Filtering information in imaging connectomics

Fabrizio De Vico Fallani\textsuperscript{1,2,*}, Vito Latora\textsuperscript{3,4}, and Mario Chavez\textsuperscript{2}

\textsuperscript{1}Inria, Aramis team, Paris, France
\textsuperscript{2}CNRS, UMR7225; Inserm, U1127; Sorbonne Universites, UPMC-P6 UMR S1127; ICM, Paris, France
\textsuperscript{3}Queen Mary University of London, School of Mathematical Sciences, London, UK
\textsuperscript{4}Università di Catania and INFN, Dipartimento di Fisica ed Astronomia, Catania, Italy

\textsuperscript{*}Corresponding author

Abstract

Mapping the connectome through non-invasive imaging tools allows to study the wiring diagram of the brain \textit{in-vivo}. The informative value of the resulting brain network depends on how much of the estimated connectivity is retained. We present a method, the efficiency cost optimization (ECO), to filter connectivity based on the optimization of the trade-off between the efficiency of a network and its wiring cost. As a result, we find an optimal connection density threshold that emphasizes the intrinsic properties of a given network, while preserving its sparsity. Such threshold can be determined \textit{a-priori}, since the number of connections to retain only depends on the network size. We validate this result on connectomes, from micro- to macro-scales, obtained with different imaging modalities. Finally, we test the potential of ECO in discriminating brain states with respect to alternative filtering approaches. This method advances our ability to analyze and compare imaging connectomes in a fast and principled way.

Keywords: Neuroimaging, Connectivity data, Complex networks, Thresholding

\* ARAMIS Inria team

\textit{Institut du Cerveau et de la Moelle épineire}
47, Boulevard de l’Hopital
75013, Paris, France
Tel.+33(0)157274294
Email. fabrizio.devicofallani@gmail.com
Introduction

Imaging connectomics uses neuroimaging techniques to map connections and/or interactions between different brain sites. Combined with tools from graph theory, imaging connectomics has considerably advanced our understanding of the brain structure and function from a system perspective.\(^1\)

Noninvasive neuroimaging is particularly attractive as it allows to map connectomes \textit{in-vivo} and quantify network organizational mechanisms underlying behavior, cognition, development as well as disease.\(^2,3\) Magnetic resonance imaging (MRI) and electro/magnetoencephalography (E/MEG) are frequently used to derive macro-scale connectomes whose nodes, or units, correspond to spatially remote brain sites.\(^4\) High-field MRI and genetically encoded calcium indicators are promising tools to image connectomes respectively at the scale of neuronal ensembles (meso-scale) and single neurons (micro-scale).\(^5\) Brain connectivity methods are typically used to estimate the links between the nodes. While anatomical connectivity (AC) measures the probability to find axonal pathways between brain areas, typically from diffusion MRI, functional connectivity (FC) rather calculates the temporal dependence between their neural processes as recorded, for instance, by functional MRI, EEG or MEG.\(^1\)

At this stage, the resulting connectomes correspond to maximally dense networks whose weighted links code for the strength of the connections between different brain nodes. Common courses in brain network analysis use thresholding procedures to filter information in these raw connectomes by retaining and binarizing a certain percentage of the strongest links (Supplementary Fig. 1). Despite the consequent information loss, these procedures are often adopted to mitigate the uncertainty carried by the weakest links and facilitate the interpretation of the network topology.\(^6\) In addition, they enable to use all the graph theoretic tools, which are mainly conceived for sparse and binary networks.\(^7\) At present, the choice of the specific value for such threshold remains arbitrary, so that scientists are obliged to explore brain network properties across a wide range of different candidates and eventually select one representative \textit{a-posteriori}.\(^8\) These approaches are extremely time-consuming for large connectomes and make difficult the comparison between many individuals or samples.\(^9\)

We introduce a criterion to select \textit{a-priori} an optimal threshold which captures the essential topology of a connectome while preserving its sparsity. Based on the perfect trade-off between two desirable but incompatible features - namely high global and local integration between nodes, and low connection density - this method is inherently motivated by the principle of efficiency and economy observed in many complex systems,\(^10\) including the brain.\(^11\)

Results

Filtering information as a network optimization problem

Global- and local-efficiency\(^12\) have revealed to be important graph quantities to characterize the structure of brain networks in terms of integration and segregation of information.\(^13\) Both anatomical and functional brain networks tend to exhibit relatively high values of global- and local-efficiency. At the same time they
also tend to minimize, for economical reasons, the number of their links leading to sparse networks.\textsuperscript{11}

Hence, we propose to determine a density threshold that filters out the weakest links and maximizes the ratio between the overall efficiency of a connectome and its wiring cost. We formally introduce a criterion to filter information in a given network by finding the optimal connection density $\rho$ that maximizes the quality function:

$$J = \frac{E_g + E_l}{\rho}$$  \hspace{1cm} (1)

where $E_g$ and $E_l$ represent respectively the global- and local-efficiency of a network (Online Method 1).

For both regular lattices and random networks, we proved analytically that the optimal density deriving from the maximization of $J$ reads $\rho \simeq 3/(n - 1)$, where $n$ is the network size, i.e., the number of nodes in the network (Supplementary Text). We confirmed this result (Supplementary Fig. 2a,b) through extensive numerical simulations (Online Methods 2), showing that it held true also in more realistic network models, such as in small-world networks\textsuperscript{14} (Fig. 1a) and in scale-free networks \textsuperscript{15} (Fig. 1b). Notably, the optimal density emphasized the intrinsic structural properties of all the implemented synthetic networks in terms of global- and local-efficiency (Fig. 1d,e and Supplementary Fig. 2d,e).

**Optimal density in connectomes derived from neuroimaging**

We computed the quality function $J$ in both micro- and macro-scale brain networks and we evaluated how the optimal density scaled as a function of the network size. We considered connectomes used in previously published studies that were obtained with different imaging modalities, from calcium imaging to EEG, and constructed with different brain connectivity methods, from Pearson-correlation to Granger-causality (Table 1).

For each connectome we applied a standard density-based thresholding. We started with the empty network by removing all the links ($\rho = 0$). Then, we reinserted and binarized one link at time, from the strongest to the weakest, until we obtained the maximally dense network ($\rho = 1$). At each step we computed $J$ and we recorded its profile as a function of $\rho$ (Fig. 1f). If a study presented more samples (individuals) with connectomes of the same size, we considered the group-averaged $J$ profile to improve the quality of the estimation (Supplementary Fig. 2c,f). The pooled density values, as returned by the maximization of $J$, followed the same scaling relationship that we reported for synthetic networks (Fig. 1c). This result confirms that also for brain networks we can assume that the optimal density threshold depends on the network size according to the same rule $\rho \simeq 3/(n - 1)$.

In conclusion, we introduced a criterion, named efficiency cost optimization (ECO), to select a threshold leading to sparse, yet informative brain networks. Such a threshold does depend neither on how the connectome is constructed nor on its underlying structure, and can be therefore selected \textit{a-priori}.
ECO discriminated network properties of different brain states

To illustrate the methodology, we considered connectomes from four different imaging modalities, namely EEG, MEG, fMRI, and DTI (Table 1). Because we do not know the true structure for these connectomes, we evaluated the ability of ECO to discriminate network properties of different brain states, i.e., healthy versus diseased, at individual level.

We characterized brain networks by calculating graph quantities at different topological scales, i.e., large (global- and local-efficiency, $E_g$ and $E_l$), intermediate (community partition, $P$; and modularity, $Q$), and small (node degree, $k_i$; and betweenness, $b_i$) (Online Method 3). To assess network differences between brain states, we measured distances between the values of the graph quantities obtained in the healthy group and those in the diseased group. We adopted the Mirkin index ($MI$) to measure distances between community partitions, and the divergent coefficient ($D$) for other graph quantities (Online Method 4).

We explored a wide range of density thresholds and, as expected, the value of the threshold affected the ability to separate network properties of different brain states (Fig. 2, Supplementary Fig. 3 and 4a). Notably, the choice $\rho = 3/(n−1)$ resulted among the best candidates in producing larger distances and improving discrimination. Furthermore, ECO overall outperformed alternative methods such as the minimum spanning tree (MST) and the planar maximally filtered graph (PMFG) (Fig. 3, Supplementary Fig. 4b, Supplementary Table 2).

Discussion

We introduced ECO to filter information in imaging connectomes whose links are statistical or probabilistic measures of brain connectivity. Conventional graph approaches evaluate brain network properties across a large and arbitrary number of thresholds. Eventually, they select a representative threshold a-posteriori that maximizes the separation between different brain states. ECO provides a theoretically grounded criterion to select an optimal threshold a-priori, drastically reducing the computational burden. Other approaches, similar in purpose to ECO, impose unnatural constraints on the filtered connectome. The minimum spanning tree (MST), for instance, leads to brain networks with a null clustering coefficient. The planar maximally filtered graph (PMFG) tries to alleviate this bias by allowing closed loops, but still forces planarity. Conversely, ECO does not impose any constraint and lets the intrinsic network structure of a connectome to emerge.

In addition, ECO is based on fundamental principles of complex systems. Maximizing global- and local-efficiency with respect to connection density means emphasizing the integration and segregation properties of a connectome while keeping a biologically plausible wiring cost. This rationale dovetails with current evidence showing that advantageous topological properties, such as economic small-world architectures, tend to be maximized in brain networks, and that, in general, sparsity increases robustness of complex systems. Other combinations could have been considered when conceiving the quality function $J$. For example, in authors introduced the cost-efficiency $E_g − \rho$, which, however, did not include the clustering counterpart. That quality function, as well as other ones that we tested, did not
exhibit an optimal density and was therefore not considered (Supplementary Text).

ECO makes use of density thresholds. Hence, filtered connectomes having same number of nodes, will have same number of links. On the one hand, this ensures that differences between brain network properties are not merely due to differences in the connection density. On the other hand, ECO does not allow a direct evaluation of neural processes altering the number of links; yet it does inform on the consequent (re)organizational mechanisms. ECO is based on a graph topological criterion and cannot filter out possible false positives (i.e., spurious links) resulting from biased brain connectivity estimates. This method assumes that the weighted links of the raw connectomes had been previously statistically validated, either maintained or canceled. We conceived ECO to filter imaging connectomes with applications ranging from cognitive to clinical and computational neuroscience. Given its generality and simplicity, we anticipate that ECO will facilitate the analysis of interconnected systems where the need of sparsity is plausible and links are weighted estimates of connectivity. This is, for example, the case of functional networks in system biology, where links are derived from transcriptional or phenotypic profiling, and genetic interactions.

Online Methods

1. Topological properties of the quality function \( J \)

The proposed quality function can be seen as a particular case of a general family of functions of the form \( f(E_g, E_l, \rho) \). When combining different graph quantities, proper normalizing procedures need to be used. By definition, each of the three quantities \( E_g, E_l \) and \( \rho \) is normalized in the range \([0, 1]\), and \( E_g \) and \( E_l \) are non-decreasing functions of \( \rho \). However, since the efficiency is based on shortest paths between nodes, a concept which is not directly captured by the density, a scaling factor might be necessary to normalize changes among those quantities.

We therefore considered a more general form of \( J \) than that in Eq. (1) by introducing two distinct dependencies on the connection density, i.e., \( \beta \rho \) and \( \rho^\beta \), where \( \beta \) is a tunable control parameter. For values around \( \beta = 1 \), the parameter had no influence on the returned optimal density (Supplementary Text 3). Hence, for the sake of simplicity, we deliberately chose \( \beta = 1 \) that corresponds to the original expression of the quality function in Eq. (1), i.e., \( J = (E_g + E_l)/\rho \).

When \( \rho = 0 \) in Eq. (1), then both global- \( E_g \) and local-efficiency \( E_l \) are null leading to an indefinite form. As density slightly increases (\( 0 < \rho < \epsilon \), with \( \epsilon \) sufficiently small) it can be demonstrated that \( J \) tends to 1. In fact, in this range, the probability to find at least three nodes connected together (a triangle) is extremely low. By definition, \( E_l = 0 \) in absence of at least one triangle and therefore \( J \simeq E_g/\rho \). By considering the definitions of \( E_g \) and \( \rho \), this quantity can be rewritten as \( E_g/\rho = 1/m \sum_{i \neq j} 1/d_{i,j} \), where \( m \) is the number of existing links and \( d_{i,j} \) is the distance between the nodes \( i \) and \( j \). In a generic network with \( m \) links there are at least \( m \) pairs of nodes directly connected (i.e., \( d_{i,j} = 1 \)). This means that the sum in the latter equation is bounded from below by \( m \) in the case of isolated pairs of connected nodes (\( m = n/2 \)) or in the trivial case of \( m = 1 \). It follows that \( J \to 1 \) when there are relatively few links in a network.
When $\rho$ tends to 1, it is trivial to see from equation (1) that $J \rightarrow 2$, as both $E_g$ and $E_l$ tend to one. For intermediate density ranges ($\varepsilon < \rho \ll 1 - \varepsilon$) the analytical estimate of $J$ is not trivial since $E_g$ and $E_l$ depend on the network topology which is, in general, unknown.

2. Numerical simulations for small-world and scale-free networks

Small-world networks were generated according to the Watts-Strogatz (WS) model \(^{14}\) with a rewiring probability $p_{ws} = 0.1$. Scale-free networks were generated according to the Barabasi-Albert (BA) model \(^{15}\).

In the first simulation, we considered undirected networks. We varied both the network size and the average node degree, i.e., $n = 16, 128, 1024, 16384$ and $k = 1, 2, 3, 4, 5$. In the WS models, $k$ is even accounting for the number of both left and right neighbors of the nodes in the initial lattice. To obtain small-world networks with $k$ odd, we first generated lattices with $k$ even and then, for each odd node (e.g., 1, 3, ...), we removed the link with its left farthest neighbor. This procedure removes in total $n/2$ links leading to a new average node degree $k' = k - 1$, while keeping a regular structure. As for BA models, we set the number of links in the preferential attachment $m_{ba} = 3$ and the initial seed was a fully connected network of $n_0 = m_{ba}$ nodes. This setting generated scale-free networks with $k = 6 - 12/n$, that is $k \geq 5$ regardless of the selected network size. We then removed at random the exceeding number of links until we reached the desired $k$ value. This procedure had the advantage to preserve the original scale-free structure.

In the second simulation, we considered directed networks to confirm and extend the results we obtained for undirected WS and BA networks. We selected eight representative network sizes, i.e., $n = 8, 16, 32, 64, 128, 256, 512, 1024$ covering the typical size of most current imaging connectomes, and we varied the connection density. Specifically, we performed a two-step procedure:

1. We fixed one-hundred $\rho$ values quadratically spaced within the entire available density interval.

2. After having identified the optimal $\rho^*$, we performed a refined research among one-hundred new values linearly spaced between the density values, in step 1, before and after $\rho^*$.

For WS models, initial lattices had $k$ equal to the nearest even integer equal or higher than $\rho(n - 1)$, with $\rho \in (0, 1)$. For BA models, the number of attaching links was $m_{ba} = \log_2 n$ to ensure an initial relatively high density; the seed was a fully connected network of $n_0 = m_{ba}$ nodes. By construction $\rho \in \left(0, \frac{2m_{ba}n + m_0}{n(n - 1)}\right)$, where $m_0 = n_0(n_0 - 1)/2$ is total number of links in the initial seed. For both models, we then removed at random the exceeding links until we reached the desired density value. For both simulation we generated one-hundred sample networks.

3. Graph analysis of brain networks

Complex networks can be analyzed by a plethora of graph quantities characterizing different topological properties. \(^{25}\) Here, we considered a subset of representative ones which have been shown to be relevant
for brain network analysis. To characterize the entire brain network (i.e., large-scale topology), we used global- and local-efficiency, which respectively read:

\[
E_g = \frac{2}{n(n-1)} \sum_{i\neq j}^{n} \frac{1}{d_{ij}}
\]

\[
E_l = \frac{1}{n} \sum_{i=1}^{n} E_g(i)
\]

(2)

where \(d_{ij}\) is the length of the shortest path between nodes \(i\) and \(j\), and \(E_g(i)\) is the global-efficiency of the \(i\)th subgraph of the network.\(^{12}\)

To characterize modules, or clusters, of brain regions (i.e., mid-scale topology), we extracted the community partition \(\mathcal{P}\) of the brain network by means of the Newman’s spectral algorithm maximizing the modularity:

\[
Q = \frac{1}{2m} Tr(G^T MG)
\]

(3)

where \(G\) is the (non-square) matrix having elements \(G_{ig} = 1\) if node \(i\) belongs to cluster \(g\) and zero otherwise, and \(M\) is the so-called modularity matrix.\(^{27}\)

To characterize individual brain areas (i.e., small-scale topology), we measured the centrality of the nodes in the brain network by means of the node degree and of the node betweenness, which respectively read:

\[
k_i = \sum_{j \neq i}^{n} A_{ij}
\]

\[
b_i = \sum_{j \neq i \neq h} \frac{\sigma_{jh}(i)}{\sigma_{jh}}
\]

(4)

where the element of the adjacency matrix \(A_{ij} = 1\) if there is a link between node \(i\) and \(j\), zero otherwise; and where \(\sigma_{jh}\) is the total number of shortest paths between nodes \(j\) and \(h\), while \(\sigma_{jh}(i)\) is the number of those paths that pass through \(i\).

4. Distances between samples and statistical analysis

To assess brain network differences between individuals (or samples) in the two groups, we measured the distance between the respective values obtained for each graph quantity. We used the Mirkin index to compute distances between two network partitions \(\mathcal{P}_u\) and \(\mathcal{P}_v:\)

\[
MI(\mathcal{P}_u, \mathcal{P}_v) = 2(n_{uv} + n_{vu})
\]

(5)
where \( n_{uv} \) is the number of pairs of nodes that are in different clusters under \( P_u \) but not in the same one under \( P_v \); and \( n_{vu} \) is the number of pairs that are in the same cluster under \( P_u \) but in different ones under \( P_v \). For all other graph quantities, we used the divergent coefficient:\(^{29}\)

\[
D(X_u, X_v) = \sqrt{\left( \frac{1}{M} \sum_{m=1}^{M} x_{u,m} - x_{v,m} \right)^2}
\]

(6)

where \( X_u = [x_{u,1}, x_{u,2}, \ldots, x_{u,M}] \) and \( X_v = [x_{v,1}, x_{v,2}, \ldots, x_{v,M}] \), contain the value(s) of the graph quantity for the \( u \)th and \( v \)th sample. Notably, \( M = 1 \) for global-, local-efficiency and modularity (i.e., \( E_g, E_l, Q \)). \( M = n \) for the node degree vector \( K = [k_1, k_2, \ldots, k_n] \) and the node betweenness vector \( B = [b_1, b_2, \ldots, b_n] \).

We used Kruskal–Wallis one-way analysis of variance to evaluate the overall effect of different thresholds, or filtering methods (i.e., MST, PMFG) on distances between individuals. A Tukey-Kramer multiple comparison test was then used to determine specific differences between pairs of thresholds or methods.\(^{30}\)

Acknowledgements

We thank all the colleagues who shared their databases (see Table 1) allowing to validate the proposed method. We are grateful to Sophie Achard, Vincenzo Nicosia, Jonas Richiardi, and Miguel Valencia for their useful comments and suggestions. V.L. and M.C. acknowledge support by the European Commission Project LASAGNE Grant 318132; V.L. acknowledges support from EPSRC project GALE Grant EP/K020633/1; F.D. and M.C. acknowledge support by French program "Investissements d’avenir” ANR-10-IAIHU-06; F.D. acknowledges support by the project NETBCI Grant ANR-15-NEUC-0006-02.

Author contributions statement

F.D. conceived the study and performed data analysis; V.L. and M.C. contributed to the analytical proofs and simulations. All authors wrote and reviewed the manuscript.

Competing financial interests

The authors declare no competing financial interest.

Materials and correspondence

Specific inquiries on experimental data can be addressed to the corresponding authors listed in Table 1.

References

1. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience 10, 186–198 (2009).
2. Park, H.-J. & Friston, K. Structural and Functional Brain Networks: From Connections to Cognition. *Science* **342**, 1238411 (2013).

3. Stam, C. J. Modern network science of neurological disorders. *Nature Reviews. Neuroscience* **15**, 683–695 (2014).

4. Craddock, R. C. et al. Imaging human connectomes at the macroscale. *Nature Methods* **524** (2013).

5. Alivisatos, A. P. et al. The Brain Activity Map Project and the Challenge of Functional Connectomics. *Neuron* **74**, 970–974 (2012).

6. De Vico Fallani, F., Richiardi, J., Chavez, M. & Achard, S. Graph analysis of functional brain networks: practical issues in translational neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**, 20130521 (2014).

7. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics. *Physics Reports* **424**, 175–308 (2006).

8. Garrison, K. A., Scheinost, D., Finn, E. S., Shen, X. & Constable, R. T. The (in)stability of functional brain network measures across thresholds. *NeuroImage* **118**, 651–661 (2015).

9. Sporns, O. Making sense of brain network data. *Nature Methods* **10**, 491–493 (2013).

10. Latora, V. & Marchiori, M. Economic small-world behavior in weighted networks. *The European Physical Journal B - Condensed Matter* **32**, 249–263 (2003).

11. Bullmore, E. & Sporns, O. The economy of brain network organization. *Nature reviews. Neuroscience* **13**, 336–349 (2012).

12. Latora, V. & Marchiori, M. Efficient behavior of small-world networks. *Physical Review Letters* **87**, 198701/1–198701/4 (2001).

13. Bassett, D. S. & Bullmore, E. Small-world brain networks. *The Neuroscientist* **12**, 512–523 (2006).

14. Watts, D. & Strogatz, S. Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442 (1998).

15. Barabási, A.-L. & Albert, R. Emergence of Scaling in Random Networks. *Science* **286**, 509–512 (1999).

16. Tumminello, M., Aste, T., Matteo, T. D. & Mantegna, R. N. A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences* **102**, 10421–10426 (2005).

17. Fornito, A., Zalesky, A. & Breakspear, M. Graph analysis of the human connectome: Promise, progress, and pitfalls. *NeuroImage* **80**, 426–444 (2013).

18. Tewarie, P., van Dellen, E., Hillebrand, A. & Stam, C. J. The minimum spanning tree: An unbiased method for brain network analysis. *NeuroImage* **104**, 177–188 (2015).
19. Tononi, G., Sporns, O. & Edelman, G. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences* **91**, 5033–5037 (1994).

20. Leclerc, R. D. Survival of the sparsest: robust gene networks are parsimonious. *Molecular Systems Biology* **4**, 213 (2008).

21. Achard, S., Salvador, R., Whitcher, B., Suckling, J. & Bullmore, E. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *Journal of Neuroscience* **26**, 63–72 (2006).

22. van Wijk, B. C. M., Stam, C. J. & Daffertshofer, A. Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory. *PLoS ONE* **5**, e13701 (2010).

23. Vidal, M., Cusick, M. & Barabási, A.-L. Interactome Networks and Human Disease. *Cell* **144**, 986–998 (2011).

24. Souza, F. S. H., Mateus, G. R. & Cunha, A. S. d. Optimization in Designing Complex Communication Networks. In Thai, M. T. & Pardalos, P. M. (eds.) *Handbook of Optimization in Complex Networks*, no. 57 in Springer Optimization and Its Applications, 3–37 (Springer US, 2012).

25. Costa, L. et al. Analyzing and modeling real-world phenomena with complex networks: A survey of applications. *Advances in Physics* **60**, 329–412 (2011).

26. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* **52**, 1059–1069 (2010).

27. Newman, M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* **103**, 8577–8582 (2006).

28. Meilă, M. Comparing clusterings—an information based distance. *Journal of Multivariate Analysis* **98**, 873–895 (2007).

29. Cox, M. A. A. & Cox, T. F. Multidimensional Scaling. In *Handbook of Data Visualization*, Springer Handbooks Comp.Statistics, 315–347 (Springer Berlin Heidelberg, 2008).

30. McDonald, J. *Handbook of Biological Statistics* (Sparky House Publishing, 2009).

31. De Vico Fallani, F., Corazzol, M., M., Sternberg, J., Wyart, C. & Chavez, M. Hierarchy of Neural Organization in the Embryonic Spinal Cord: Granger-Causality Graph Analysis of Calcium Imaging Data. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **PP**, 1–1 (2014).

32. Plomp, G., Quairiaux, C., Michel, C. M. & Astolfi, L. The physiological plausibility of time-varying Granger-causal modeling: Normalization and weighting by spectral power. *NeuroImage* **97**, 206–216 (2014).

33. Teller, S. et al. Emergence of Assortative Mixing between Clusters of Cultured Neurons. *PLoS Computational Biology* **10**, e1003796 (2014).
34. Niu, H. et al. Test-Retest Reliability of Graph Metrics in Functional Brain Networks: A Resting-State fNIRS Study. *PLoS ONE* **8**, e72425 (2013).

35. O’Reilly, J. X. et al. Causal effect of disconnection lesions on interhemispheric functional connectivity in rhesus monkeys. *Proceedings of the National Academy of Sciences* **110**, 13982–13987 (2013).

36. De Vico Fallani, F. et al. Multiscale topological properties of functional brain networks during motor imagery after stroke. *NeuroImage* **83**, 438–449 (2013).

37. Achard, S. et al. Hubs of brain functional networks are radically reorganized in comatose patients. *Proceedings of the National Academy of Sciences* **109**, 20608–20613 (2012).

38. Chavez, M., Valencia, M., Navarro, V., Latora, V. & Martinerie, J. Functional Modularity of Background Activities in Normal and Epileptic Brain Networks. *Physical Review Letters* **104**, 118701 (2010).

39. Besson, P. et al. Structural connectivity differences in left and right temporal lobe epilepsy. *NeuroImage* **100**, 135–144 (2014).

40. De Vico Fallani, F. et al. Community structure in large-scale cortical networks during motor acts. *Chaos, Solitons & Fractals* **45**, 603–610 (2012).
Figures

Figure 1. Optimal connection density in synthetic networks and brain networks. (a-b) Blue curves show the trends of the optimal density $\rho$ for one-hundred generated small-world $p_{ws} = 0.1$ and scale-free $m_{ba} = 9$ networks along different sizes $n$. Blue squares spot out the average optimal $\rho$ values. The black line shows the fit $\rho = c / (n - 1)$ to the data, with $c = 3.258$ for small-world networks and $c = 3.215$ for scale-free networks (Supplementary Table 1). The background color codes for the average value of the quality function $J$. Insets indicate that the optimal average node degree converges to $k = 3$ for large network sizes ($n = 16834$). (c) Optimal density values obtained from group-averaged $J$ profiles for different brain networks. Imaging connectomes come from previously published studies (Table 1). The fit $\rho = c / (n - 1)$ to the pooled data gives $c = 3.06$ (adjusted $R^2 = 0.994$). The inset shows a sharp distribution for the optimal average node degree, with a mode $k = 3$. (d-e) Average $J$ profile (black curves) for simulated small-world and scale-free networks as a function of the network size ($n$) and of the density ($\rho$). $J$ values are represented in normalized units (n.u.), having scaled them by the global maximum obtained for $n = 1024$. Blue and red curves show respectively the profiles of global- ($E_g$) and local-efficiency ($E_l$). (f) Group-averaged $J$ profile for fMRI connectomes (Table 1). The grey dashed line indicates the actual density maximizing $J$, i.e., $\rho = 0.035$, corresponding to an average node degree $k = 3.115$. The graph illustrates the brain network of a representative healthy subject (lateral view, frontal lobe on the left $L_x$).
Figure 2. Statistical comparison of brain network distances across different thresholds. (a-c) Top panels show group-averaged connectomes filtered with ECO for the healthy (blue links) and diseased (red links) group, in three representative imaging modalities, i.e., EEG, fMRI, and DTI (Table 1). Lower panels show distances between brain network properties across different thresholds for global-efficiency $E_g$, community partition $P$, and node degree vector $K = [k_1, ..., k_n]$. Thresholds are given by the average node degree $k$, which corresponds to a connection density $\rho = k/(n-1)$. Circles correspond to medians; horizontal grey lines correspond to lower and upper quartiles; bar colors shade after quartiles. Overall, the distance significantly depends on the threshold value (Kruskalwallis test, $P < 0.001$; Supplementary Table 2). Grey circles represent distances for the optimal threshold $k = 3$. White circles denote threshold values for which distances are not significantly different from $k = 3$ (Tukey-Kramer post-hoc test, $P \geq 0.001$). Transparent circles denote threshold values for which distances are significantly lower than $k = 3$ (Tukey-Kramer post-hoc test, $P < 0.001$). Insets show the $P$-values resulting from the Tukey-Kramer post-hoc comparison of distances between all the threshold values.
Figure 3. Statistical comparison of brain network distances across different filtering methods. Bar plots show the medians of distance between brain network properties of samples in the healthy and diseased group. Vertical bars denote lower and upper quartiles. Medians and quartiles are normalized for the sake of representation. Overall, the choice of the filtering method significantly affects distances between samples (Kruskal-Wallis test, \( P < 0.001 \); Supplementary Table 3). For all graph quantities, ECO gives significantly larger distances as compared to MST (Tukey-Kramer post hoc test, \( P < 0.01 \)) and, with minor extent, to PMFG (Supplementary Table 4). By construction, MST gives null distances for local-efficiency as there are no triangles in tree-like networks.
| Imaging modality | Group(s)     | Species       | Samples x Group | Condition          | Nodes   | Connectivity method                      | Domain               | Links      |
|------------------|--------------|---------------|-----------------|--------------------|---------|------------------------------------------|----------------------|------------|
| Ca⁺S             | Healthy      | Zebrafish     | 5               | Spontaneous        | [9,21]  | Granger causality                        | Time                 | Directed   |
| eEEG             | Healthy      | Rodent        | 1               | Evoked potential   | 15      | Partial directed coherence               | Time/Freq. (8 ms/14-29 Hz) | Directed   |
| Ca⁺M             | Culture      | 2             | Spontaneous     | [19,32]            | Time delay | Pearson’s correlation                   | Time                 | Directed   |
| fNIRS            | Healthy      | Human         | 2               | Resting state      | 46      | Pearson’s correlation                    | Time                 | Undirected |
| fMRS             | Healthy      | Primate       | 3               | Resting state      | 56      | Imaginary coherence                      | Frequency (14-29 Hz) | Undirected |
| EEG              | Healthy      | Human         | 20              | Motor imagery      | 61      | Wavelet correlation                      | Time                 | Undirected |
| fMRI             | Healthy,     | Human         | 17              | Resting state      | 90      | Spectral coherence                       | Frequency (5-14 Hz)  | Undirected |
| MEG              | Healthy,     | Human         | 5               | Resting state      | 149     | Fractional anisotropy                    | -                    | Undirected |
| DTI              | Healthy,     | Human         | 19              | -                  | 164     | Wavelet correlation                      | Time                 | Undirected |
| fMRI L           | Healthy,     | Human         | 17              | Resting state      | 417     | Time delay                               | Time                 | Directed   |
| Ca⁺L             | Culture      | 6             | Spontaneous     | [562,1107]         | Time delay | Imaginary Coherence                     | Frequency (13-30 Hz) | Undirected |
| EEG L            | Healthy      | Human         | 5               | Motor execution    | 4094    | -                                        | -                    | Undirected |

**Table 1.** Experimental details and network characteristics of imaging connectomes.
Supplementary Figure 1. From imaging connectomes to brain networks. As a result of measurements, a raw imaging connectome is mathematically described by full and weighted connectivity matrices $W$. To obtain a sparse brain network, the raw information is filtered and binarized by applying a threshold either on the weights (i.e., the connectivity strength) or on the percentage (i.e., the connection density) of strongest weights to retain in the adjacency matrix $A$. 
Supplementary Figure 2. Optimal connection density in synthetic networks and brain networks. (a-b) Blue curves show the trends of the optimal density $\rho$ for one-hundred generated lattices and random networks along different sizes $n$. Blue squares spot out the average optimal $\rho$ values. The black line shows the fit $\rho = c/(n-1)$ to the data, with $c = 3.265$ for lattices and $c = 2.966$ for random networks (Supplementary Table 1). The background color codes for the average value of the quality function $J$. Insets indicate that the optimal average node degree converges to $k = 3$ for large network sizes ($n = 16834$). (c) Optimal density values obtained from individual $J$ profiles for different brain networks. Imaging connectomes come from previously published studies (Table 1). A larger variability can be observed with respect to the values we obtained when considering group-averaged $J$ profiles Fig. 1c. The inset confirms a more uniform distribution for the optimal average node. Nevertheless, the fit $\rho = c/(n-1)$ to the pooled data gives $c = 2.87$ (adjusted $R^2 = 0.946$), which is in practice very close to $k = 3$. (d-e) Average $J$ profile (black curves) for simulated lattices and random networks as a function of the network size $(n)$ and of the density $(\rho)$. $J$ values are represented in normalized units (n.u.), having scaled them by the global maximum obtained for $n = 1024$. Blue and red curves show respectively the profiles of the global- ($E_g$) and local-efficiency ($E_l$). (f) Individual $J$ profile for a representative fMRI connectome (Table 1). The grey dashed line indicates the actual density maximizing $J$, i.e., $\rho = 0.008$, corresponding to an average node degree $k = 0.712$. This value was very far from the expected optimal $k = 3$. Indeed, we noticed that for very low density values the intrinsic brain network structure could not completely emerge and spurious peaks could appear. To improve the quality of the estimate, we averaged the $J$ profiles over the samples in the same group (Fig. 1). The graph illustrates the brain network of a representative healthy subject (lateral view, frontal lobe on the left $L_x$).
Supplementary Figure 3. Statistical comparison of brain network distances across different thresholds for local-efficiency $E_l$, modularity $Q$, and node betweenness vector $B = [b_1, ..., b_n]$. Panel (a) show distances for EEG connectomes, (b)-fMRI, (c)-DTI, (d)-MEG. Thresholds are given by the average node degree $k$, which corresponds to a connection density $\rho = k/(n - 1)$. Circles correspond to medians; horizontal grey lines correspond to lower and upper quartiles; bar colors shade after quartiles. Overall, the distance significantly depends on the threshold value (Kruskal-Wallis test, $P < 10^{-54}$; Supplementary Table 2). Grey circles represent distances for the optimal threshold $k = 3$. White circles denote threshold values for which distances are not significantly different from $k = 3$ (Tukey-Kramer post-hoc test, $P \geq 0.001$). Transparent circles denote threshold values for which distances are significantly lower than $k = 3$ (Tukey-Kramer post-hoc test, $P < 0.001$). Insets show the $P$-values resulting from the Tukey-Kramer post-hoc comparison of distances between all the threshold values.
Supplementary Figure 4. Statistical comparison of brain network distances across thresholds and filtering methods for MEG connectomes. (a) The top-left panel shows the group-averaged MEG connectomes filtered with ECO for the healthy (blue links) and diseased (red links) (Tab. 1). Other panels show brain network distances across different thresholds for global-efficiency $E_g$, community partition $P$, and node degree vector $K = [k_1, ..., k_n]$ (Online Method 3). Thresholds are given by the average node degree $k$, which corresponds to a connection density $\rho = k/(n-1)$. Circles correspond to medians; horizontal grey lines correspond to lower and upper quartiles; bar colors shade after quartiles. Overall, the distance significantly depends on the threshold value (Kruskal wallis test, $P < 10^{-15}$; Supplementary Table 3). Grey circles represent distances for the optimal threshold $k = 3$. White circles denote threshold values for which distances are not significantly different from $k = 3$ (Tukey-Kramer post-hoc test, $P \geq 0.001$). Transparent circles denote threshold values for which distances are significantly lower than $k = 3$ (Tukey-Kramer post-hoc test, $P < 0.001$). Insets show the $P$-values resulting from the Tukey-Kramer post-hoc comparison of distances between all the threshold values. (b) Medians of the distances between brain network properties of samples in the healthy and diseased group for MEG connectomes. Vertical bars denote lower and upper quartiles. Medians and quartiles are normalized for the sake of representation. For all graph quantities ECO gives significantly larger distances as compared to MST (Tukey-Kramer post hoc, $P < 0.05$) and, with minor extent, to PMFG (Supplementary Table 4). Notably, MST gives null distances for local-efficiency $E_l$ as there are no triangles in tree-like networks.
Supplementary Figure 5. Optimal density obtained with parametric quality functions in syntethic networks. Background colors code for the average $J$ values in the logarithmic scale. Blue circles spot out the maximum as a function of the parameter $\beta$ (y-axis). The black line shows the reference optimal density value $\rho = \frac{3}{(n-1)}$ (x-axis). (a) $J' = \frac{E_g + E_l}{\beta \rho}$; (b) $J'' = \frac{E_g + E_l}{\rho^\beta}$.
**Supplementary Tables**

|       | Lattice | Small-world | Random | Scale-free |
|-------|---------|-------------|--------|------------|
| $c$   | 3.265   | 3.258       | 2.966  | 3.215      |
| $R^2$ | 0.999   | 0.997       | 0.993  | 0.998      |

**Supplementary Table 1.** Statistics of data fitting $\rho = c/(n - 1)$ for synthetic networks. The fit’s constant $c$ and the adjusted $R^2$ coefficient are reported along different network models.
### Supplementary Table 2.

$P$-values from Kruskalwallis tests on brain network distances across different thresholds. Three asterisks denote significant differences with a statistical percentile $P < 0.001$; one asterisk corresponds to $P < 0.05$.

|      | EEG     | fMRI    | DTI     | MEG     |
|------|---------|---------|---------|---------|
| $E_g$| 2.98E-36*** | 3.74E-57*** | 9.53E-41*** | 2.82E-2* |
| $E_l$| 1.66E-113*** | 6.60E-83*** | 4.65E-82*** | 3.99E-4*** |
| $P$  | 1.09E-185*** | 5.03E-139*** | 1.09E-192*** | 8.10E-12*** |
| $Q$  | 3.51E-106*** | 1.94E-130*** | 1.07E-62*** | 1.15E-10*** |
| $k_i$| 9.57E-214*** | 6.70E-151*** | 1.21E-203*** | 8.25E-14*** |
| $b_i$| 7.19E-162*** | 3.49E-125*** | 6.35E-184*** | 3.38E-11*** |
|       | EEG          | fMRI         | DTI          | MEG          |
|-------|--------------|--------------|--------------|--------------|
| $E_g$ | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $4.905E-43^{***}$ |
| $E_i$ | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $3.521E-54^{***}$ |
| $P$   | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $5.196E-61^{***}$ |
| $Q$   | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $7.722E-59^{***}$ |
| $k_i$ | $3.143E-21^{***}$ | $5.240E-20^{***}$ | $1.948E-18^{***}$ | $1.523E-15^{***}$ |
| $b_i$ | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $< 1E-255^{***}$ | $4.069E-63^{***}$ |

**Supplementary Table 3.** $P$-values from Kruskalwallis tests on brain network distances across different filtering methods. Three asterisks denote significant differences with a statistical percentile $P < 0.001$. 
|        | EEG     | fMRI    | DTI     | MEG     |
|--------|---------|---------|---------|---------|
| E_{st} | 3.041E-03** | 9.561E-10*** | 9.561E-10*** | 9.561E-10*** |
| E_{t}  | -       | 1.060E-10*** | -       | 1.060E-10*** |
| T      | 9.561E-10*** | 9.561E-10*** | 9.561E-10*** | 9.561E-10*** |
| Q      | 9.561E-10*** | 9.561E-10*** | 9.561E-10*** | 9.561E-10*** |
| k_{t}  | 9.561E-10*** | 9.561E-10*** | 9.561E-10*** | 9.561E-10*** |
| b_{t}  | 9.561E-10*** | 1.886E-01 | 9.561E-10*** | 6.936E-03** |

Supplementary Table 4. P-values from Tukey-Kramer post-hoc tests on brain network distances across different filtering methods. Three asterisks denote significant differences with a statistical percentile \( P < 0.001 \); two asterisks correspond to \( P < 0.01 \), while one asterisk corresponds to \( P < 0.05 \). For local-efficiency (\( E_t \)), tests were not performed when comparing ECO vs MST, as by construction, MST gives null \( E_t \) values.
Supplementary Text

1. Optimal density in regular lattices and random networks

Global- $E_g$ and local-efficiency $E_l$ can be well approximated, respectively, by the inverse of the characteristic path length $L$ and by the clustering coefficient $C$. We can then rewrite the quality function in Eq. (1) as $J = (L^{-1} + C)/\rho$. In some simple cases, $L$ and $C$ can be expressed as a function of the average node degree $k$ and of the number of nodes $n$ of a network.

In particular, for large 1D regular lattices, corresponding to Watts-Strogatz (WS) model with a rewiring probability $p_{ws} = 0$, we have $L_0 = n/2k$ and $C_0 = 3(k-2)/(4(k-1))$; while for large random networks ($p_{ws} = 1$), we have $L_1 = \log n / \log k$ and $C_1 = k/n$. Since the relation $\rho = k/(n-1)$ holds for any network, we can rewrite the quality function $J$ for regular lattices as:

$$J_0 = 2 \frac{(n-1)}{n} + \frac{3(n-1)\rho - 6}{4(n-1)^2 - 4\rho}$$

and for random networks:

$$J_1 = \frac{\log [(n-1)\rho]}{\rho \log n} + \frac{(n-1)}{n}$$

Notice that the leading contribution to $J$ is coming from $C$ in the case of a lattice and from $L^{-1}$ in the case of a random network.

For these synthetic networks we can determine analytically the optimal value of $\rho$ that maximizes the quality function $J$. In particular, by deriving with respect to $\rho$ and equating to zero, we obtained for lattices the following solution which satisfies the requirement of connectedness ($\rho \geq 2/n$):

$$\rho_0 = \frac{2 + \sqrt{2}}{n-1}$$

As for random networks we obtained:

$$\rho_1 = \frac{e}{n-1}$$

It follows from Eq. (3) and Eq. (4) that in both cases the optimal density can be written as $k/(n-1)$, where the constant (i.e., the average node degree) is approximately equal to 3. Without loss of generality we referred here to undirected networks. For directed networks the same results hold, $k$ corresponding to the average node in/out-degree.
2. Optimal density in alternative quality functions

We considered alternative quality functions combining $E_g$, $E_l$ and $\rho$. For each new $J$, we checked analytically the existence of an optimal density in both regular lattices and random networks. We excluded the quality functions for which we obtained indefinite, non-closed or trivial solutions.

- $J = E_g + E_l - \rho = 1/L + C - \rho$

  By substituting the expression of $L$ and $C$ for random networks we obtain:

  $$J_1 = \frac{\log k}{\log n} + \frac{k}{n} - \frac{k}{n-1}$$

  (5)

  For the sake of simplicity we solved with respect to $k$, knowing that $\rho = k/(n-1)$. When derivating and equating to zero we had $k = \frac{n(n-1)}{\log n}$, which leads to the impossible condition $\rho = n/\log n > 1$.

- $J = E_g - \rho = 1/L - \rho$

  By using the expression of $L$ for lattices we obtain:

  $$J_0 = \rho \left(2 \left(\frac{n-1}{n}\right) - 1\right)$$

  (6)

  When derivating and equating to zero we had $n = 2$ and we cannot solve with respect to $\rho$.

- $J = E_l - \rho = C - \rho$

  By substituting the expression of $C$ for random networks we obtain:

  $$J_1 = \frac{k}{n} - \frac{k}{n-1}$$

  (7)

  When derivating and equating to zero we had $1/n = 1/(n-1)$ and we could not solve with respect to $\rho$.

- $J = E_g E_l / \rho = (C/L)/\rho$

  By substituting the expression of $L$ and $C$ for random networks we obtain:

  $$J_1 = \frac{n-1}{n \log n} \log k$$

  (8)

  When derivating and equating to zero we had the trivial solution $k = 0$.

Similarly, it is easy to prove that neither $J = E_g / \rho$ nor $J = E_l / \rho$ admitted meaningful solutions in lattices or random networks.
3. Optimal density in parametric quality functions

We introduced a parameter $\beta$ to tune the contribution of the connection density in $J$. In particular, we evaluated two quality functions:

\[
J' = \frac{E_g + E_l}{\beta \rho}
\]  

and

\[
J'' = \frac{E_g + E_l}{\rho^\beta}
\]

where we fixed $\beta \in [0, 2]$. In both cases, we obtained the original expression of $J$ when $\beta = 1$. We evaluated the above quality functions over one-hundred synthetic networks generated according to Online method 2. For the sake of simplicity we only considered networks with size $n = 512$. Results show that for $J'$ the parameter $\beta$ has no effect on the optimal density, which is always $\rho = 3/(n-1)$ Supplementary Fig. 5a. As for $J''$, the optimal density can change when we move far from $\beta = 1$ Supplementary Fig. 5b. This is particularly true for both random networks and scale-free networks.