subject (Fig. 1d,e). This approach has been described elsewhere [12, 13] and shown to be superior to distance-agnostic thresholding procedures [14].

Network analysis

Inter-areal networks were represented as weighted connectivity matrices, $\mathbf{W} \in \mathbb{R}^{N \times N}$, where the element w_{ij} denoted the strength of the connection between brain areas (nodes) i and j. We encoded spatial relationships between nodes with Euclidean distance matrices, $\mathbf{E} \in \mathbb{R}^{N \times N}$, where the element e_{ij} denoted the straight-line distance between the physical locations of areas i and j.

Shortest weighted paths. The shortest path between two areas represents the most direct channel by which they can communicate, with shorter paths implying enhanced communication capacity. One of the challenges of studying shortest path structure in weighted complex networks is that in order to compute shortest paths, we must first map connection weights from a measure of nodes' affinities for one another to a measure of length. In principle, this mapping would be data-driven and, in the case of neural systems, informed by neurobiology. Typically, however, the most neurophysiologically accurate mapping for a particular data type, spatial resolution, and organism, is unknown, which motivates the exploration of different heuristic mappings. A common choice is to compute length as the reciprocal of weight, i.e. $l_{ij} = w_{ij}^{-1}$. While effective, this mapping is pragmatic and there is no particular motivation for why weights are raised to a power of -1 as opposed to some other power.

While it is infeasible to explore all possible weight-to-length mappings, it is essential to test several candidates in order to increase confidence in the generalizability of results, and to ensure that our findings are not idiosyncrasies of a particular choice for how to map weights to lengths. Accordingly, we consider a family of parameterized weight-to-length mappings: $l_{ij} = \frac{1}{w_{ij}^{-\alpha}}$. As α increases, the weights of the strongest connections become exponentially greater than the weights of the weakest connections, and a larger proportion of shortest paths are funnelled through those (strong) connections. As $\alpha \to 0$, the ratio grows smaller and shortest paths are more evenly distributed over connections. Conveniently, this mapping also allows us to recover, as special cases, the binary shortest paths ($\alpha = 0$) and the reciprocal mapping noted earlier, $\alpha = 1$. Once weights were transformed into lengths, the network's shortest path structure was computed and stored in the distance matrix, $\mathbf{D} \in \mathbb{R}^{N \times N}$, whose element $d_{st} = l_{si} + l_{ij} + \ldots + l_{kt}$ encoded the length of the weighted shortest path between source area s and target area t [15].

Mean weighted path length. Given a network's shortest path structure, we calculated a number of useful metrics. The simplest was the average length of shortest paths:

$$\langle L \rangle = \frac{2}{N(N-1)} \sum_{i,j>i} d_{ij}. \quad (1)$$

This measure tells us, on average, the cost of using shortest paths for communication.

While we focus on the $w_{ij}^{-\alpha}$ mapping in the main text, we also explore an alternative mapping in this supplement. Specifically, we consider a "negative log" mapping: $L_{ij} = -\log\left(\frac{W_{ij}}{\max_{ij}(W_{ij})}\right)$. Using this mapping, we find results comparable to those reported in the main text, namely that shortest paths are composed, mostly, of strong short-range connections. We show these results in Fig. 2 and Fig. 3.

Edge betweenness centrality. We also calculated the contributions made to a network's shortest path structure by its connections [16]. Let $\pi_{s \to t} = \{\{s,i\},\{i,j\},\ldots,\{k,t\}\}$ be the sequence of connections traversed along the shortest path from a source node s to a target node t. A connections's betweenness centrality, $\mathcal{B}_{ij}$, measures the fraction of all shortest paths that include the connection $\{i,j\}$; its value can be interpreted as a measure of a connection's importance for communication along a network's shortest paths.

Alternative communication schemes. In the main text of this paper we consider inter-areal patterns of communication taking place along the shortest weighted path between pairs of brain regions. While shortest path communication is among the most widely studied mechanisms [17], there are others that we do not explicitly consider. In this section we test the effect of short-range and long-distance connections on two other communication mechanisms: communicability [18, 19] and navigability [20–22].

Whereas shortest-paths routing assumes that communication between two regions occurs along a single path composed of strong connections, communicability assumes that communication occurs along all paths of all lengths while exponentially discounting longer, multi-step paths. Formally, communicability is calculated as the matrix exponential $\mathbf{G} = e^{\mathbf{W}}$, so that $G_{ij} = \sum_{k=0}^{\infty} \frac{(W^k)_{ij}}{k!}$. Accordingly, communicability can be calculated for all pairs $\{i,j\}$, even if those nodes are not directly connected, and can be thought of intuitively as the capacity for those nodes to communicate along all pathways. In addition to pairwise communicability, we also computed an average communicability among all pairs of nodes. Note that to ensure that the matrix exponential does not "blow up", we divide all elements of $\mathbf{W}$ by $\mathbf{W}$'s largest eigenvalue, λ_{max} before computing $\mathbf{G}$.

Navigability – or "navigation routing" – refers to an algorithm in which a message is sent from a source node i to a target node j. Starting with node i and at each subsequent step the message is passed along its current node's outgoing links to the neighbor nearest its target in terms of Euclidean distance (note: other distance metrics

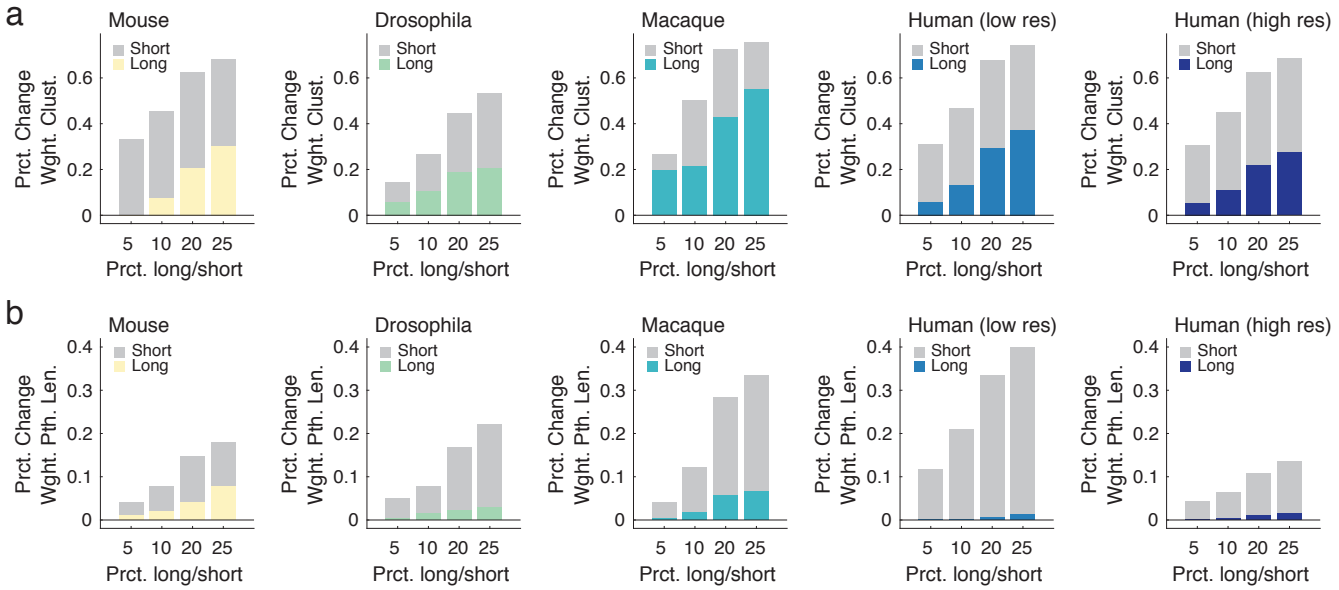

FIG. 2. **Change in weighted clustering coefficient and path length after removing short-range and long-distance connections.** (**a**) Change in weighted clustering coefficient as a result of removing the 5%, 10%, 20%, and 25% shortest and longest connections. (**b**) Change in weighted path length as a result of removing the 5%, 10%, 20%, and 25% shortest and longest connections. Note that shortest paths were computed based on a logarithmic weight-to-length mapping.

can be considered, e.g. those induced after embedding a network into a hyperbolic space [23]). Unlike shortest paths routing, navigation routing can fail to pass the message to its target, e.g. if the message gets stuck in a cycle. We can compute a number of measures based on this routing mechanism, including (i) an edge-wise usage statistic tabulating the number paths in which each edge is involved, (ii) the length of each completed path in number of steps, and (iii) the total number of failures across all pairs of nodes.

Our comparison of shortest paths routing with communicability and navigation routing included two distinct components. First, we aimed to show that measures of communicability and navigability both exhibit distance dependencies. Second, we aimed to show that not only do these communication strategies depend upon distance, but that weak long-distance connections play either a minor or complementary role to strong short-range connections.

To demonstrate that both communicability and navigation routing depend on Euclidean distance, we computed the correlation of Euclidean distance with pairwise communicability and edge usage (Fig. 4). In both cases, we observed robust negative correlations. These supplementary results indicate that, even under different routing schemes, the capacity for communication among distant nodes tends to be weaker than for nodes located nearer to one another.

Next, we systematically removed long-distance and short-range connections from the network and computed what effect their removal had on average pairwise communicability in the case of communication by communicability, and on average path length and failure count in the case of navigability. In line with the main text, we found that removing strong, short-range connections resulted in greater decreases in communicability compared to the removal of weak long-distance connections (Fig. 5). On the other hand, because navigation routing can efficiently take advantage of long-distance pathways to direct a message towards the neighborhood of its target, we found that removing long connections consistently increased path length (Fig. 6a). However, short-range connections also play a critical role. We observed that removing short-range connections resulted in an increased number of failures (Fig. 6b).

Collectively, these findings paint a complicated picture of communication mechanisms. Shortest paths routing and communicability both depend upon connection weight; perhaps for this reason results obtained using these methods are fairly consistent with one another, suggesting that strong short-range connections play an outsized role in shaping communication patterns. Navigability on the other hand, depends on the spatial embedding of the brain to facilitate communication, exploiting long-distance pathways to ensure few processing steps while depending on short-range connections for endpoint routing. These communication strategies differ not only in terms of mechanism, but also in terms of biological realism. Shortest paths routing requires that a message's target destination be known ahead of time and assumes com-

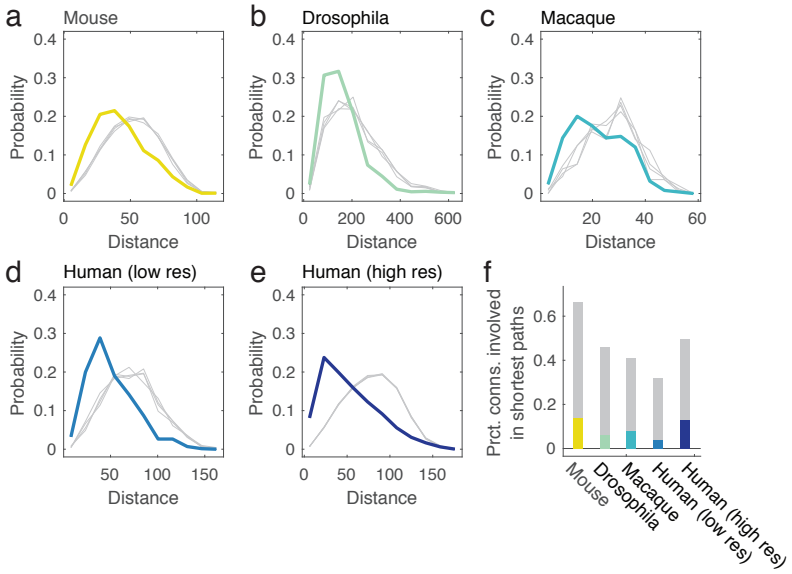

FIG. 3. **Shortest path usage using logarithmic weight-to-length mapping.** (**a,b,c,d,e**) Euclidean distance of edges used in shorest paths for the empirical network (colored) and randomized null model (gray). (**f**) For each network, we calculated the fraction of total connections involved in shortest paths (gray) and of those connections, the fraction that were among the 25% longest (colored).

plete knowledge of the network's topology, which may be an inaccurate assumption for neural systems. Similarly, navigation routing requires that a message's target destination be known ahead of time. However, rather than requiring global knowledge of all shortest paths, at each step navigation routing relies only on local information, i.e. distance to target, to determine the connection along which a message or signal is propagated. Communicability, finally, takes advantage of all available pathways simultaneously, which may require a prohibitive amount of energy and exceed the brain's metabolic limits [24]. In any case, these findings extend past work investigating the brain's strategies for communication [12, 19, 21, 22] and motivate future work linking computational and algorithmic approaches to known neurobiology [17].

Interareal similarity. A brain area's functionality is derived from its connectivity profile, i.e. its pattern of incoming and outgoing connections [25, 26]. The connectivity profile of area i is defined as the vector $\mathbf{w}_i = [w_{i1}, \ldots, w_{iN}]$. Regions with similar connectivity profiles have the capacity to receive and deliver similar input and output signals, and are therefore thought to play roughly equivalent functional roles within the network [27–30]. To measure the functional relatedness of two areas i and j, we can calculate the similarity of their connectivity profiles as the cosine of the angle, θ_{ij}, formed by the vectors $\mathbf{w}_i$ and $\mathbf{w}_j$:

$$S_{ij} = cos(\theta_{ij}) = \frac{\mathbf{w}_i \cdot \mathbf{w}_j}{\|\mathbf{w}_i\|\|\mathbf{w}_j\|}. \quad (2)$$

Note that for directed networks, we define an area's connectivity profile to include both its incoming and outgoing connections. That is, $\mathbf{w}_i = [w_{i1}, \ldots, w_{iN}, w_{1i}, \ldots, w_{Ni}]$.

Short- and long-distance connections. While a brain area's functionality depends on its own connectivity profile, it also depends on the connectivity profiles of its neighbors. An area's neighbors can have connectivity profiles dissimilar from its own and therefore can contribute unique inputs to that area or offer novel targets for that area's outgoing connections.

We measure the uniqueness of inputs and outputs using cosine similarity. Specifically, we compare the connectivity profiles of area i's neighbors linked by short- and long-distance connections. Let $\Gamma_i = \{j : w_{ij} \neq 0\}$ be the set of i's neighbors. These neighbors can be subdivided into short- and long-distance subsets: $\Gamma_i^{\text{short}} = \{j : w_{ij} > 0, e_{ij} \leq \tau_e^{\text{short}}\}$ and $\Gamma_i^{\text{long}} = \{j : w_{ij} > 0, e_{ij} \geq \tau_e^{\text{long}}\}$. Here, $\tau_e^{\text{short/long}}$ represent distance cutoffs below or above which we consider a neighbor to be short- *versus* long-distance with respect to area i's location.

To demonstrate that area i's short- *versus* long-distance neighbors have dissimilar connectivity profiles and therefore unique inputs and outputs, we computed the cosine similarity of their mean connectivity profiles.

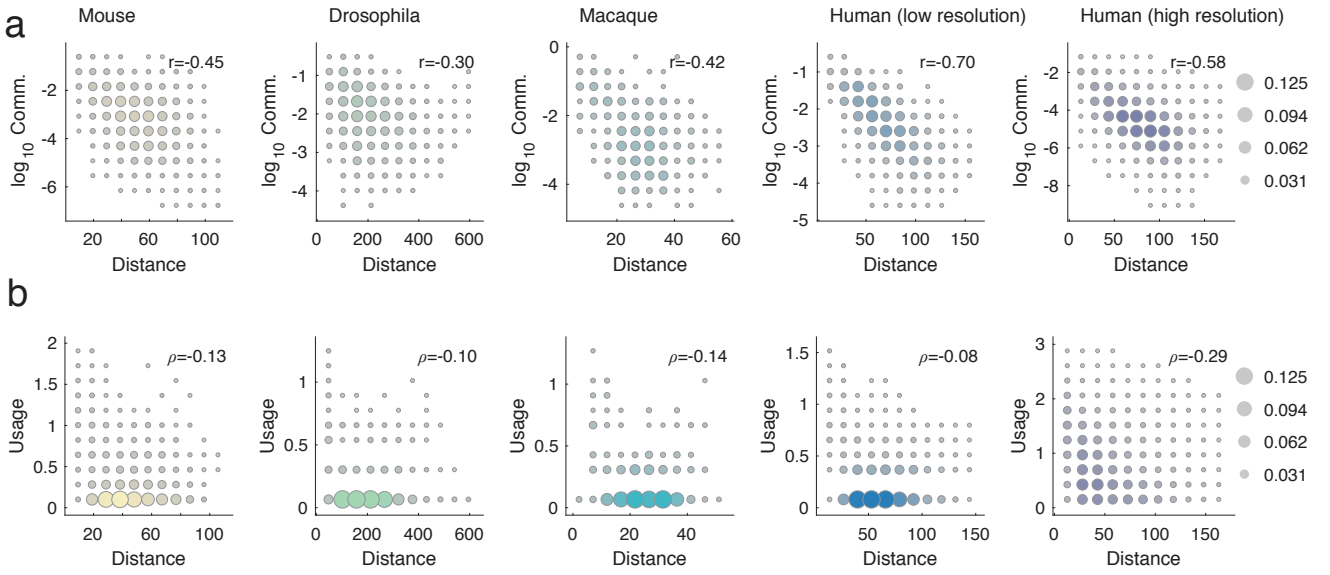

FIG. 4. **Relationship of communicability and navigation routing edge usage with Euclidean distance.** (**a**) Scatterplot and Pearson and Spearman correlation coefficients of interregional communicability with Euclidean distance for each network dataset. (**b**) Scatterplot of edge usage (number of times an edge is utilized by navigation routing) with Euclidean distance for each network dataset. The size and color of each point is proportional to density of data points in that neighborhood.

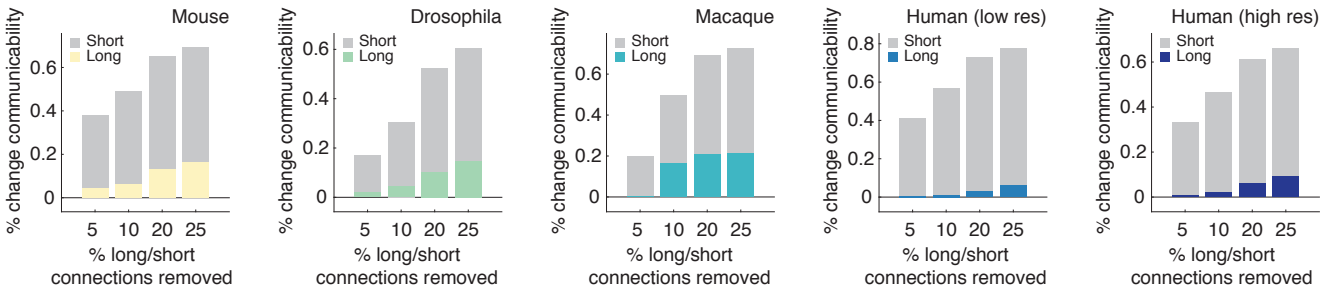

FIG. 5. **Effect of removing connections on mean inter-areal communicability.** Each panel depicts the percent change in mean interregional communicability as a result of removing either 5%, 10%, 20%, or 25% of the longest/shortest connections in the network. Note: all changes in communicability are negative, for the sake of visualization we flipped all signs.

The mean connectivity profile of $\Gamma_i^{\text{short/long}}$, is defined as:

$$\mathbf{w}_i^{\text{short/long}} = \sum_{j \in \Gamma_i^{\text{short/long}}} \mathbf{w}_j. \tag{3}$$

The similarity of $\mathbf{w}_i^{\text{short}}$ and $\mathbf{w}_i^{\text{long}}$, $S_{i^{\text{short}}i^{\text{long}}}$ was compared to a randomized null model, in which the network's topology was kept fixed, but where nodes' locations were randomly permuted. This procedure tests the null hypothesis that the (dis)similarity of connectivity profiles from short-range and long-distance neighbors could arise under random spatial embeddings as a result of the network's topology alone.

Redundancy of long-distance connections. Complex networks are subject to perturbations and their components can degrade over time, processes that compromise network function [31]. To counter these processes, many systems exhibit structural degeneracy in which a multiplicity of components play the same or similar functional roles [32]. In the event that some of these components are damaged, system function is maintained by the remaining undamaged components. We hypothesized that if the brain's long-distance connections were organized to provide unique and specific inputs and outputs, then these pathways should exhibit structural degeneracy. To test this hypothesis, we determined whether ar-

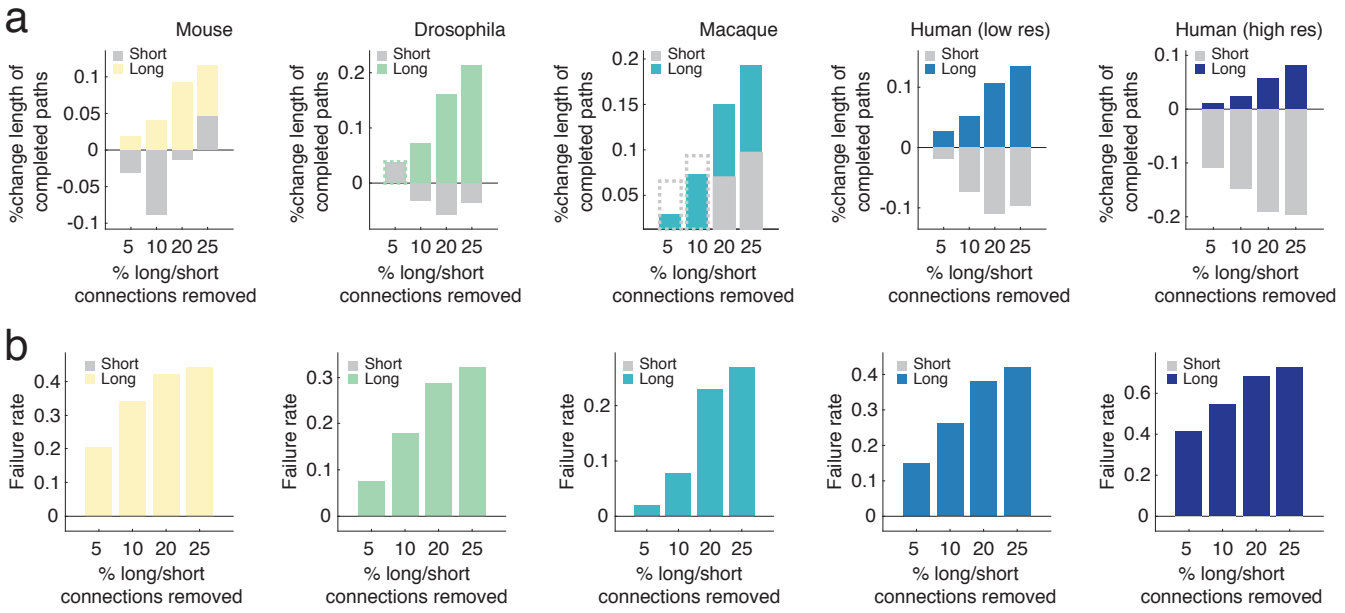

FIG. 6. **Effect of removing connections on navigation routing.** (**a**) Each panel depicts the percent change in mean path length as a result of removing either 5%, 10%, 20%, or 25% of the longest/shortest connections in the network. Note that mean path length is only computed over source/target pairs for which navigation routing was successful. (**b**) Failure rate as a result of removing short/long connections. Note that for all but the high-resolution network, removing long-distance connections resulted in a failure rate of 0% (all successes). For the high resolution network the failure rates were all less than 0.0007 (seven hundredths of a percent).

eas' long-distance connectivity profiles were more similar to one another than expected in a randomized null model in which the network's degree sequence, edge-length distribution, and edge-weight distribution were maintained exactly. This object was accomplished by retaining connections and weights whose lengths, $e_{ij} \geq \tau_e^{\text{long}}$. Using this network of long-distance connections only, we computed the cosine similarity for every pair of brain area connectivity profiles and computed the average similarity across all pairs:

$$\langle S^{\text{long}} \rangle = \frac{2}{N(N-1)} \sum_{i,j>i} S_{ij}^{\text{long}}. \tag{4}$$

Larger values of $\langle S^{\text{long}} \rangle$ indicate greater levels of structural degeneracy.

Modularity maximization. Many complex systems, including brain networks, exhibit rich meso-scale structure such that their nodes can be meaningfully partitioned into clusters [33]. Modular organization, in which clusters represent weakly interacting sub-systems called modules or communities, is a well-described phenomenon in both structural and functional brain networks [34]. Here, we use modularity maximization to uncover network modules [35]. Modularity maximization seeks to partition nodes into modules such that the intra-modular density of connections maximally exceeds that of a null connectivity model. This is accomplished by heuristically maximizing the modularity quality function:

$$Q(\gamma) = \sum_{ij} b_{ij} \delta(c_i, c_j). \tag{5}$$

Here, $b_{ij} = w_{ij} - \gamma \cdot p_{ij}$, where w_{ij} and p_{ij} are the observed and expected weights of the connection between nodes i and j. The resolution parameter, γ, scales the relative importance of p_{ij} and determines the number and size of detected modules. Node i's module assignment is encoded as $c_i \in \{1, \ldots, K\}$. Here, $\delta(c_i, c_j)$ is the Kronecker delta function and is equal to 1 when $c_i = c_j$ and zero otherwise. Effectively, Q is computed as a summation over node pairs, $\{i, j\}$, that fall within modules and is maximized when these pairs are more strongly connected than anticipated.

Structural network modules. We applied modularity maximization to each of the five structural network datasets. For structural brain networks, we used a null connectivity model that preserves nodes' binary and weighted degrees but otherwise allowed connections to be formed at random. For undirected networks we defined a modularity matrix with elements $b_{ij} = w_{ij} - \gamma p_{ij}$, where $p_{ij} = k_i k_j / 2m$. For the directed case, we defined, first, an asymmetric modularity matrix $b_{ij} = w_{ij} - k_i^{in} k_j^{out} / 2m$.

Following reference [36] exactly, we then defined a symmetric modularity matrix with elements $b'_{ij} = b_{ij} + b_{ji}$. Modularity, Q, was maximized with respect to this symmetric matrix.

We used a generalized version of the Louvain algorithm to maximize $Q(\gamma)$ [37], varying γ from 0.5 to 4.0 in increments of 0.1. At every value of γ we repeated the Louvain algorithm 100 times with random initial conditions. We selected the optimal value of γ by computing the pairwise similarity (z-score of the Rand index [38]) of partitions and focusing on local maxima of the γ *versus* median similarity curve. At local maxima, we generated a representative consensus partition from the partitions produced by the Louvain algorithm (see [13] for more details).

Functional network modules. The focus of this paper was on structural brain networks. However, to facilitate more direct functional interpretation of results obtained from analyses of human networks, we sought estimates of brain areas' functional system assignments. To obtain such estimates, we applied modularity maximization to correlation matrices constructed from independently acquired task-free fMRI BOLD data, the details of which have been described elsewhere [12, 39]. Briefly, this dataset comprised 40 subjects that underwent 9-minute resting-state scans, following which BOLD time series were extracted from the same $N_{low} = 82$ and $N_{high} = 1000$ areas as described above and subsequently averaged across subjects. We refer to these group-level correlation matrices as $\mathbf{r}_{low}$ and $\mathbf{r}_{high}$.

Applying modularity maximization to functional brain networks to obtain system labels requires extra care for several reasons. First, establishing a consistent set of system identities across scales is non-trivial because low- and high-resolution functional network datasets are generally constructed and clustered independently from one another. Here, we mitigate this issue using a multi-layer network model and a generalization of modularity maximization to cluster both matrices simultaneously [40]. Specifically, we treated $\mathbf{r}_{low}$ and $\mathbf{r}_{high}$ as differently-sized layers inter-linked to one another through the rectangular correlation matrix, $\mathbf{r}_{low/high}$, whose elements encoded the correlation magnitude of activity in low-resolution areas with that of high-resolution areas. The resulting multi-layer network was flattened to have dimensions $\mathbf{r} \in \mathbb{R}^{1068\times1068}$, where the first 68 nodes represent cortical areas from the low-resolution functional network and the final 1000 nodes represent cortical areas from the high-resolution network.

Second, the null connectivity model must be compatible with whatever measure was used to define functional connection strength (in this case a Pearson correlation coefficient) [41]. As suggested by Bazzi *et al.* [42], we defined $p_{ij} = 1$ for all node pairs $\{i, j\}$, so that $b_{ij} = r_{ij} - \gamma$. This expression corresponds to a null model in which BOLD activity of all nodes is uniformly correlated with a magnitude of γ. As before, the free parameter γ determines the number and size of communities. We tested $\gamma \in [0, 0.3]$ in increments of 0.006 for a total of 51 possible values. We used the same strategy described earlier to identify γ values of interest and to obtain consensus partitions. This analysis resulted in nine modules consistent across low- and high-resolution datasets. Based on spatial topography and visual inspection, we named these modules: visual (VIS), temporal + precuneus (T+P), dorsal attention (DAN), somatomotor (SMN), salience (SAL), default mode (DMN), frontal (FR), control (CONT), and limbic networks (LIM) (Fig. 7b).

Participation coefficient. A network's meso-scale organization highlights groups of brain areas thought to perform similar functions. Once those groups were identified, we further characterized the functional roles of individual brain areas based on their structural interactions with modules. One popular measure for doing so is an area's participation coefficient, which measures the extent to which an area's connections are concentrated within a few modules or are distributed more evenly across many modules [43].

$$\mathcal{P}_i = 1 - \sum_{s=1}^{K} \left(\frac{k_{is}}{k_i} \right)^2. \tag{6}$$

Here $k_i = \sum_j w_{ij}$ is node i's weighted degree and $k_{is} = \sum_{j\in s} w_{ij}$ is the total weight of node i's connections to module s. Participation coefficients range from 0 to 1, where larger values indicate that connections are evenly spread over modules. An area's participation coefficient can be interpreted as a measure of its diversity of function [44].

Simulated functional connectivity. The network analyses described above are based on *structural* connectomes – network representations of the brain's physical and material pathways. These pathways are often interpreted in terms of their theoretical contributions to network function, e.g. by facilitating signalling between connected brain areas. This interpretation of brain structure influencing communication patterns can be made more explicit by using connectomes to constrain dynamical systems [45]. In this framework brain areas are associated with internal states that represent their instantaneous level of activity, which evolves over time as a function of their current state and that of their connected neighbors. These dynamics induce statistical inter-dependencies, e.g. covariances or correlations, among brain areas' activity over time. These inter-dependencies are then interpreted as estimates of the brain's *functional* connectivity.

Here, we follow recent work [46, 47] by estimating the inter-areal covariance matrix, $\mathbf{C}$, based on a linearization of Wilson-Cowan dynamics for neuronal populations [48]. The element $c_{ij} \in \mathbf{C}$ denotes the covariance of activity in area i with that of area j (Fig. 8a). In more detail, we

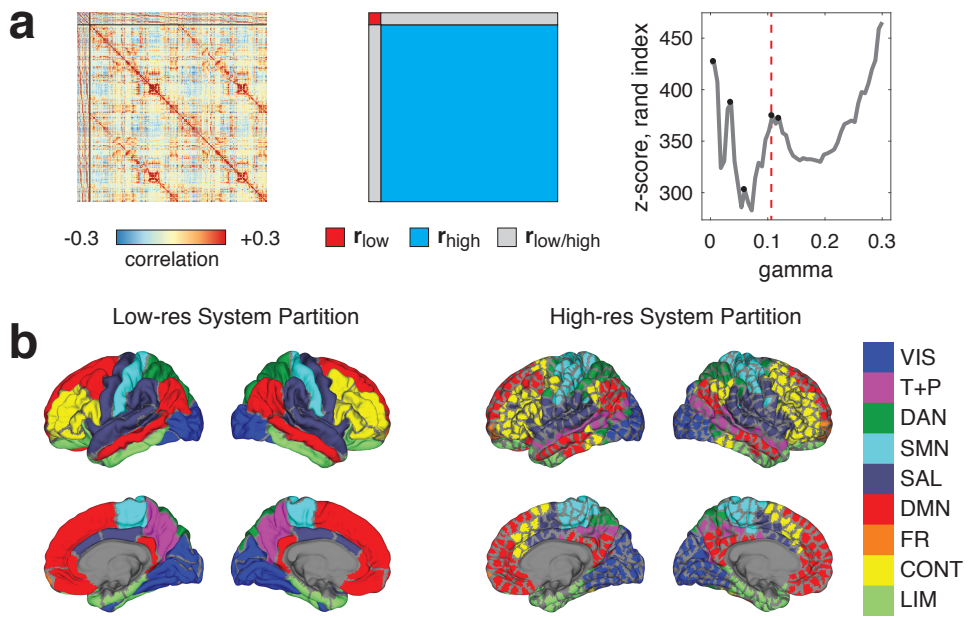

FIG. 7. **Multi-layer modularity maximization for functional systems.** (*a,left*) Flattened multi-layer correlation matrix, $\mathbf{r}$. (*a,middle*) Illustration of multi-layer matrix divided into low- and high-resolution layers and inter-layer coupling. (*a,right*) Similarity of detected partitions as a function of the resolution parameter, γ. Local maxima are shown as black circles. The dotted red line depicts the value of γ used to construct the system partitions. *b* Spatial topography of low- and high-resolution versions of the system partitions.

let $\mathbf{u}(t) = \{u_1(t), \ldots, u_N(t)\}$ be the vector of brain areas' states (activity levels) at time t. Under these dynamics, brain areas' states evolve as:

$$\mathbf{u}(t + \Delta t) = \mathbf{A}\mathbf{u}(t) + \xi(t), \tag{7}$$

where $\xi(t)$ is uncorrelated Gaussian noise and Δt is a single time step. Here, the generalized coupling matrix, $\mathbf{A}$, is based on the structural connectivity matrix, $\mathbf{W}$, and was defined as:

$$\mathbf{A} = (1 - \alpha\Delta t)\mathbf{I} + \mathbf{W}\Delta t, \tag{8}$$

where α is a leak variable within each brain area and $\mathbf{I}$ is the identity matrix. As in Honey *et al.* [47], we fixed $\alpha = 2$.

Conveniently, Galán [48] showed that brain areas' pairwise covariances (summarized by the matrix $\mathbf{C} \in \mathbb{R}^{N\times N}$) can be estimated directly from the spectral properties of $\mathbf{A}$ and the covariance of the noise terms $\xi(t)$. As with covariance matrices estimated from recorded time series of brain activity, we interpret $\mathbf{C}$ as an estimate of functional connectivity. See Galán [48] for more details.

We can use $\mathbf{C}$ to estimate each brain areas' functional diversity. This entailed defining network modules based on structural connectivity, and then using those modules in conjunction with $\mathbf{C}$ to calculate each brain area's participation coefficient, $\mathcal{P}_i$. Intuitively, $\mathcal{P}_i$ is close to 1 when its activity covaries uniformly with other brain areas in different modules, and is close to 0 when its activity covaries mostly with brain areas in its own module. Therefore, the average participation coefficient over all brain regions, $\frac{1}{N}\sum_i \mathcal{P}_i$, serves an estimate of functional diversity in the system.

Space- and topology-preserving null model. An essential component of network analysis is the contextualization of summary statistics by comparing their observed values against a distribution generated under a null connectivity model. Of such models, perhaps the best known is the edge swapping algorithm in which a network's edges are randomly rewired while exactly preserving each node's degree [49]. Effectively, this model tests the hypothesis that an observed summary statistic is common in the space of all networks with the same degree sequence as the observed network. This model, however, rewires networks without taking into account nodes' locations in space and the influence of inter-nodal distances on network topology.

Because this paper focuses on spatial statistics, we found it imperative to explore null models that preserve both topological and spatial characteristics of the ob-

served network. Accordingly, we developed a novel model in which we preserve a network's edge weight distribution and each node's degree exactly, while approximately preserving the edge length distribution and the relationship of edge length and weight. Intuitively, our model worked by swapping edges as in [49], but restricting pairs of edges that can be swapped so that the two new edges created *via* swapping have approximately the same lengths as those being replaced. This procedure takes place on a binary, unweighted network. After rewiring a predefined number of edges, weights are added back to the network. Because the rewired network has approximately the same edge length distribution, the original edges can be added back using a rank-preserving procedure (i.e. the shortest connection in the rewired network gets the weight of the shortest connection in the original network, the second shortest connection in the rewired network gets the weight of the second shortest connection in the original network, etc.). As a result, the rewired network preserves a select set of topological and spatial features of the original network (Fig. 8). Code for implementing this model is available at `https://www.richardfbetzel.com/code`.

In the main text we calculated the distribution of the lengths of all edges involved in shortest paths. We contextualized these results by comparing observed distributions against those obtained from a null model in which the network's connectivity structure was preserved but where nodes' spatial locations were randomly permuted. This model tests the hypothesis that the observed edge length distribution could arise given the topology of the network and random spatial embeddings alone. The results of this test demonstrated that the observed distributions exhibited far more short-range connections than would be expected under the null model, suggesting that shortest paths favor strong, short-range connections.

In addition, we tested a more conservative null model in which we randomly rewired networks while preserving their degree sequence and edge weight distribution exactly, and while preserving edge length distribution and the weight-length relationships approximately. For each random network we recalculated its shortest path structure and the length distribution of edges involved. We found that, in general, the median length of shortest path edges in the observed networks was smaller than that of the randomly rewired networks (Fig. 9). This result held for all datasets and for all values of α with the exception being the macaque dataset with $\alpha > 2$. Above this α value, the distribution of edge lengths involved in shortest paths in the observed network was not distinguishable from that of the null model. It is important to note, however, that the macaque dataset is incomplete (no available tract-tracing data from 64 of the 93 total regions) and should therefore be approached with greater caution and skepticism than the other datasets. The convergence of results using both null models increases confidence that our findings are robust and not driven by idiosyncrasies of any single null model.

Thresholding connectomes. Analysis of brain networks is complicated by false positives that arise from inaccurate reconstructions or noise in the acquisition process. The presence of spurious connections can alter statistics estimated on a network, resulting in incorrect or misleading conclusions [50, 51]. One strategy for mitigating the unwanted influence of false positives is to simply discard all connections whose weight is below some pre-defined threshold. While this approach has been shown to improve network analysis [52], how to best choose this threshold remains an open question.

Here, we explore the effect of removing weak connections on our conclusions. In particular, we discarded the weakest 1%, 2%, 4%, and 8% and reanalyzed the thresholded networks. First, we calculated weighted shortest paths and edge usage as a function of distance. Because weak connections rarely factor into shortest paths, thresholding them from the network had almost no influence on edge usage in shortest paths. This effect gets even stronger by the weight-to-length remapping ($L_{ij} = W_{ij}^{-\alpha}$), which placed added emphasis on strong connections compared to weak connections. As a result, we found that edge usage statistics (Fig. 10) and the lengths of weighted shortest paths (Fig. 11) remained dominated by short-range connections, as reported in the main text. In fact, for many networks the shortest path structure was entirely unchanged. We also computed the similarity of areas' long-distance connectivity profiles and compared those results to that of a null model (Fig. 12). Once again, and in agreement with the main text, we found that our results are unchanged qualitatively after thresholding away weak connections. These findings support the hypothesis that the results reported in the main text are robust to reasonable choices of edge weight thresholds.

[1] Seung Wook Oh, Julie A Harris, Lydia Ng, Brent Winslow, Nicholas Cain, Stefan Mihalas, Quanxin Wang, Chris Lau, Leonard Kuan, Alex M Henry, *et al.*, "A mesoscale connectome of the mouse brain," Nature **508**, 207 (2014).

[2] Mikail Rubinov, Rolf JF Ypma, Charles Watson, and Edward T Bullmore, "Wiring cost and topological participation of the mouse brain connectome," Proceedings of the National Academy of Sciences **112**, 10032–10037 (2015).

[3] J Z Kim, J M Soffer, A E Kahn, J M Vettel, F Pasqualetti, and D S Bassett, "Role of graph architecture in controlling dynamical networks with applications to neural systems," Nature Physics **14**, 91–98 (2018).

[4] Richard F Betzel, John D Medaglia, and Danielle S Bassett, "Diversity of meso-scale architecture in hu-

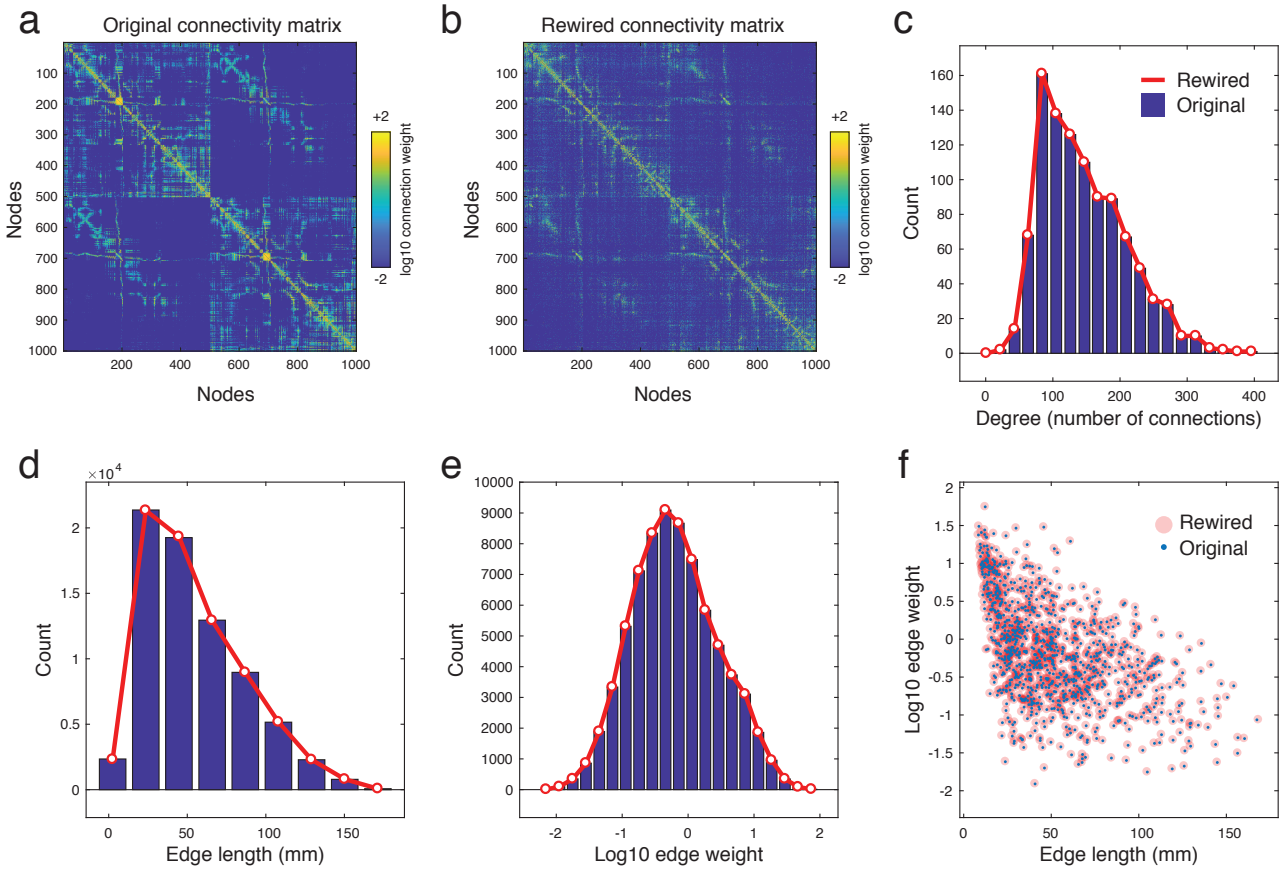

FIG. 8. **Rewiring model.** (**a**) Original network. (**b**) Rewired network. (**c**) Degree, (**d**) edge length, and (**e**) edge weight distribution of observed (blue) and rewired network (red). (**f**) Weight-length relationship is preserved. Panel depicts a randomly selected subsample of 1000 edges; this choice enhanced clarity of the presentation. Red circles are obtained from the rewired network while blue dots were obtained from the original network.

man and non-human connectomes," arXiv preprint arXiv:1702.02807 (2017).

[5] Ann-Shyn Chiang, Chih-Yung Lin, Chao-Chun Chuang, Hsiu-Ming Chang, Chang-Huain Hsieh, Chang-Wei Yeh, Chi-Tin Shih, Jian-Jheng Wu, Guo-Tzau Wang, Yung-Chang Chen, *et al.*, "Three-dimensional reconstruction of brain-wide wiring networks in drosophila at single-cell resolution," Current Biology **21**, 1–11 (2011).

[6] Chi-Tin Shih, Olaf Sporns, Shou-Li Yuan, Ta-Shun Su, Yen-Jen Lin, Chao-Chun Chuang, Ting-Yuan Wang, Chung-Chuang Lo, Ralph J Greenspan, and Ann-Shyn Chiang, "Connectomics-based analysis of information flow in the drosophila brain," Current Biology **25**, 1249–1258 (2015).

[7] Jacob C Worrell, Jeffrey Rumschlag, Richard F Betzel, Olaf Sporns, and Bratislav Mišić, "Optimized connectome architecture for sensorymotor integration," Network Neuroscience (2017).

[8] Nikola T Markov, MM Ercsey-Ravasz, AR Ribeiro Gomes, Camille Lamy, Loic Magrou, Julien Vezoli, P Misery, A Falchier, R Quilodran, MA Gariel, *et al.*, "A weighted and directed interareal connectivity matrix for macaque cerebral cortex," Cerebral cortex **24**, 17–36 (2012).

[9] Nikola T Markov, Maria Ercsey-Ravasz, Camille Lamy, Ana Rita Ribeiro Gomes, Loïc Magrou, Pierre Misery, Pascale Giroud, Pascal Barone, Colette Dehay, Zoltán Toroczkai, *et al.*, "The role of long-range connections on the specificity of the macaque interareal cortical network," Proceedings of the National Academy of Sciences **110**, 5187–5192 (2013).

[10] Fang-Cheng Yeh, Van Jay Wedeen, and Wen-Yih Isaac Tseng, "Estimation of fiber orientation and spin density distribution by diffusion deconvolution," Neuroimage **55**, 1054–1062 (2011).

[11] Leila Cammoun, Xavier Gigandet, Djalel Meskaldji, Jean Philippe Thiran, Olaf Sporns, Kim Q Do, Philippe Maeder, Reto Meuli, and Patric Hagmann, "Mapping the human connectome at multiple scales with diffusion spectrum mri," Journal of neuroscience methods **203**, 386–397 (2012).

[12] Bratislav Mišić, Richard F Betzel, Azadeh Nematzadeh, Joaquin Goñi, Alessandra Griffa, Patric Hagmann, Alessandro Flammini, Yong-Yeol Ahn, and Olaf Sporns,

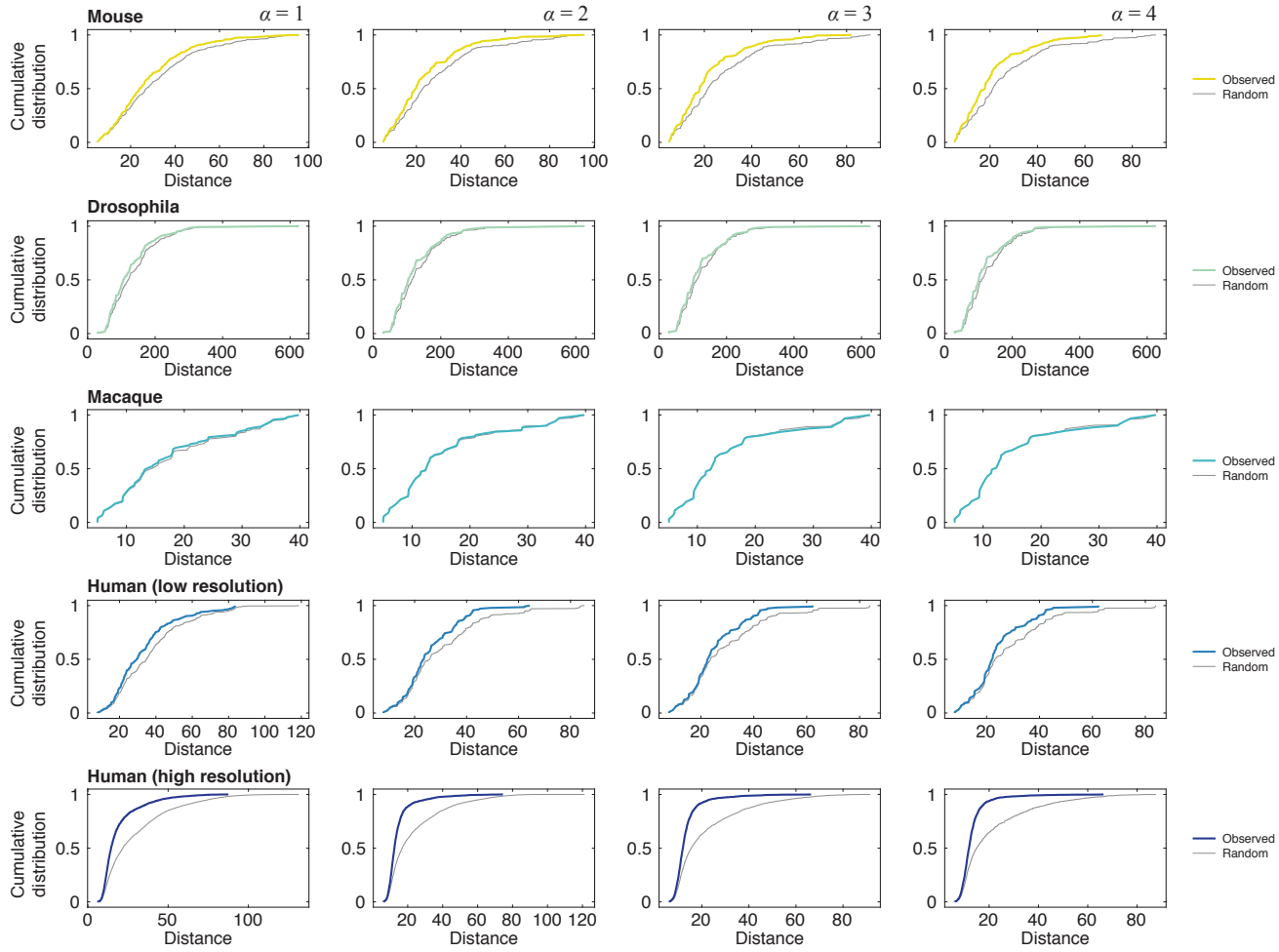

FIG. 9. **Shortest path usage by connection length.** We computed shortest path structure for observed and randomly rewired networks. We aggregated the lengths of all connections that participated in at least one shortest path and examined their cumulative distribution (observed distribution shown in color; mean cumulative distribution over 100 realizations of the null model shown in gray). We repeated this analysis for different values of the weight-to-length parameter, α. To compare observed and random curves statistically we computed p-values non-parametrically as the fraction of rewired networks whose median connection length was smaller than that of the observed networks. With the exception of the macaque dataset, all other networks exhibited shorter connection lengths than the corresponding rewired networks ($p < 0.05$).

"Cooperative and competitive spreading dynamics on the human connectome," Neuron **86**, 1518–1529 (2015).

[13] Richard F Betzel, John D Medaglia, Lia Papadopoulos, Graham L Baum, Ruben Gur, Raquel Gur, David Roalf, Theodore D Satterthwaite, and Danielle S Bassett, "The modular organization of human anatomical brain networks: Accounting for the cost of wiring," Network Neuroscience (2017).

[14] James A Roberts, Alistair Perry, Gloria Roberts, Philip B Mitchell, and Michael Breakspear, "Consistency-based thresholding of the human connectome," Neuroimage **145**, 118–129 (2017).

[15] Edsger W Dijkstra, "A note on two problems in connexion with graphs," Numerische mathematik **1**, 269–271 (1959).

[16] Ulrik Brandes, "A faster algorithm for betweenness centrality," Journal of mathematical sociology **25**, 163–177 (2001).

[17] Andrea Avena-Koenigsberger, Bratislav Misic, and Olaf Sporns, "Communication dynamics in complex brain networks," Nature Reviews Neuroscience **19**, 17 (2018).

[18] Ernesto Estrada and Naomichi Hatano, "Communicability in complex networks," Physical Review E **77**, 036111 (2008).

[19] Jonathan J Crofts and Desmond J Higham, "A weighted communicability measure applied to complex brain networks," Journal of the Royal Society Interface , rsif–2008 (2009).

[20] Marian Boguna, Dmitri Krioukov, and Kimberly C Claffy, "Navigability of complex networks," Nature

Physics **5**, 74–80 (2009).
[21] Caio Seguin, Martijn P van den Heuvel, and Andrew Zalesky, "Navigation of brain networks," arXiv preprint arXiv:1801.07938 (2018).
[22] Antoine Allard and M Serrano, "Navigable maps of structural brain networks across species," arXiv preprint arXiv:1801.06079 (2018).
[23] Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguná, "Hyperbolic geometry of complex networks," Physical Review E **82**, 036106 (2010).
[24] Simon B Laughlin and Terrence J Sejnowski, "Communication in neuronal networks," Science **301**, 1870–1874 (2003).
[25] Semir Zeki and Stewart Shipp, "The functional logic of cortical connections," Nature **335**, 311–317 (1988).
[26] Richard E Passingham, Klaas E Stephan, and Rolf Kötter, "The anatomical basis of functional localization in the cortex," Nature reviews. Neuroscience **3**, 606 (2002).
[27] Claus C Hilgetag, Rolf Kötter, Klaas E Stephan, and Olaf Sporns, "Computational methods for the analysis of brain connectivity," in *Computational neuroanatomy* (Springer, 2002) pp. 295–335.
[28] H Barbas and DN Pandya, "Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey," Journal of Comparative Neurology **286**, 353–375 (1989).
[29] Malcolm P Young, "The organization of neural systems in the primate cerebral cortex," Proceedings of the Royal Society of London B: Biological Sciences **252**, 13–18 (1993).
[30] Jack W Scannell, Colin Blakemore, and Malcolm P Young, "Analysis of connectivity in the cat cerebral cortex," Journal of Neuroscience **15**, 1463–1483 (1995).
[31] Réka Albert, Hawoong Jeong, and Albert-László Barabási, "Error and attack tolerance of complex networks," arXiv preprint cond-mat/0008064 (2000).
[32] Giulio Tononi, Olaf Sporns, and Gerald M Edelman, "Measures of degeneracy and redundancy in biological networks," Proceedings of the National Academy of Sciences **96**, 3257–3262 (1999).
[33] Mark EJ Newman, "Communities, modules and large-scale structure in networks," Nature physics **8**, 25 (2012).
[34] Olaf Sporns and Richard F Betzel, "Modular brain networks," Annual review of psychology **67**, 613–640 (2016).
[35] Mark EJ Newman and Michelle Girvan, "Finding and evaluating community structure in networks," Physical review E **69**, 026113 (2004).
[36] Elizabeth A Leicht and Mark EJ Newman, "Community structure in directed networks," Physical review letters **100**, 118703 (2008).
[37] Inderjit S Jutla, Lucas GS Jeub, and Peter J Mucha, "A generalized louvain method for community detection implemented in matlab," URL http://netwiki. amath. unc. edu/GenLouvain (2011).
[38] Amanda L Traud, Eric D Kelsic, Peter J Mucha, and Mason A Porter, "Comparing community structure to characteristics in online collegiate social networks," SIAM review **53**, 526–543 (2011).
[39] Richard F Betzel, Alessandra Griffa, Andrea Avena-Koenigsberger, Joaquín Goñi, Jean-Philippe Thiran, Patric Hagmann, and Olaf Sporns, "Multi-scale community organization of the human structural connectome and its relationship with resting-state functional connectivity," Network Science **1**, 353–373 (2013).
[40] Peter J Mucha, Thomas Richardson, Kevin Macon, Mason A Porter, and Jukka-Pekka Onnela, "Community structure in time-dependent, multiscale, and multiplex networks," science **328**, 876–878 (2010).
[41] Mel MacMahon and Diego Garlaschelli, "Community detection for correlation matrices," arXiv preprint arXiv:1311.1924 (2013).
[42] Marya Bazzi, Mason A Porter, Stacy Williams, Mark McDonald, Daniel J Fenn, and Sam D Howison, "Community detection in temporal multilayer networks, with an application to correlation networks," Multiscale Modeling & Simulation **14**, 1–41 (2016).
[43] Roger Guimera and Luis A Nunes Amaral, "Functional cartography of complex metabolic networks," nature **433**, 895 (2005).
[44] MA Bertolero, BTT Yeo, and M D'Esposito, "The diverse club: The integrative core of complex networks," arXiv preprint arXiv:1701.01150 (2017).
[45] Mark W Woolrich and Klaas E Stephan, "Biophysical network models and the human connectome," Neuroimage **80**, 330–338 (2013).
[46] Andrea Avena-Koenigsberger, Joaquín Goñi, Richard F Betzel, Martijn P van den Heuvel, Alessandra Griffa, Patric Hagmann, Jean-Philippe Thiran, and Olaf Sporns, "Using pareto optimality to explore the topology and dynamics of the human connectome," Phil. Trans. R. Soc. B **369**, 20130530 (2014).
[47] CJ Honey, O Sporns, Leila Cammoun, Xavier Gigandet, Jean-Philippe Thiran, Reto Meuli, and Patric Hagmann, "Predicting human resting-state functional connectivity from structural connectivity," Proceedings of the National Academy of Sciences **106**, 2035–2040 (2009).
[48] Roberto F Galán, "On how network architecture determines the dominant patterns of spontaneous neural activity," PloS one **3**, e2148 (2008).
[49] Sergei Maslov and Kim Sneppen, "Specificity and stability in topology of protein networks," Science **296**, 910–913 (2002).
[50] Andrew Zalesky, Alex Fornito, Ian H Harding, Luca Cocchi, Murat Yücel, Christos Pantelis, and Edward T Bullmore, "Whole-brain anatomical networks: does the choice of nodes matter?" Neuroimage **50**, 970–983 (2010).
[51] Stamatios N Sotiropoulos and Andrew Zalesky, "Building connectomes using diffusion mri: Why, how and but," NMR in Biomedicine (2017).
[52] Andrew Zalesky, Alex Fornito, Luca Cocchi, Leonardo L Gollo, Martijn P van den Heuvel, and Michael Breakspear, "Connectome sensitivity or specificity: which is more important?" Neuroimage **142**, 407–420 (2016).

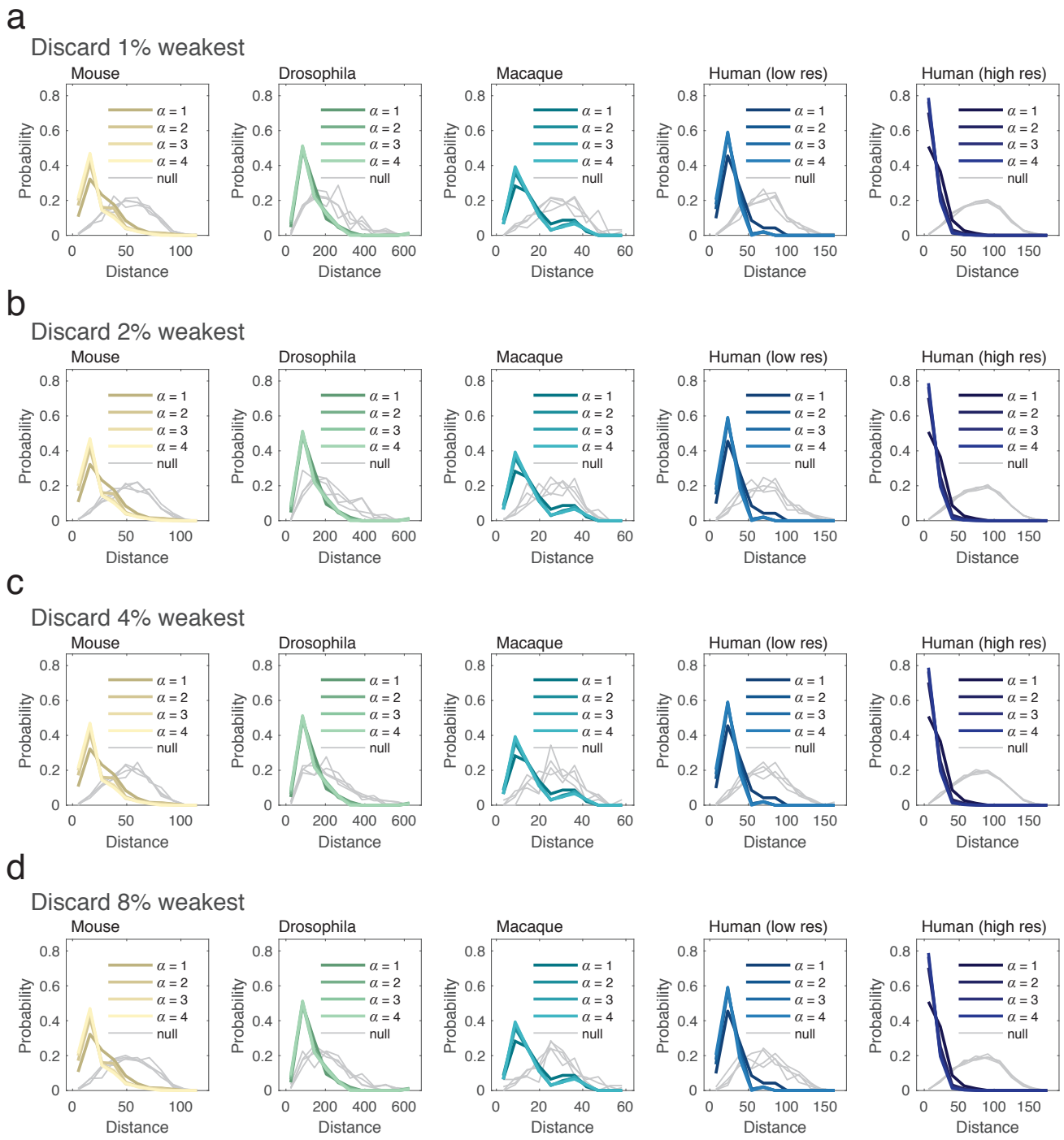

FIG. 10. **Effect of thresholding on lengths of connections involved in shortest paths.** Each panel depicts the edge length distribution of connections included in shortest paths in each of the five network datasets while thresholding away the 1%, 2%, 4%, and 8% weakest connections. Note that because shortest paths seek out the strongest connections (which are preserved even after thresholding), the shortest path structure and involved edges changed minimally across all datasets.

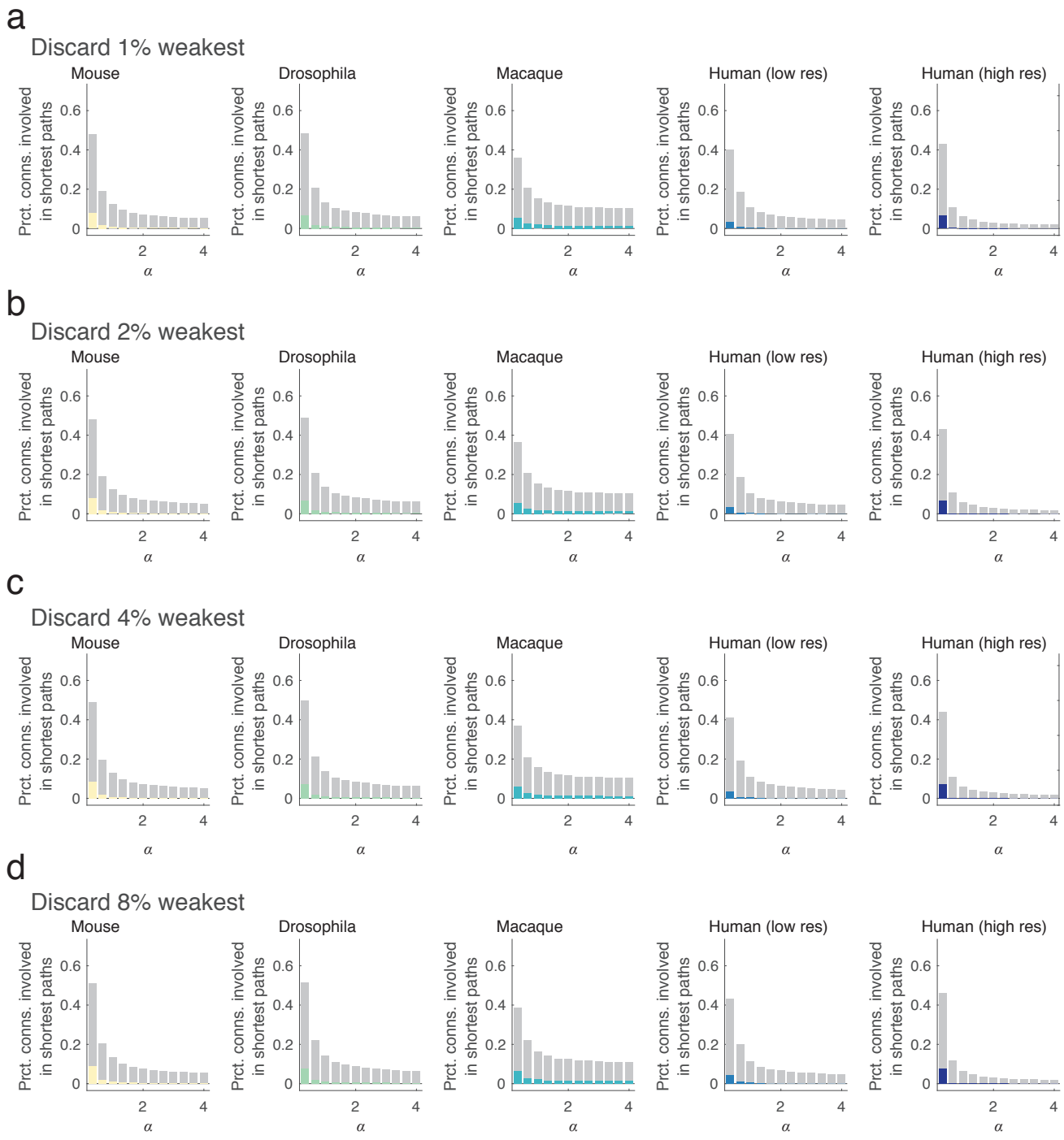

FIG. 11. **Effect of thresholding on composition of shortest paths.** Each panel depicts, in gray, the fraction of all edges involved in shortest paths. Shown in color are the fraction of those connections that are among the top 25% longest in the network. Rows depict the effect of thresholding the 1%, 2%, 4%, or 8% weakest connections on length composition of shortest paths.

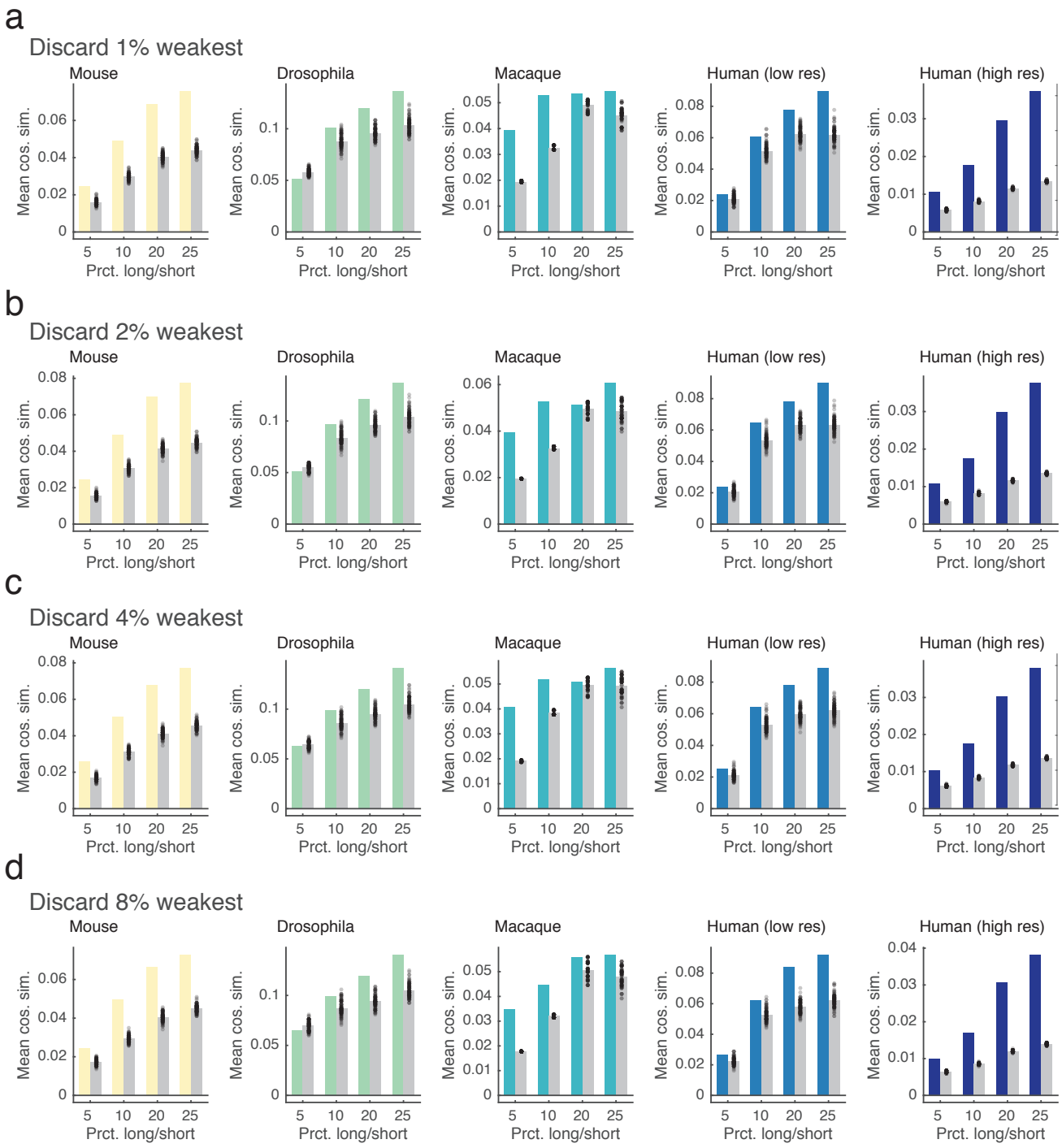

FIG. 12. **Effect of thresholding on similarity of long-distance inputs/outputs.** Each panel depicts the mean similarity of nodes' long-distance connectivity profiles (in color) and what is expected under a conservative null model (gray) in which nodes degree sequences and edge weight distributions are preserved exactly, and in which edge length distribution and weight-distance relationships are preserved to an excellent approximation. Each point represents a single random network realized under this null model. Rows depict the effect of thresholding the 1%, 2%, 4%, or 8% weakest connections on similarity of long-distance connectivity profiles.