A method for multiscale community detection in brain networks

Lazaro M. Sanchez-Rodriguez[a,*], Yasser Iturria-Medina[b], Pauline Mouches[a,c], Roberto C. Sotero[a,c*]

[a]Department of Radiology and Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada

[b]Montreal Neurological Institute, Montreal, Quebec, Canada

[c]Biomedical Engineering Graduate Program, University of Calgary, Calgary, Alberta, Canada

[*]Corresponding author: Department of Radiology and Hotchkiss Brain Institute, University of Calgary, 3330 Hospital Drive NW Calgary, AB T2N 4N1

E-mail addresses: lazaro.sanchezrodrgu@ucalgary.ca (L.M. Sanchez-Rodriguez) and roberto.soterodiaz@ucalgary.ca (R.C. Sotero)

Abstract

The identification of community structure in graphs continues to attract great interest in several fields. Network neuroscience is particularly concerned with this problem considering the key roles communities play in brain processes and functionality. Most methods used for community detection in brain graphs are based on the maximization of a highly parameter-dependent modularity function. In practice, the parametrization of this function often obscures the physical meaning and hierarchical organization of the partitions of network nodes. In this work, we present a new method able to detect communities at different scales in a natural, unrestricted way. First, to obtain an estimation of the information flow in the network we release random walkers to freely move over it. The activity of the walkers is separated into oscillatory modes by using empirical mode decomposition. After grouping nodes by their co-occurrence at each time scale, the so-called *k*-modes clustering returns the desired partitions. Our algorithm was first tested on benchmark graphs with favorable performance. We used the method on brain networks, including the anatomical connectivity of the macaque and human brains and a model for the interactions between nodes. We found a repertoire of community structures in the anatomical and functional networks, with a clear link existing between these two. The observed partitions range from the evident division in two hemispheres –in which all processes are managed globally– to specialized communities seemingly given by physical proximity and shared function. Our results stimulate the research of hierarchical community organization in terms of temporal scales of information flow.

Highlights

- Oscillatory modes of networks' signals carry information on architectural rules.
- Meaningful partitions of the brain network are found over different temporal scales.
- The multiscale organization of the brain responds to the function of its components.

Keywords

Community detection algorithms

Brain networks

Random walkers

Empirical mode decomposition

k-modes clustering

1. Introduction

Network community detection constitutes a problem of current vital importance. Among all the nodes and interactions constituting a network, structures of subdivisions exist. In each of these communities (also referred as groups or clusters), nodes have a greater probability of being locally connected than to nodes in other groups (Fortunato & Hric, 2016; Garcia, Ashourvan, Muldoon, Vettel, & Bassett, 2018). One example with several applications in the literature (Girvan & Newman, 2002; Porter, Mucha, Newman, & Warmbrand, 2005; Traud, Kelsic, Mucha, & Porter, 2011) is the tight-knit of a person's friendships and the exchanges they have with other groups of friends. The identification of community structures provides insights into organizational principles, not only in terms of isolation of the clusters per se but also for the collective dynamical spreading of processes over the network (Fortunato & Hric, 2016).

In the brain, neural units connect to one another over different spatio-temporal scales in intriguing and fascinating ways (Breakspear, Sporns, Honey, & Ko, 2007; Moradi, Dousty, & Sotero, 2019; Sotero & Trujillo-Barreto, 2008; Valdes-Sosa et al., 2009). The modularity of a such system is believed to critically impact the phenomena of segregation (processes occurring in groups of heavily interconnected brain units) and integration (the combination of information exclusive to specialized brain regions) (Rubinov & Sporns, 2010; Sporns, 2013). Other advantages of a community structure relate to adaptability, robustness to failure and the reduction of wiring costs –see (Garcia et al., 2018) and (Betzel et al., 2017) and the references therein. Additionally, grouping exists across different levels (a hierarchy) for supporting rapid responses to changes (Garcia et al., 2018). As an illustration, consider the large community of neural conglomerates in one cerebral hemisphere. This can be broken into smaller communities according to the functional role of their members (Thomas Yeo et al., 2011). An important initial step for the study of brain

structure, however, is the definition of the nodes and edges in the graph and the scale to be considered. This selection relies on the data available, which depends on the imaging modality used to record it. For example, anatomical associations can be examined through diffusion weighted magnetic resonance imaging data (DWMRI) and functional neuroimaging or electrophysiological methods, e.g., fMRI and electroencephalogram provide insights into the dynamic interactions between brain regions (Y. Iturria-Medina et al., 2007; Valdes-Sosa et al., 2009).

Regardless of the network data, the bulk of community studies in the brain use variants of Newman's modularity function (Newman, 2006) and its maximization through Louvain-like algorithms (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) for the detection of clusters of regions (Sporns, 2013). The partitions obtained through these methods maximize intra-community edge weights relative to a specific random network null model (Bassett et al., 2013; Garcia et al., 2018). Overall, these algorithms are problematic in that the output structure depends on the chosen null model and on a resolution parameter, γ, as well. Exploration of the resolution parameter space yields several structures which occasionally present hierarchy (Bassett, Khambhati, & Grafton, 2017). How could one set γ so that a meaningful set of communities –and not any partition– is revealed? In many instances, researchers exclusively report the partition obtained for $\gamma = 1$ (Fukushima et al., 2018). Nevertheless, it is known that high-modularity partitions can be found for $\gamma = 1$ in random, unstructured graphs, where no community structure should be detected (Fortunato & Hric, 2016). More recently, a useful heuristic has been introduced to retain the so-called graph's most salient partition (Bassett et al., 2013; Garcia et al., 2018). In brief, a grid search is performed on the resolution parameter to find the value that generates the set of partitions with the greatest similarity. However, modularity maximization tends to split large communities in

smaller pieces, which is a consequence of the choice of the null model. This effect is not solved by multi-resolution approaches (Fortunato & Hric, 2016). These techniques have also been adapted to generate hierarchical output structures (Ashourvan, Telesford, Verstynen, Vettel, & Bassett, 2019; Jeub, Sporns, & Fortunato, 2018) though the limitations with regard to the choice of null models and resolution parameters persist.

Other algorithms exist with somewhat fewer applications in brain research (Fortunato & Hric, 2016; Gates, Henry, Steinley, & Fair, 2016). Given the connectivity characteristics of communities, the utilization of random walkers for their identification is fairly straight-forward. Walkers tend to stay trapped in a cluster before transitioning to a different group (Fortunato & Hric, 2016). Walktrap (Pons & Latapy, 2005) and Infomap (Rosvall & Bergstrom, 2008) are examples of detection methods that employ random walk dynamics. The former is a costly, parameter-dependent method that exploits the probability of transition between two nodes in a certain number of steps as a measure of vertex similarity to group nodes. In the latter, a codeword is assigned to each vertex the walker encounters. Infomap considers networks with community structure to be analogous to geographical maps: unique codewords (street names) are only necessary to identify nodes (streets) in one specific community (city). Although Infomap has proven effective in artificial benchmark graphs and large datasets, it has performed more poorly in classical real networks traditionally utilized for testing algorithms (Gates et al., 2016; Hric, Darst, & Fortunato, 2014), e.g. the Zachary karate club (Girvan & Newman, 2002; Zachary, 1977). Those networks, for which ground-truth partitions are known, resemble some commonly analyzed brain graphs in that they have relatively small size and present various types of adjacency matrices, e.g., sparse, like those obtained from DWMRI or dense, from fMRI (Gates et al., 2016).

The description in terms of dynamical flows, as utilized in Walktrap and Infomap despite the above-mentioned limitations, is one that appeals to neuroscientists. In the first place, structural features like node degrees, number of edges of the brain graph, etc., condition the dynamics of network processes (Fortunato & Hric, 2016). Secondly, the transmission of information in the brain is, obviously, a dynamical process (Sotero, Sanchez-Rodriguez, Dousty, Iturria-Medina, & Sanchez-Bornot, 2019), brain connectivity being adaptive and function-sensitive within the context of structural constraints (Friston, 2011). For these reasons, we believe that the analysis of community structure and the identification of hierarchical architectures in brain networks can benefit from taking into account the dynamical aspects of its information flow. Thus, in this paper, we present a novel approach to community detection specifically designed for brain graphs, although not limited to them.

We build on the methodology introduced by Sotero et al. in recent work (Sotero et al., 2019). These authors studied information flow in brain networks by using the fraction of walkers that a given one finds at each node as the variable describing the evolution of the walker over the network. This function of time, taken as the network's signal, can be decomposed into its constituent frequencies by using empirical mode decomposition (EMD) (N. E. Huang et al., 1996). Each of these oscillatory modes then associates with the notion of a temporal scale. Here, we incorporate a final step for performing data partitioning through k-modes clustering (Z. Huang, 1998). The arrangement of the nodes visits recorded throughout the walkers' flows at the different temporal scales allows the unveiling of a hierarchical organization. Intuitively, a walker would spend considerable times in large communities, which is seen in slow oscillatory modes. Analogously, fast modes could reflect the motion over smaller clusters. Initially, we test the algorithm on benchmarks and real networks with known community structure such as Girvan-

Newman (Girvan & Newman, 2002), Lancichinetti-Fortunato-Radicchi (Lancichinetti, Fortunato, & Radicchi, 2008) and the Zachary karate club (Zachary, 1977). We then proceed to extract communities existing in macaque and human anatomical connectivity matrices, as well as *in-silico* functional connectivity graphs built over the human brain anatomical network. Meaningful patterns of communities obtained here support the reliability of our method.

2. **Materials and methods**

2.1 *The network's signal*

Let us imagine a network of n nodes, possibly presenting community structure, in which a random walker is set free. The walker moves over the edges available to it. In general, the probability of transitioning from node j towards node i in the next time step is given by (Zhang, Shan, & Chen, 2013):

$$p_{j\to i} = \frac{c_{ji}}{\sum_{j=1}^{n} c_{ji}} \tag{1}$$

where C_{ji} is the weight of the connection from area j to area i. The walker tends to visit the nodes in a community before a route takes it to an outsider, a member of a different community (Fortunato & Hric, 2016). This is because of the predominantly local connectivity pattern of communities. Now, suppose that $W (W \gg n)$ walkers simultaneously move over the same network. Each time one walker appears in a node, it finds fellow walkers, while others visit different nodes. Let us compute, for each walker, the fraction of the total number of other walkers that it encounters at each time step. After T time steps, there exist W time series reflecting different realizations of the flow of information in the network (Sotero et al., 2019). Those series incorporate information on the structure of the network (e.g., the number of walkers at a hub is expected to be persistently high), and the paths therein existing (i.e., the random walk itself). Finally, for generalization purposes –as the ratio of walkers would depend on the size of the network– we

standardize such time series. Fig. 1a shows an exemplary signal corresponding to one of the walkers flowing over one of the networks considered in this paper. The horizontal axis is two-fold, showing both the temporal iteration (lower) and the indexes of the nodes the selected walker visits at each time step. Given the size of the graphs in this study, i.e., brain networks with $n \sim 10^2$, we fix $W = 1000$.

The representation of the network we come up with can be decomposed to obtain oscillatory modes at different time scales. The temporal scales of random walks processes in complex networks depend on the network structure (Sotero et al., 2019). In other words, the network structure can be seen through different dynamical levels ranging from slow time scales in which walkers practically travel through the entire network to faster scales consisting, for instance, of the abrupt transitions from one node to the next. Empirical mode decomposition (EMD) (N. E. Huang et al., 1996) solves the problem of finding a nearly orthogonal basis for any complicated, nonlinear and non-stationary process without the need for any predefined model. The components into which the signal is broken down representing temporal scales are conventionally called intrinsic mode functions (IMFs). Each of these satisfy that: 1) the number of zero-crossings and extrema of the function are either equal or differ by one, and 2) the mean of its upper and lower envelopes is zero. Components are different as to conveyance of information (Sotero, 2016). For notation purposes, IMF_1 denotes the fastest mode (highest frequency). The rest are named accordingly.

In our interpretation oriented to community detection, IMFs of the fraction of walkers-signal reflect hierarchical organization of the network. For example, fast temporal scales of the fraction of walkers could associate with partitions composed of small groups of nodes. In slower modes, the walker may get to visit all the nodes existing in larger communities. To give an example, as shown in Fig. 1b, a walker may quickly transition over nodes 74, 76, 65, 54 and 60 (seen with

IMF_2) or more slowly appear at those but also at others (IMF_4). The five elements mentioned above may represent a community; those and the ones identified by the green color, may constitute a broader community.

Fig. 1 goes around here

Fig. 1. A typical network's signal. a) Standard fraction of other random walkers one walker finds while flowing in a network. The lower horizontal axis shows the temporal progression of the walker. The upper axis shows the succession of nodes it visits. b) Empirical mode decomposition of the signal in (a). The activity of five nodes that are found between two zero-crossings of IMF_2 is highlighted in blue. Together with those, other nodes appear between zero-crossings of IMF_4 (in green). Only the first 30 time steps of $T = 5000$ are shown for visualization purposes.

2.2 *Finding nodes that cluster together*

The following step consists of exploiting the features of the IMFs and grouping nodes together. For each oscillatory mode and walker, we take chunks of data consisting of the network nodes seen between zero crossings. In our previous example, 74, 76, 65, 54 and 60 would be one of such data chunks for the fast IMF_2 (Fig. 1b). Other sets of nodes will appear in different portions. One may think of our selection of the chunks in terms of the oscillations of a spring-mass system. There, points to the right/left of the zero reference cluster together (the spring is stretched/compressed). Each time the signal for the displacement of the mass passes through the equilibrium position it is also switching from one 'community' to the other. The zero crossings-analogy bases on the interpretation and symmetry properties of the IMFs. In practice, nodes outside a certain true community C_q, may occasionally pertain to a chunk of otherwise genuine members

of C_q, given the existence of edges to that community. Likewise, all nodes belonging to a community do not necessarily have to appear together between two contiguous zero crossings. The problem is how to identify authentic node clusters over the effects of noise with the information available from the IMFs. To address this issue, we turn to unsupervised learning, particularly to clustering. In clustering analysis, the goal is to group objects based on the information available – features describing the data (Assent, 2012; Ronan, Qi, & Naegle, 2016; Steinbach, Ertöz, & Kumar, 2004). The more similar items in a group are and the more different to those in other groups, the better the clustering (Steinbach et al., 2004).

To capture the natural structure of the data, we proceed to cluster network nodes –or objects, in conventional clustering jargon– given their co-appearance between zero-crossings of the IMFs –the features. Features are binary vectors encoding the positions of the nodes appearing together in the between zero-crossings chunks. This yields a complete representation of the IMFs corresponding to each walker in terms of binary variables (Fig. 2a). To ensure a proper sampling of all nodes and their co-appearances, we select the features corresponding to a high number of walkers (200 out of the 1000 simulated). This is a random selection, in the same way that subsets of variables are sometimes chosen when clustering data with multiple independent signals (Jiliang Tang, Salem Alelyani, & Huan Liu, 2014; Ronan et al., 2016).

Here, we employ the so-called *k*-modes clustering algorithm (Z. Huang, 1998). *K*-modes is an extension of the popular *k*-means method (Macqueen, 1967) for categorical variables, binary features being a particular case of those. A simple matching dissimilarity is used as notion of distance in *k*-modes. Two objects, $\boldsymbol{x}$ and $\boldsymbol{y}$ are far from each other by a quantity that equals the number of mismatching features (of M features), namely:

$$d(\boldsymbol{x},\boldsymbol{y}) = \sum_{l=1}^{M} \delta(\boldsymbol{x}_l,\boldsymbol{y}_l),\ \delta(\boldsymbol{x}_l,\boldsymbol{y}_l) = \begin{cases} 0 & (\boldsymbol{x}_l = \boldsymbol{y}_l) \\ 1 & (\boldsymbol{x}_l \neq \boldsymbol{y}_l) \end{cases} \tag{2}$$

The algorithm minimizes a cost function after a vector (the mode[1]) for each of the *k* clusters has been selected and objects grouped around it such that their dissimilarity is minimal. Alike *k*-means, Huang's *k*-modes yields locally optimal solutions depending on the starting conditions. Thus, one necessary first step is running the algorithm several times to select the solution with the lowest overall cost. The number of initial conditions for the clustering algorithm is empirically set to 50 in this work, based on the consistency of the solutions obtained. With all these considerations *k*-modes is run (Fig. 2b). The implementation of *k*-modes we used is available from https://github.com/nicodv/kmodes.

2.3 *Accepting/rejecting hierarchical partitions*

Clustering algorithms generally find clusters even in absence of underlying structure, highlighting the necessity of validating solutions (Ronan et al., 2016). In *k*-modes, objects are allocated in *k* clusters exactly. However, unless the user has prior knowledge on the distribution of the data –which rarely occurs– *k* is a parameter to be determined. Several metrics are intended to elucidate the correct number of clusters in the data from running the algorithm over a range of *k*. These metrics use measures of separation, compactness, or both. Various studies, most famously the one by Milligan and Cooper (Milligan & Cooper, 1985), have looked at the performance of indexes for assessing the results over numerical data and Euclidean distances. To the best of our knowledge, such measures have not been transformed to account for categorical data (binary, in particular) and matching dissimilarity, as recommended by Huang in his seminal paper (Z. Huang, 1998). Therefore, in Appendix B, we briefly describe six measures with satisfactory performance for recovering true cluster structure, i.e.: the Calinski-Harabasz index (Anderson, 2001; Calinski

[1]The terms employed here are mostly faithful to the ones used in each of the parts combined. This is why the word mode appears with two meanings: one refers to the oscillatory functions in which a signal can be decomposed (the intrinsic mode functions), the other represents 'the centroids' of the clusters of nodes obtained with binary features (as in *k*-modes clustering).

& Harabasz, 1974), the C-index (Milligan & Cooper, 1985), a modified Duda-Hart criterion (Duda, Hart, & Stork, 2001), silhouette width (Kaufman & Rousseeuw, 1990), one of the family of Dunn indexes (Bezdek & Pal, 1998; Dunn, 1973) and the Davies-Bouldin index (Davies & Bouldin, 1979; Dubes, 1987). Details on their utilization and customization to account for binary data, if applicable, are also included. Importantly, as the success ratio of an individual index in determining the true number of clusters is limited and may depend on the data (Dubes, 1987; Milligan & Cooper, 1985), here, we adopt the criterion of selecting the best number of clusters based on a majority rule (Charrad, Ghazzali, Boiteau, & Niknafs, 2014). Failure to establish a majority of the indexes indicating the same correct number of clusters may hint at the lack of a definitive community pattern. This is usually accompanied by at least one of the indexes rejecting the existence of community structure altogether (see Appendix C). Fig. 2c shows an example in which the majority of the validation measures indicate the existence of $k = 4$ clusters, to finalize the illustration of the basic pipeline of our method. For the networks considered in this paper, distributions of nodes in up to 20 clusters only were investigated.

Fig. 2 goes around here

Fig. 2. Clustering network nodes. a) For each IMF, 'features' are constructed so that nodes appearing together between zero-crossings of the IMF are assigned a logical 1. For example, Feature 2 here corresponds to the highlighted nodes for IMF_4 of Fig. 1b. Such description is extended through all zero-crossings of one IMF and to other walkers to guarantee a proper sampling of the network nodes and their close acquaintances in the temporal scale. b) The obtained data matrix feeds a *k*-modes clustering algorithm. For one instance of the data and $k = 4$ clusters, the algorithm returns the solution shown (e.g. nodes 76, 77 and 78 are in "community 3"). c) After

exploring the k-space, all the solutions are considered according to several validation indexes. In the example, the Calinski-Harabasz index and the silhouette width, both presenting an absolute maximum, together with C-index and the Duda-Hart measure (having absolute minimums), suggest the existence of 4 clusters in the data. This is sufficient to come to a conclusion although the resting indexes (Dunn and Davies-Bouldin) present relative maximum and minimum at $k = 4$, respectively, which according to their definition may as well indicate the presence of network organization in 4 communities. Values of k for which k-modes yields singleton communities are not shown ($k = [13, 20], k \in \mathbb{N}$).

The process illustrated in this figure should be performed for many combinations of walkers (to build consensus partitions) and for all the IMFs (to unveil hierarchical organization).

The exploration over values of k should be performed for each of the IMFs. The following intuitive rule of thumb is followed. First, take the slowest IMF in the decomposition of the signal corresponding to the fraction of walkers. After running the clustering algorithm, determine the number of communities existing. Move on to the next IMF and determine its corresponding partition and the number of clusters present. If this number is equal to any obtained beforehand, reject the previous partition and accept only the current one. Else, retain both partitions (with different numbers of clusters). Proceed with the analysis until all the IMFs have been considered. This way, a hierarchical organization is unveiled.

One last issue to consider is the stochasticity that is inherent to most community detection techniques (Bassett et al., 2013; Fortunato & Hric, 2016). In our method, this is expressed as sets of randomly chosen walkers which are considered to build clustering features. The application of the clustering algorithm over those subsets of features can yield slightly different partitions. To

present unique partitions, we use consensus clustering (Strehl & Ghosh, 2003). One starts by building a consensus matrix, $\boldsymbol{T}$, that accounts for the co-occurrence of nodes in communities. Non-significant relationships between nodes are removed by thresholding the co-occurrence matrix. Such threshold is set to the highest value of all co-occurrences in the association matrix resulting from random permutations of the original partitions (Bassett et al., 2013). Then, the algorithm is applied over $\boldsymbol{T}$ until all the partitions are identical (the 'true-partition'). The results reported in this paper correspond to the consensus partition after applying *k*-modes clustering to 50 sets of independent features for each of the networks. Their similarity with known ground-truth communities is analyzed by using adjusted mutual information, AMI (N. Vinh, Epps, & Bailey, 2010) (see Appendix D).

2.4 *Data description and processing*

2.4.1 *Benchmarks*

Artificially generated graphs and a real network with known group structure were used to assess the performance of our algorithm. We chose simple benchmark graphs with features alike the brain networks for which the application was intended, e.g. similar number of nodes.

2.4.1.1 *Girvan-Newman benchmarks (GN)*

These graphs are random with known community structure. A GN benchmark consists of 128 nodes and four communities, with 32 nodes each. The average expected degree of a node is 16 (Girvan & Newman, 2002). A fraction of those connections (μ) is made to vertices in other communities. As such fraction increases, algorithms usually struggle to pinpoint the underlying community structure. Here, we set $\mu = 0.1$ (Lancichinetti & Fortunato, 2009) and applied our detection algorithm over both binary and weighted versions of the GN model. One limitation of GN is its inability to reproduce the scale-free property of real networks (heterogeneous node degree

distributions, node degree and community sizes following a power-law distribution) (Lancichinetti et al., 2008).

2.4.1.2 *Lancichinetti-Fortunato-Radicchi benchmarks (LFR)*

A more realistic benchmark, LFR does account for the heterogeneous and skewed distribution of the degree and community size. Both these parameters are chosen from power-law distributions. Networks are built by joining stubs at random (Fortunato & Hric, 2016; Lancichinetti et al., 2008). We kept the value of the mixing parameter at 0.1. The average degree in a network of $n = 100$ nodes was set to 13 and the upper extreme of the degree distribution to 27. Consequently, the randomly generated networks (binary and weighted) utilized here had 5 communities, with sizes [26, 23, 21, 20, 10].

The GN and LFR networks used in this work were generated by using code available from (https://sites.google.com/site/santofortunato/inthepress2). All the code parameters were set to their default values except for the ones above-mentioned.

2.4.1.3 *Zachary's karate club*

The karate club network collects the interactions of 34 individuals over three years (Zachary, 1977). A conflict over the price of the karate lessons escalated and provoked the fission of the group as the supporters of the club's instructor formed a new organization, separate from the original one that stayed with the president. Thus, Zachary's data encompasses one of the few examples of nearly-definitive ground-truth communities, the two resulting groups (Hric et al., 2014). Many of the detection algorithms existing in the literature are tested on their ability to recuperate Zachary's factions. Such task is usually performed over a binary connectivity matrix for the members of the club (Girvan & Newman, 2002; Hric et al., 2014; Newman, 2004). In this study, instead, we used the weighted version provided by Zachary, in which the strength of an edge

is given by the number of external contexts where interactions between two individuals were observed (see Appendix A, Fig. S1, for the adjacency matrix that fixes some inconsistencies in Zachary's report). The weights were normalized to the [0,1] interval.

2.4.2 *Brain networks*

The community structure of several brain graphs was investigated. The anatomical connections of both the macaque and human brains as well as a model of functional interactions based on Kuramoto oscillators that were considered here are detailed in what follows.

2.4.2.1 *Macaque visual and sensorimotor anatomical network*

Cortico-cortical connections existing between large-scale areas of the macaque neocortex have been identified through anatomical tracing studies (Malcolm P Young, 1993). Among all the areas and pathways summarized in Young's paper, only those lying in the cortical visual and somatosensory-motor systems are considered here (see Fig. S2). This connectivity matrix, with 46 nodes, is only slightly different than the one utilized in the network structure study by Honey et al. (Honey, Kotter, Breakspear, & Sporns, 2007), where visual areas were labelled following (Felleman & Van Essen, 1991). Several connections are reciprocal. However, in general, the network is directed and binary, with 1's in a row indicating the efferent projections reported for the given area –see (Van Essen & Felleman, 1991; M P Young, 1993) for more details on the cortical areas.

2.4.2.2 *Human brain anatomical network*

An average human brain anatomical network (Yasser Iturria-Medina, Sotero, Toussaint, & Evans, 2014) was also constructed and analyzed in this study. The original data is freely available by The Cognitive Axon (CoAx) Lab, in the Center for the Neural Basis of Cognition and Department of Psychology at Carnegie Mellon University

(http://www.psy.cmu.edu/~coaxlab/data.html), who acquired and processed the data. Participants in the study included 60 subjects (29 males and 31 females; ages 18 to 45 years, mean 26 ± 6), recruited from the local Pittsburgh community and the Army Research Laboratory in Aberdeen Maryland. All subjects were neurologically healthy, with no history of either head trauma or neurological or psychiatric illness.

2.4.2.2.1 *Ethics statement*

The procedure was approved by the institutional review board at Carnegie Mellon University. Participants provided informed consent to participate in the study and consent to publish any research findings based on their provided data (Dunovan, Lynch, Molesworth, & Verstynen, 2015).

2.4.2.2.2 *Image acquisition*

Participants were scanned on a Siemens Verio 3T system in the Scientific Imaging & Brain Research (SIBR) Center at Carnegie Mellon University using a 32-channel head coil. Image collection was performed with the following parameters: 50 min, 257-direction DSI scan using a twice-refocused spin-echo EPI sequence and multiple q values (TR = 9916 ms, TE= 157 ms, voxel size = 2.4×2.4×2.4 mm, FoV = 231×231 mm, b-max = 5,000 s/mm^2, 51 slices). Head movement was minimized during the scan.

2.4.2.2.3 *Image processing*

All images were processed using a q-space diffeomorphic reconstruction method (Yeh & Tseng, 2011) to register the voxel coordinates into MNI space (Evans, Kamber, Collins, & MacDonald, 1994). Orientation distribution functions (ODFs) were reconstructed to a spatial resolution of 2 mm^3. As a result of the processing across all 60 subjects, a final template image

(CMU-60 DSI) was created by averaging the ODF maps. This template constitutes a detailed and unbiased representative map of the nervous fiber orientations in the young healthy brain.

Next, we estimated probabilistic axonal connectivity values between each brain voxel and the surface of each considered gray matter region (voxel-region connectivity) using a fully-automated fiber tractography algorithm (Y. Iturria-Medina et al., 2007) and the intravoxel fiber ODFs of the CMU-60 DSI Template. The tracking parameters were imposed as follows: a maximum of 500 mm trace length and a curvature threshold of ±90º. The anatomical regions were defined following the labeling procedure by Klein & Tourville (Klein & Tourville, 2012), from which 78 regions were considered –see (Y Iturria-Medina et al., 2016; Sanchez-Rodriguez et al., 2018) for more details. Based on the resulting voxel-region connectivity maps, the anatomical connection probability between any pair of regions i and j ($0 \leq ACP_{ji} \leq 1, ACP_{ji} = ACP_{ij}$) was calculated as the maximum voxel region connectivity value between both regions. For any pair of regions i and j, the ACP_{ji} measure (Y. Iturria-Medina et al., 2007) reflects the degree of evidence supporting the existence of the hypothetical white matter connection, independently of the density/strength of this connection. A network backbone, containing the dominant connections in the regional connectivity map, was computed using a minimum-spanning-tree based algorithm (Rubinov & Sporns, 2010). It was the resulting minimum spanning tree the network that we used (Fig. 3a).

2.4.2.3 *Human brain functional network*

To construct a representation of functional interactions in the brain, simulations of the Kuramoto model (Kuramoto, 1975) were performed. The Kuramoto model is a classical dynamical system that describes the behavior of a set of coupled oscillators. For the sake of consistency and contrast, the anatomical parcellation described in the previous section was conserved, while the

relative coupling between two nodes in the network of oscillators corresponded to the backbone-ACP measure between regions j and i. The evolution of the phase of the i-th oscillator, θ_i, is given by (Daffertshofer & van Wijk, 2011):

$$\dot{\theta}_i = \omega_i + \frac{\kappa}{n}\sum_{j=1}^{n} ACP_{ji} \sin\left(\theta_j - \theta_i\right) \quad (3)$$

where κ is a global coupling strength and ω_i is the intrinsic frequency of node i. In all our simulations, we drew the natural frequencies from a standard Gaussian distribution and the initial conditions from a uniform distribution in the interval $[0, 2\pi)$. A total of 250 sets of natural frequencies and initial conditions were used. System (3) was numerically solved via an Euler scheme, with time step $\Delta t = 0.001s$ and $t_{total} = 50s$. The first simulated $10s$ were discarded in all occasions to reduce the effect of transients in the results.

Intuitively, the collective behavior of the system depends on the parameter κ. Stronger interactions (high κ) overcome the dispersion of the intrinsic frequencies yielding coherence in the network, whereas in the low$-\kappa$ regime oscillators tend to remain asynchronous (Breakspear, Heitmann, & Daffertshofer, 2010; Daffertshofer & van Wijk, 2011). The degree of synchrony of the oscillators is quantified through the phase uniformity (K. V. Mardia, 1975):

$$R(t) = \frac{1}{n}\left|\sum_{j=1}^{n} e^{\iota\theta_j(t)}\right| \quad (4)$$

In our calculations, a grand-average phase uniformity value for each coupling strength, κ, was obtained by averaging $R(t)$ in the considered time interval across all simulations with such κ (Fig. S3). Similarly, a so-called κ-dependent functional connectivity matrix was also calculated. To do so, we followed Cabral et al. (Cabral, Hugues, Sporns, & Deco, 2011) and assumed an electrophysiological measure of the brain activity, such as a the mean firing rate or excitatory postsynaptic potential over the brain region, to be given as $y_i(t) = y_0 \sin\left(\theta_i(t)\right)$. Functional connectivity for a pair of nodes was then defined as the Pearson correlation between their $y_i(t)$ and

$y_j(t)$ signals, for each simulation (Bordier, Nicolini, & Bifone, 2017). The representative interaction matrix associated with κ was finally obtained after Fisher-transforming the pairwise correlation coefficients, averaging, performing one sample *t*-tests ($\alpha = 0.05$), correcting by false-discovery rate and applying the inverse transformation. Functional connectivity matrices for $\kappa = 5$ ($R \approx 0.12$), $\kappa = 30$ ($R \approx 0.47$) and $\kappa = 150$ ($R \approx 0.98$) are shown in Fig. 3b, Fig. 3c and Fig. 3d, respectively.

Fig. 3 goes around here

Fig. 3. Human brain networks used in the study. a) Anatomical connections between 78 brain areas. b) Functional connectivity obtained from superimposing Kuramoto oscillators to the matrix in (a). The global coupling strength is $\kappa = 5$. c) As in (b), with $\kappa = 30$. d) As in (b) and (c), with $\kappa = 150$.

2.4.3 *Data and code availability statement*

The datasets and codes analyzed during the current study are available from public repositories, which have been referenced throughout the paper. A specific set of codes containing a demonstration on how to concatenate the method pipeline is offered at https://www.soterolab.com/software. All calculations but the detection of clusters were performed in MATLAB R2018a (The MathWorks Inc., Natick, MA, USA). Python 2.7 (Python Software Foundation, https://www.python.org/) was used for an implementation of the *k*-modes algorithm.

3. **Results**

3.1 *Community detection in benchmark graphs*

Table 1 shows the results of the application of our method in benchmark graphs (see *Methods, Benchmarks*). The AMI value (see *Appendix D*), appearing in the last column, illustrates the degree of similarity between the obtained partitions and the ground-truth community structure known for each of the graphs. For both instances of the GN model (binary and weighted), the right partition was found over a range of IMFs. In the case of the LFR benchmarks, our method unveiled the 5 communities planted at IMF_6. Over slower IMFs than the ones reported in Table 1, coarser organization of the networks was in some cases observed, e.g. one of the ground-truth communities stood alone and the rest merged. The analysis of faster IMFs did not return any community structure (see *Appendix B* and *Appendix C*). Finally, in the case of the karate club, the two known fractions in which it split were nearly obtained over IMF_3.

Network	**No. Nodes**	**No. Communities**	**IMF found**	**AMI**
GN (binary)	128	4	8-4	1
GN (weighted)	128	4	8-4	1
LFR (binary)	100	5	6	1
LFR (weighted)	100	5	6	1
Zachary's	34	2	3	0.83

Table 1. Characteristics of the benchmarks and results of the application of our detection method. For each network, the number of nodes and known communities existing are given. The IMFs over which our method finds the right number of communities appear in the fourth column. In the last column, the adjusted mutual information values quantifying the degree of similarity between the solutions returned by the algorithm and the ground-truth structures are shown.

The results for the network of Zachary's karate club are further illustrated in Fig. 4. A schematic representations of the two-communities structure that was revealed appears in Fig. 4a. Fig. 4b contains information regarding the validation measures, showing the selection, by a majority rule, of two clusters in the data (see *Appendix B*). Likewise, other runs of the algorithm signaled the existence of two clusters. The left panel in Fig. 4c shows 50 of such partitions (one per row). In some of those, node 10 was assigned to the community we have called "1" (in blue). Thus, a consensus matrix (Fig. 4c, center panel) basically consists of binary values for the co-occurrences of all nodes in communities but those including node 10. Re-running the algorithm yielded 50 identical partitions (Fig. 4c, right). This partition (Fig. 4a) corresponds to the division reported by Zachary through observations of the karate club except for one member (node 9). This result is expected, according to the original paper and many others in which the karate club has been analyzed (Girvan & Newman, 2002; Hric et al., 2014), as the data apparently supports node 9's membership to the wrong faction.

The other structure (Fig. 4d) was obtained at IMF_2. This consists of three communities and suggests a pattern in which the two leaders (node 1, the instructor, and 34, the president) often interact with what presumably are their intimate friendship circles (nodes colored in blue and green, respectively) and the rest of the network conforms a different group. It is important to bear in mind that our analysis was performed over a weighted matrix accounting for several contexts in which the members of the club were seen interacting. Thus, the broader community found may represent a set of passive actors in the fission of the social network, some who "sit and wait" for the inputs coming from the rapidly exchanging groups of leaders and close followers. Therefore, the consideration of temporal scales –essential to our methodology– could be a key aspect to uncover new and interesting phenomena.

Fig. 4 goes around here

Fig. 4. Communities of Zachary's karate club. a) Representation of the community structure obtained over IMF_3. The two groups in which the network split after the conflict largely coincide with this pattern. The instructor's (president's) faction is shown in blue (green). The node colored in olive is misclassified as belonging to the president's faction, when compared to the ground-truth. The edges drawn are proportional to the weights of the connections. b) Validation indexes supporting the selection of $k = 2$ clusters in the data corresponding to IMF_3. Values of k for which k-modes yields singleton communities are not shown ($k = [4, 20], k \in \mathbb{N}$). c) Consensus clustering for partitions obtained with 50 different sets of random features of IMF_3. d) Over IMF_2, a new partition of three communities is obtained with small clusters including the instructor and the president.

Visualization of the community structures was achieved by means of SpringVisCom (Jeub, Balachandran, Porter, Mucha, & Mahoney, 2015).

3.2 *Community detection in brain networks*

3.2.1 The macaque visual and sensorimotor network

After testing the reliability of our method in several networks for which the community structure is known, we proceeded to its application to brain graphs. The first network considered was that of binary connections between cortical structures of the macaque visual and somatosensory-motor systems (see *Methods*, *Macaque visual and sensorimotor anatomical network)*. As such, a certain distribution of network nodes between those two functional systems was expected. Fig. 5 shows the hierarchical tree returned as consensus clustering for the macaque

anatomical network. At the highest level (two-clusters partition), the communities found correspond with the documented distinction between visual and sensorimotor areas (Hilgetag, Burns, O'Neill, Scannell, & Young, 2000; Van Essen & Felleman, 1991; M P Young, 1993). The sensorimotor system retained a single hierarchy, comprised of areas 3a, 3b, 1, 2, 5, Ri, S2, 7b, IG, ID, 4, 6 and SMA, at the following level whereas the other community split in two (showed in variations of blue). The first of the groups is composed of areas V1, V2, V3, VP, V3a, V4, V4t, MT, MSTd, MSTl, FST, PO, PIP, LIP, VIP and DP. The following cortical regions appear in the other set discovered: VOT, PITd, PITv, CITd, CITv, AITd, AITv, STPp, STPa, TF, TH, 7a, FEF, 46, TGV, ER and 35. These two smaller clusters largely resemble the traditional anatomical subdivision of the primate visual system in groups of 'ventral' and 'dorsal' areas (Hilgetag et al., 2000).

To further explore the performance of the algorithm herein introduced, we compared our results to the more conventional Louvain-like community detection methods –Newman-Girvan null model, implemented in the Brain Connectivity Toolbox (Rubinov & Sporns, 2010). Fig. S4 shows the results of 50 initial runs and the consensus partitions obtained in both cases. Both widely used criteria for the selection of the resolution parameter of the Louvain algorithm, i.e., $\gamma = 1$ or γ chosen as the value for which partitions are more similar, yielded the same result, which is a three-clusters structure. This organizational pattern echoed our last result (Fig. 5, Fig. S4b), with the exception of the ventral occipitotemporal (VOT) cortex which in some partitions appeared together with the ventral cortex though was eventually grouped with most dorsal areas.

Fig. 5 goes around here

Fig. 5. Dendrogram for the hierarchical consensus clustering of the primate visual and sensorimotor cortex. Somatosensory and motor areas are colored in green. Regions largely regarded as part of the visual system appear in blue. These are divided in two groups for predominantly dorsal and ventral anatomical areas.

3.2.2 The human brain networks

We have also applied our community detection algorithm to the network of cortical and subcortical neural conglomerates of the human brain (see *Methods*, *Human brain anatomical network* and *Human brain functional network*). Fig. 6 summarizes the results. Firstly, the anatomical connectivity matrix was considered. We obtained two organizational levels, which are depicted in Fig. 6a and 6b. The highest of the two (Fig. 6a) consists of two communities which are the left and right hemisphere of the brain. Running the clustering algorithm with the features of a different IMF yielded a structure of subdivisions of the two hemispheres (Fig. 6b). This four-community organization is practically symmetrical with the exception of the postcentral gyrus, the pallidum and the thalamus proper, which switch communities from one hemisphere to the other (see Fig. 6b and Table S1). The two communities to the top of the brain representation in the panel are mainly part of the frontal lobe, the cingulate cortex and the basal ganglia. On the other hand, those shown toward the bottom generally correspond with parietal, occipital and temporal areas (Klein & Tourville, 2012; Lanciego, Luquin, & Obeso, 2012). An instance of the validation indexes supporting the existence of four communities in this data was given as a demonstrative example in Fig. 2c.

The three synthetic functional networks (Fig 3b-d) resulting from superimposing Kuramoto oscillators to the matrix of anatomical connections were explored lastly. In Fig. 6c, the only

communities found in the $\kappa = 5$ case are shown. Two of those communities are a set of neural structures belonging to either the left (in blue) or right (in green) hemisphere. However, a third community (in maroon) consists of fifteen inter-hemispherical regions, all of which except for the right posterior cingulate appeared in a symmetrical manner in both hemispheres, including the totality of the occipital areas (see also Table S1). Fig. S5 shows the consensus clusters identified by using the two standard criteria for the resolution parameter in Louvain-modularity maximization. While three regions are grouped inter-hemispherically with those of their kind, in general, modularity maximization seems to fail at recognizing the functional relationships that are supposed to exist in this data, e.g., the mixed community pinpointed by our method. The tendency to split communities to obtain higher modularity values is also observed in Fig. S5 as anatomical communities are divided in a virtually arbitrary way (compare to Fig. 6b, for example). For $\kappa = 30$, our algorithm (and Louvain-maximization) returned the same two-hemispheres structure illustrated in Fig. 6a. Nevertheless, for $\kappa = 150$, no community structure was found over any of the IMFs and combinations of walkers considered (one cluster encompassed all nodes). This conclusion was reached by applying the criteria of *Appendix C*.

Fig. 6. Representation of the communities of the human brain anatomical and functional networks. a) Two-communities structure obtained for the anatomical network. b) Four-communities structure obtained for the anatomical network. c) Three-communities structure obtained for the functional network with global coupling parameter of the Kuramoto oscillators $\kappa = 5$. Colored nodes correspond to communities and their location, to average coordinates of the brain regions in MNI space.

Visualization of the community structures was achieved by means of BrainNet Viewer (Xia, Wang, & He, 2013).

4. **Discussion**

The problem of identifying modular structures at different scales of a network has captured the attention of the neuroscience community in recent times. Notably, Jeub et al. (Jeub et al., 2018) and Ashourvan et al. (Ashourvan et al., 2019) have introduced variants for the sweep through the Newman-Girvan modularity's γ-space eventually yielding hierarchical architecture. These methods have been tested in brain networks with encouraging results. Inherent limitations exist however, as the algorithms build on multi-scale modularity functions. Consequently, the exposed structures depend on the selection of several parameters and a null model, as in regular Louvain-like community detection (Blondel et al., 2008). Specifically, authors tend to recommend the utilization of null models that suit the characteristics of the data perfectly (Betzel et al., 2017). However, null models appear as abundant as detection algorithms in the literature oftentimes, making its selection a key step for the success or failure of the application of an algorithm (Sporns, 2013). Ideally, one would like to provide the user with minimum-input tools that can reveal the underlying structures of the data in natural ways.

In the recent past, much of the discussion as to the directions of neuroscience research was centered on avoiding univariate statistical comparisons and, instead, looking at the network interactions as a whole (Telesford, Simpson, Burdette, Hayasaka, & Laurienti, 2011). What is more, we believe that to better capture the complexities of a system like the brain, with multiple spatio-temporal scales and dynamic reconfigurations, the mere application of generic network science methods is not sufficient. Tools must be developed to account for the brain's unique

characteristics. Thus, in this paper, we have searched for organizational hierarchies through the temporal scales of the network's random walker signals, without necessitating to fix any parameter model. By doing so, the characteristics of the information flow in the brain are also incorporated (Sotero et al., 2019). The integration of the network's architecture with the dynamical interactions of the oscillatory modes in the brain is consequently suggested as an important consideration in clustering techniques.

4.1 *The brain organizes according to function*

In discussing the results of applying our methodology to the macaque visual and sensorimotor network, several insights can be gleaned. Firstly, the obvious and most straightforward precedent is the classification of areas according to the functional neural system they belong to, visual and sensorimotor (Van Essen & Felleman, 1991; M P Young, 1993). Additionally, two anatomical pathways have been identified in the visual system (Mishkin, Ungerleider, & Macko, 1983), which are usually known as dorsal (originating in the occipital cortex and terminating in the parietal lobe) and ventral (from occipital to temporal). These anatomically constrained divisions constituted the rationale for expecting the separation between somatosensory-motor and visual areas (which in turn was further divided into dorsal and ventral) by our community detection algorithm, given the numerous connections existing between areas in a functional system and somehow less connections with outsiders (Honey, Thivierge, & Sporns, 2010). We highlight, however, that limbic structures like the entorhinal cortex (ER) and the perirhinal cortex (A35), usually considered together with the sensorimotor system (Van Essen & Felleman, 1991), were clustered with most visual areas. Also, the ventral occipitotemporal area (VOT) appeared in an otherwise dorsal community. This result is analogous to the one described by Hilgetag et al. (Hilgetag et al., 2000) in a study on cluster organization of a similar, larger

network using now obsolete techniques. Moreover, in that work, prototypical ventral and dorsal areas V4 and 7a were clustered with the opposite streams. Notwithstanding the slight differences, the subdivision of the visual community obtained here closely resembles the ventral and dorsal streams reported by Hilgetag et al. Several more recent studies utilizing modularity maximization methods have detected similar sets of communities, yet different macaque datasets have been used (Sporns & Betzel, 2016).

We believe that the minor discrepancies in the analysis of the primate cortical network commented in the above paragraph are due to two main issues. First of all, we should revisit the limitations that exist intrinsic to each community detection method. These, of course, affect the partitions returned given any connectivity matrix, revealing the necessity of continuing to develop tools and migrating to more comprehensive approaches that use as much valuable network information as possible. Another important matter to recall is that imperfections exist in data as well. For example, the matrix of connections we used (M P Young, 1993) encompasses reports from several studies, which sometimes even employed different anatomical parcellations. Also, this matrix accounts only for the existence or absence of reports of links between areas of the macaque cortex, without considering the strength of the connections. Whether datasets accurately reflect the particularities of the connections existing in the brain or not will remain a fundamental question in neuroscience.

The next application of our newly introduced community-finding method was to human brain networks. We considered a connectivity matrix in which each entry reflects the evidence of the existence of a white matter link between two brain regions (Y. Iturria-Medina et al., 2007), given a template of such connections in the young healthy brain (Dunovan et al., 2015). Dominant connections were retained through a minimum spanning tree algorithm. Although the minimum

spanning tree trims connections and does not contain loops, it is believed to provide a correct representation of any denser brain network to which it is applied, retaining paramount topological characteristics like its small-worldness and scale-freeness (Tewarie, van Dellen, Hillebrand, & Stam, 2015).

We have found two partitions of the anatomical network, which seemingly follow physical proximity and functional specialization rules. In the case of the first partition, commissural fibers appear to act as those rare links to members of other communities for the two brain hemispheres were perfectly separated. Each hemisphere split into two communities over the other partition found. The four-communities structure was almost bilaterally symmetrical, as only the postcentral gyrus, pallidum and thalamus proper exchanged membership, i.e., they grouped with most frontal, cingulate and basal ganglia regions in the right hemisphere and with parietal-occipital-temporal areas in the left. Variations in local connections within hemispheres may be the reason why these regions behaved in such a way. In principle, the thalamus, as universal relay station, and the pallidum projecting to the thalamus (Lanciego et al., 2012) should have no constraints to belong to one or the other intra-hemisphere community found. The postcentral gyrus, although deemed part of the parietal lobe is located in the vicinity of the frontal lobe, possibly explaining its grouping with such neural structures. The association of brain regions to perform processes and functions could also be reflected in the anatomical network and, consequently, in the communities obtained. For instance, having the basal forebrain clustered with frontal areas may be justified by the fact that its projections to the prefrontal cortex are paramount for attention, learning and memory, and decision-making (Tashakori-Sabzevar & Ward, 2018).

The study of the matrices for the interaction of oscillators over the anatomical frame yielded stimulating results. For one thing, the mechanism for the transitions between functional states

relates to the tuning of the coupling parameter in the model. Three communities appeared in the low-coupling regime ($\kappa = 5$), one of them presenting areas from both brain hemispheres in a close to symmetrical pattern. When the coupling strength was raised to $\kappa = 30$, the mixed community was destroyed and the only recognizable pattern was the one of two separate brain hemispheres, which was one already existing in the anatomical network. This is because the connectivity matrix of a set of Kuramoto oscillators overcomes the dispersion of natural frequencies for higher values of the coupling parameter (Breakspear et al., 2010). All in all, our results suggest additional evidence for the known link between structure and function in the brain (Honey et al., 2010). Higher functional couplings amplify the anatomical subdivision of the network in two hemispheres.

Other characteristics of the clusters of the functional networks are also noteworthy. For example, all the occipital areas (Klein & Tourville, 2012), responsible for vision (Johns, 2014), appeared in the mixed community of the $\kappa = 5$ -regime. Previous studies of modular organization in functional graphs found robust grouping in the occipital cortex (Meunier, 2009). Among the other areas present in the intra-hemispherical cluster: the superior parietal lobe has abundant connections with the occipital lobe and participates in visuospatial perception (Johns, 2014); the precuneus has a major role in visuo-spatial imagery (Cavanna & Trimble, 2006) and the posterior cingulate cortex is considered a core node of the default mode network (DMN) and to be involved in many tasks (Yasser Iturria-Medina et al., 2014). The inter-hemispherical community obtained appears to be one extended circuit concerned with the function of vision. The rest of the nodes within each hemisphere may consequently process all the non-visual stimuli, possibly constituting an optimal configuration for speedy and accurate performance on cognitive tasks (Garcia et al., 2018). We believe that the detection of these communities supports the notion of functional

integration in the brain, whereas evidence for segregation can also be found in a division that isolates units specialized in handling with visual stimuli.

4.2 *On the strengths and limitations of the algorithm*

To conclude this section we would like to highlight important features of our method. One interesting scenario is the one of functional interactions with $\kappa = 150$. The large synchronization seen there yields a close-to- n-regular graph (Fig. 3d) that goes together with the algorithm identifying a single group of nodes. In fact, one recommended practice for testing new community detection algorithms is checking that they do not return group structure in the absence of it (Hric et al., 2014). Of further value is the effective performance shown when searching for the community structure of the benchmark graphs, all of which presented different network characteristics. On this topic, one must also mention the diversity of the real networks whose organization in groups was explored. The karate club is a small, weighted and undirected network. On the other hand, the macaque visual and sensorimotor network is binary and directed. The human brain networks, with larger dimension, had different levels of sparsity. Many algorithms, e.g. Infomap, are initially designed for a specific type of graph (Fortunato & Hric, 2016) and only later extended. However, our method seems to be primed to perform reliable community detection in moderate-size networks like the ones associated with the brain's large-scale activity.

It was not in our interest, though, to test its suitability for high-dimensional graphs. In that case, the demand for computing resources would grow. Firstly, the computational cost of *k*-modes scales linearly with the number of objects and many random initializations of the modes are required to find a reliable clustering solution (Nguyen, 2017). Secondly, the number of possible combinations of nodes appearing between zero-crossings of an IMF would increase as well, so the implementation of *k*-modes must be optimal to handle a large number of binary features. We

bypassed some of these complications by using the University of Calgary's (https://hpc.ucalgary.ca/resources) and Compute Canada's (https://docs.computecanada.ca/wiki/Getting_started) computing clusters resources, where calculations were run in parallel. The other two stages of the algorithm, namely the random walks and empirical mode decomposition are already fast enough through built-in functions in MATLAB. Other speeding-up alternatives for the clustering problem must be explored, however. One solution for reducing dimensionality is the selection of relevant features in the data (Ronan et al., 2016). Because the random choice of features (i.e., the set of walks) could also generate solutions that differ from the existing structure if many unrepresentative features were combined, the identification of a relevant set is also desired for the stability of the method. Nevertheless, dimensionality reduction by feature selection is a complicated matter on its own (Steinbach et al., 2004), especially in unsupervised learning, and state-of-the-art techniques degrade the results in some situations (Ronan et al., 2016). In short, we do not recommend the use of the method presented in this paper on large networks until further steps towards optimization are taken.

5. **Conclusion**

In summary, we have introduced an approach for the detection of modular organization by considering the temporal scales of the information flow over the networks of interest. This new tool insinuates particularly useful for the analysis of large-scale brain graphs, for which: 1) the transmission of information is a process of paramount importance and 2) a desirable balance between accuracy and computational complexity of the community detection algorithm can be achieved given the current implementation state. We find several organizational patterns existing in the brain anatomical and functional networks –also in the social network that we study. These structures may coexist together, in a dynamical way that is given by the temporal scales of the

activity they produce, guaranteeing functional independence and coordination. Our results promise a shift of focus in the discussion surrounding the occurrence of community structure.

Appendix A. Supplementary material

Supplementary data related to this article can be found at [*insert link here*].

Appendix B. Cluster validation indexes

In what follows, n is the number of objects to be clustered and k is the number of such clusters. Let $\{C_1, C_2, \cdots, C_k\}$ be a partition of the integers from 1 to n such that $i \epsilon C_q$ if the ith object, $\boldsymbol{x}_i$, belongs to the qth cluster. The centroid (mode) of a subset C_q of n_q objects is the vector $\boldsymbol{m}_q$ that minimizes the sum of the distances to all the objects in C_q. Supposing the (mismatching similarity) distances between every i and j, d_{ij} are known, which also applies for the distances between objects and their clusters' centroids, then:

- The *Calinski-Harabasz* index (Calinski & Harabasz, 1974), also known as pseudo-F ratio, is defined as:

$$F(k) = \frac{SSA/(k-1)}{SSW/(n-k)} \qquad \text{Eq. (A.1)}$$

SSW and SSA are the within-group sum of squares and the among-group sum of squares, respectively. Adapting (Anderson, 2001), these quantities are obtained from the matrix of distances between pair of objects:

$$SSW = \sum_{q=1}^{k} \frac{1}{n_q} \sum_{\substack{i,j \in C_q \\ i<j}} d_{ij}^2 \qquad \text{Eq. (A.2)}$$

$$SSA = SST - SSW, SST = \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d_{ij}^2 \qquad \text{Eq. (A.3)}$$

The among-group distances are large compared to the within-group distances in the case of high separateness and compactness. Thus, maximum values are taken to represent the correct number of clusters (Milligan & Cooper, 1985).

- The *C-Index* (Milligan & Cooper, 1985) is calculated as:

$$C(k) = \frac{SW - S_{min}}{S_{max} - S_{min}} \qquad \text{Eq. (A.4)}$$

where SW is the sum of the within-cluster distances:

$$SW = \sum_{q=1}^{k} \sum_{\substack{i,j \in C_q \\ i<j}} d_{ij} \qquad \text{Eq. (A.5)}$$

and $S_{min}(S_{max})$ is the sum of the N_W smallest (largest) distances in the dataset. N_W is the total number of pair of objects in the same cluster, $N_W = \sum_{q=1}^{k} \frac{n_q(n_q-1)}{2}$ (Charrad et al., 2014). C-Index is restricted to the interval $(0,1)$ and its minimum value suggests the optimal number of clusters (Milligan & Cooper, 1985).

- The *Duda-Hart* (Duda et al., 2001) score is inspired in the fact that the sum of squared-errors corresponding to a partition decreases with k. Thus, in conventional Euclidean-distance clustering problems, the optimal number of clusters is the smallest k such that $\frac{J(k)}{J(k-1)}$ is smaller than certain critical value (Milligan & Cooper, 1985). In our case of binary distances, we limit ourselves to request that ratio to be minimal, indicating a possible correct number of clusters, and define $J(k)$ as a "sum of mismatching similarity distances error":

$$J(k) = \sum_{q=1}^{k} \sum_{i \in C_q} d(\boldsymbol{x_i}, \boldsymbol{m}_q) \qquad \text{Eq. (A.6)}$$

- The *Silhouette* width (Kaufman & Rousseeuw, 1990) is calculated with the following expression:

$$S(k) = \frac{1}{n} \sum_{i=1}^{n} \frac{b_i - a_i}{\max\{a_i, b_i\}} \qquad \text{Eq. (A.7)}$$

here, a_i is the average distance from the ith point to every other object in its cluster: $a_i = \frac{1}{n_q - 1} \sum_{j \in \{C_q \setminus i\}} d_{ij}$ and b_i is the minimum average distance from the ith object to all objects of other clusters, minimized over the clusters, namely: $b_i = \min_{s \neq q}\{d_{iC_s}\}$; $d_{iC_s} = \frac{1}{n_s} \sum_{j \in C_s} d_{ij}$ (Charrad et al., 2014). The index can take values in the interval $[-1,1]$ with negative values indicating the

clustering solution is not accurate, and understandably so, as the minimum average distance from many objects to other clusters would be bigger than the dissimilarity to objects of the clusters where they belong to. On the other hand, the maximum value is taken to represent the optimal number of clusters in the data (Charrad et al., 2014).

- The *Dunn* index (Dunn, 1973) is generally defined as:

$$D(k) = \frac{\min_{1 \le r < s \le k} \mathrm{I}(C_r, C_s)}{\max_{1 \le q \le k} \Delta_q} \qquad \text{Eq. (A.8)}$$

where Δ_q is the diameter of the qth cluster and $\mathrm{I}(C_r, C_s)$ is the intercluster distance between C_r and C_s. Out of the many variants available for computing both these quantities, we use the ones recommended by (Bezdek & Pal, 1998):

$$\Delta_q = \frac{1}{n_q} \sum_{i \in C_q} d\left(\boldsymbol{x}_i, \boldsymbol{m}_q\right) \qquad \text{Eq. (A.9)}$$

$$\mathrm{I}(C_r, C_s) = \frac{1}{n_r n_s} \sum_{\substack{i \in C_r \\ j \in C_s}} d_{ij} \qquad \text{Eq. (A.10)}$$

D is maximized when the clusters are compact (the diameter is small) and separate (the intercluster distance is large) (Milligan & Cooper, 1985).

- The *Davies-Bouldin* index (Davies & Bouldin, 1979) is also a function of the ratio of within-cluster dispersions and the between-clusters separation. When using mismatching dissimilarities, it can be calculated as:

$$DB(k) = \frac{1}{k} \sum_{q=1}^{k} \max_{q \neq l} \left(\frac{\Delta_q + \Delta_l}{\sigma_{ql}} \right) \qquad \text{Eq. (A.11)}$$

where Δ_q is defined as in Eq. (A.9) and σ_{ql} is the distance between the centroids of clusters C_q and C_l, $\sigma_{ql} = d\left(\boldsymbol{m}_q, \boldsymbol{m}_l\right)$ (Charrad et al., 2014). The smaller $DB(k)$, the better the partition (Dubes, 1987).

The diameter, Δ_q, is zero for clusters with single members. Thus, as highlighted by Davies & Bouldin, theirs and most measures, have limited meaning for singleton clusters.

Appendix C. Decision rules for random data

The above-mentioned decision rules are not adequate for identifying the correct number of clusters in the limit case of two versus one cluster, as such indexes are not defined for partitions of a lone community (Dubes, 1987). Although the Duda-Hart index was originally designed to reject the existence of only one cluster in the data, the critical value used for such means was obtained by supposing that data came from a normal distribution (Duda et al., 2001), which does not hold in our case of binary variables. Two of the other measures had either simplistic or well-established criteria that could be applied when they and the majority of the indexes signalled the presence of two communities. The first one is inspecting for negative silhouette values (Kaufman & Rousseeuw, 1990). The second rule checks for the presence of a significant drop in the curve of $DB(k)$ at $k = 2$ (Dubes, 1987). An user can conclude one cluster is present when $DB(k)$ has a minimum at 2 but

$$DB(3) - DB(2) < \frac{2}{3}\sum_{k=1}^{3}|DB(k_{max} - k) - DB(k_{max} - k - 1)| \qquad \text{Eq. (C.1)}$$

These two criteria were applied to resolve the optimal number of clusters in the limit case situation over all the networks. Consistently, spurious two-clusters partitions over different IMFs were rejected (one single cluster existed). On the contrary, meaningful or real two-communities patterns were indicated as correct by our method, and restated by the silhouette and the DB-based limit-case criteria.

Appendix D. Estimating the similarity of partitions

To determine the effectiveness of our clustering technique in retrieving a planted structure, we computed the adjusted mutual information, which establishes a measure of similarity between

two partitions based on information theory while adjusting for chance (N. Vinh et al., 2010; Weir, Emmons, Gibson, Taylor, & Mucha, 2017):

$$AMI(X,Y) = \frac{MI(X,Y) - E\big(MI(X,Y)\big)}{\max\big(H(X),H(Y)\big) - E\big(MI(X,Y)\big)} \qquad \text{Eq. (D.1)}$$

where $MI(X,Y)$ is the mutual information between random variables X and Y, $H(X)$ is the entropy of X and the expected mutual information, $E\big(MI(X,Y)\big)$, is obtained as in (N. X. Vinh, Epps, & Bailey, 2009). A MATLAB implementation of AMI is available from the Network Community Toolbox (http://commdetect.weebly.com/).

Acknowledgments

We are grateful to Anthony Winder for his comments on the manuscript. We also thank the High Performance Computing group at the University of Calgary (UofC) for their support accessing UofC's computing resources.

Funding resources

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2015-05966).

References

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*(1), 32–46. http://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Ashourvan, A., Telesford, Q. K., Verstynen, T., Vettel, J. M., & Bassett, D. S. (2019). *Multi-scale detection of hierarchical community architecture in structural and functional brain networks*. *PLoS ONE* (Vol. 14). http://doi.org/10.1371/journal.pone.0215520

Assent, I. (2012). Clustering high dimensional data. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(4), 340–350. http://doi.org/10.1002/widm.1062

Bassett, D. S., Khambhati, A. N., & Grafton, S. T. (2017). Emerging Frontiers of Neuroengineering: A Network Science of Brain Connectivity. *Annual Review of Biomedical Engineering*, *19*(1), 327–352. http://doi.org/10.1146/annurev-bioeng-071516-044511

Bassett, D. S., Porter, M. A., Wymbs, N. F., Grafton, S. T., Carlson, J. M., & Mucha, P. J. (2013). Robust detection of dynamic community structure in networks. *Chaos*, *23*(1), 1–16. http://doi.org/10.1063/1.4790830

Betzel, R. F., Medaglia, J. D., Papadopoulos, L., Baum, G. L., Gur, R., Gur, R., … Bassett, D. S. (2017). The modular organization of human anatomical brain networks: Accounting for the cost of wiring. *Network Neuroscience*, *1*(1), 42–68. http://doi.org/10.1162/NETN_a_00002

Bezdek, J. C., & Pal, N. R. (1998). Some new indexes of cluster validity. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, *28*(3), 301–315. http://doi.org/10.1109/3477.678624

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10). http://doi.org/10.1088/1742-5468/2008/10/P10008

Bordier, C., Nicolini, C., & Bifone, A. (2017). Graph analysis and modularity of brain functional connectivity networks: Searching for the optimal threshold. *Frontiers in Neuroscience*, *11*(AUG), 1–9. http://doi.org/10.3389/fnins.2017.00441

Breakspear, M., Heitmann, S., & Daffertshofer, A. (2010). Generative models of cortical oscillations: neurobiological implications of the kuramoto model. *Frontiers in Human Neuroscience*, *4*(November), 190. http://doi.org/10.3389/fnhum.2010.00190

Breakspear, M., Sporns, O., Honey, C. J., & Ko, R. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales.

Cabral, J., Hugues, E., Sporns, O., & Deco, G. (2011). Role of local network oscillations in resting-state functional connectivity. *NeuroImage*, *57*(1), 130–139. http://doi.org/10.1016/j.neuroimage.2011.04.010

Calinski, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics - Theory and Methods*, *3*(1), 1–27. http://doi.org/10.1080/03610927408827101

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*, *129*(3), 564–583. http://doi.org/10.1093/brain/awl004

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust : An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, *61*(6), http://www.jstatsoft.org/. http://doi.org/10.18637/jss.v061.i06

Daffertshofer, A., & van Wijk, B. C. M. (2011). On the Influence of Amplitude on the Connectivity between Phases. *Frontiers in Neuroinformatics*, *5*(July), 6. http://doi.org/10.3389/fninf.2011.00006

Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-1*(2), 224–227.

http://doi.org/10.1109/TPAMI.1979.4766909

Dubes, R. C. (1987). How many clusters are best? - An experiment. *Pattern Recognition*, *20*(6), 645–663. http://doi.org/10.1016/0031-3203(87)90034-3

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern Classification* (2nd Ed). New York: John Wiley & Sons.

Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, *3*(3), 32–57. http://doi.org/10.1080/01969727308546046

Dunovan, K., Lynch, B., Molesworth, T., & Verstynen, T. (2015). Competing basal ganglia pathways determine the difference between stopping and deciding not to go. *eLife*, *4*, 1–24. http://doi.org/10.7554/elife.08723

Evans, A. C., Kamber, M., Collins, D. L., & MacDonald, D. (1994). An MRI-Based Probabilistic Atlas of Neuroanatomy. In *Magnetic Resonance Scanning and Epilepsy* (pp. 263–274). Boston, MA: Springer US. http://doi.org/10.1007/978-1-4615-2546-2_48

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, *1*(1), 1–47. http://doi.org/10.1093/cercor/1.1.1

Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *Physics Reports*, *659*, 1–44. http://doi.org/10.1016/j.physrep.2016.09.002

Friston, K. (2011). Functional and effective connectivity: a review. *Brain Connectivity*, *1*(1), 13–36. http://doi.org/10.1089/brain.2011.0008

Fukushima, M., Betzel, R. F., He, Y., van den Heuvel, M. P., Zuo, X. N., & Sporns, O. (2018). Structure–function relationships during segregated and integrated network states of human

brain functional connectivity. *Brain Structure and Function*, *223*(3), 1091–1106. http://doi.org/10.1007/s00429-017-1539-3

Garcia, J. O., Ashourvan, A., Muldoon, S., Vettel, J. M., & Bassett, D. S. (2018). Applications of Community Detection Techniques to Brain Graphs: Algorithmic Considerations and Implications for Neural Function. *Proceedings of the IEEE*, *106*(5), 846–867. http://doi.org/10.1109/JPROC.2017.2786710

Gates, K. M., Henry, T., Steinley, D., & Fair, D. A. (2016). A Monte Carlo Evaluation of Weighted Community Detection Algorithms. *Frontiers in Neuroinformatics*, *10*(November). http://doi.org/10.3389/fninf.2016.00045

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, *99*(12), 7821–7826. http://doi.org/10.1073/pnas.122653799

Hilgetag, C., Burns, G. A. P. C., O'Neill, M. A., Scannell, J. W., & Young, M. P. (2000). Anatomical connectivity defines the organization of clusters of cortical areas in the macaque and the cat. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1393), 91–110. http://doi.org/10.1098/rstb.2000.0551

Honey, C. J., Kotter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, *104*(24), 10240–10245. http://doi.org/10.1073/pnas.0701519104

Honey, C. J., Thivierge, J. P., & Sporns, O. (2010). Can structure predict function in the human brain? *NeuroImage*, *52*(3), 766–776. http://doi.org/10.1016/j.neuroimage.2010.01.071

Hric, D., Darst, R. K., & Fortunato, S. (2014). Community detection in networks: Structural communities versus ground truth. *Physical Review E - Statistical, Nonlinear, and Soft*

Matter Physics, *90*(6), 1–19. http://doi.org/10.1103/PhysRevE.90.062805

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Yen, N., … Liu, H. H. (1996). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Royal Society of London Proceedings Series A*, *454*(1), 903–995. http://doi.org/10.1098/rspa.1998.0193

Huang, Z. (1998). Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values. *Data Mining and Knowledge Discovery*, *2*(3), 283–304. http://doi.org/10.1023/A:1009769707641

Iturria-Medina, Y., Canales-Rodríguez, E. J., Melie-García, L., Valdés-Hernández, P. A., Martínez-Montes, E., Alemán-Gómez, Y., & Sánchez-Bornot, J. M. (2007). Characterizing brain anatomical connections using diffusion weighted MRI and graph theory. *NeuroImage*, *36*(3), 645–660. http://doi.org/10.1016/j.neuroimage.2007.02.012

Iturria-Medina, Y., Sotero, R. C., Toussaint, P. J., & Evans, A. C. (2014). Epidemic Spreading Model to Characterize Misfolded Proteins Propagation in Aging and Associated Neurodegenerative Disorders. *PLoS Computational Biology*, *10*(11). http://doi.org/10.1371/journal.pcbi.1003956

Iturria-Medina, Y., Sotero, R. C., Toussaint, P. J., Mateos-Perez, J. M., Evans, A. C., & Initiative, T. A. D. N. (2016). Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat Commun*, *7*(May), 11934. http://doi.org/10.1038/ncomms11934

Jeub, L. G. S., Balachandran, P., Porter, M. A., Mucha, P. J., & Mahoney, M. W. (2015). Think locally, act locally: Detection of small, medium-sized, and large communities in large networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *91*(1), 1–29.

http://doi.org/10.1103/PhysRevE.91.012821

Jeub, L. G. S., Sporns, O., & Fortunato, S. (2018). Multiresolution Consensus Clustering in Networks. *Scientific Reports*, *8*(1), 1–16. http://doi.org/10.1038/s41598-018-21352-7

Jiliang Tang, Salem Alelyani, & Huan Liu. (2014). Feature Selection for Classification: A Review. *Data Classification: Algorithms and Applications*, 571–605. http://doi.org/10.1201/b17320

Johns, P. (2014). Functional neuroanatomy. In *Clinical Neuroscience* (pp. 27–47). Elsevier. http://doi.org/10.1016/B978-0-443-10321-6.00003-5

K. V. Mardia. (1975). Statistics of Directional Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, *37*(3), 349–393. Retrieved from https://www.jstor.org/stable/2984782

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data*. (L. Kaufman & P. J. Rousseeuw, Eds.). Hoboken, NJ, USA: John Wiley & Sons, Inc. http://doi.org/10.1002/9780470316801

Klein, A., & Tourville, J. (2012). 101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol. *Frontiers in Neuroscience*, *6*(DEC), 1–12. http://doi.org/10.3389/fnins.2012.00171

Kuramoto, Y. (1975). Self-entrainment of a population of coupled non-linear oscillators. *International Symposium on Mathematical Problems in Theoretical Physics*, *39*, 420–422. http://doi.org/10.1007/BFb0013365

Lancichinetti, A., & Fortunato, S. (2009). Community detection algorithms: A comparative analysis. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *80*(5), 1–11. http://doi.org/10.1103/PhysRevE.80.056117

Lancichinetti, A., Fortunato, S., & Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *78*(4), 1–5. http://doi.org/10.1103/PhysRevE.78.046110

Lanciego, J. L., Luquin, N., & Obeso, J. A. (2012). Functional neuroanatomy of the basal ganglia. *Cold Spring Harbor Perspectives in Medicine*, *2*(12), 1–20. http://doi.org/10.1101/cshperspect.a009621

Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability* (pp. 281–297).

Meunier, D. (2009). Hierarchical modularity in human brain functional networks. *Frontiers in Neuroinformatics*, *3*(October), 1–12. http://doi.org/10.3389/neuro.11.037.2009

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, *50*(2), 159–179. http://doi.org/10.1007/BF02294245

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, *6*(C), 414–417. http://doi.org/10.1016/0166-2236(83)90190-X

Moradi, N., Dousty, M., & Sotero, R. C. (2019). Spatiotemporal empirical mode decomposition of resting-state fMRI signals : application to global signal regression, 1–23.

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *69*(6), 5. http://doi.org/10.1103/PhysRevE.69.066133

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the*

National Academy of Sciences, *103*(23), 8577–8582. http://doi.org/10.1073/pnas.0601602103

Nguyen, H. H. (2017). Clustering categorical data using community detection techniques. *Computational Intelligence and Neuroscience*, *2017*. http://doi.org/10.1155/2017/8986360

Pons, P., & Latapy, M. (2005). Computing Communities in Large Networks Using Random Walks. In *Download.Springer.Com* (pp. 284–293). http://doi.org/10.1007/11569596_31

Porter, M. A., Mucha, P. J., Newman, M. E. J., & Warmbrand, C. M. (2005). A network analysis of committees in the U.S. House of Representatives. *Proceedings of the National Academy of Sciences*, *102*(20), 7057–7062. http://doi.org/10.1073/pnas.0500191102

Ronan, T., Qi, Z., & Naegle, K. M. (2016). Avoiding common pitfalls when clustering biological data. *Science Signaling*, *9*(432), 1–13. http://doi.org/10.1126/scisignal.aad1932

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, *105*(4), 1118–1123. http://doi.org/10.1073/pnas.0706851105

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. http://doi.org/10.1016/j.neuroimage.2009.10.003

Sanchez-Rodriguez, L. M., Iturria-Medina, Y., Baines, E. A., Mallo, S. C., Dousty, M., & Sotero, R. C. (2018). Design of optimal nonlinear network controllers for Alzheimer's disease. *PLOS Computational Biology*, *14*(5), e1006136. http://doi.org/10.1371/journal.pcbi.1006136

Sotero, R. C. (2016). Topology, Cross-Frequency, and Same-Frequency Band Interactions Shape the Generation of Phase-Amplitude Coupling in a Neural Mass Model of a Cortical

Column. *PLoS Computational Biology*, *12*(11), 1–29. http://doi.org/10.1371/journal.pcbi.1005180

Sotero, R. C., Sanchez-Rodriguez, L. M., Dousty, M., Iturria-Medina, Y., & Sanchez-Bornot, J. M. (2019). Cross-Frequency Interactions During Information Flow in Complex Brain Networks Are Facilitated by Scale-Free Properties. *Frontiers in Physics*, *7*(July), 1–16. http://doi.org/10.3389/fphy.2019.00107

Sotero, R. C., & Trujillo-Barreto, N. J. (2008). Biophysical model for integrating neuronal activity, EEG, fMRI and metabolism. *NeuroImage*, *39*, 290–309. http://doi.org/10.1016/j.neuroimage.2007.08.001

Sporns, O. (2013). Structure and function of complex brain networks. *Dialogues in Clinical Neuroscience*, *15*(3), 247–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24174898

Sporns, O., & Betzel, R. F. (2016). Modular Brain Networks. *Annual Review of Psychology*, *67*(1), 613–640. http://doi.org/10.1146/annurev-psych-122414-033634

Steinbach, M., Ertöz, L., & Kumar, V. (2004). The Challenges of Clustering High Dimensional Data. *New Directions in Statistical Physics*, 273–309. http://doi.org/10.1007/978-3-662-08968-2_16

Strehl, A., & Ghosh, J. (2003). Cluster ensembles - A knowledge reuse framework for combining multiple partitions. *Journal of Machine Learning Research*, *3*(3), 583–617. http://doi.org/10.1162/153244303321897735

Tashakori-Sabzevar, F., & Ward, R. D. (2018). Basal Forebrain Mediates Motivational Recruitment of Attention by Reward-Associated Cues. *Frontiers in Neuroscience*, *12*(October), 1–11. http://doi.org/10.3389/fnins.2018.00786

Telesford, Q. K., Simpson, S. L., Burdette, J. H., Hayasaka, S., & Laurienti, P. J. (2011). The Brain as a Complex System: Using Network Science as a Tool for Understanding the Brain. *Brain Connectivity*, *1*(4), 295–308. http://doi.org/10.1089/brain.2011.0055

Tewarie, P., van Dellen, E., Hillebrand, A., & Stam, C. J. (2015). The minimum spanning tree: An unbiased method for brain network analysis. *NeuroImage*, *104*, 177–188. http://doi.org/10.1016/j.neuroimage.2014.10.015

Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., … Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. http://doi.org/10.1152/jn.00338.2011

Traud, A. L., Kelsic, E. D., Mucha, P. J., & Porter, M. A. (2011). Comparing Community Structure to Characteristics in Online Collegiate Social Networks. *SIAM Review*, *53*(3), 526–543. http://doi.org/10.1137/080734315

Valdes-Sosa, P. A., Sanchez-Bornot, J. M., Sotero, R. C., Iturria-Medina, Y., Aleman-Gomez, Y., Bosch-Bayard, J., … Ozaki, T. (2009). Model driven EEG/fMRI fusion of brain oscillations. *Human Brain Mapping*, *30*(9), 2701–2721. http://doi.org/10.1002/hbm.20704

Van Essen, D. C., & Felleman, D. J. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex*, *1*, 1–47. Retrieved from http://www.cns.nyu.edu/~tony/vns/readings/felleman-vanessen-1991.pdf

Vinh, N., Epps, J., & Bailey, J. (2010). Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *The Journal of Machine Learning Research*, *11*, 2837–2854. http://doi.org/10.1182/blood-2008-03-145946

Vinh, N. X., Epps, J., & Bailey, J. (2009). Information theoretic measures for clusterings

comparison. In *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09* (pp. 1–8). New York, New York, USA: ACM Press. http://doi.org/10.1145/1553374.1553511

Weir, W. H., Emmons, S., Gibson, R., Taylor, D., & Mucha, P. J. (2017). Post-processing partitions to identify domains of modularity optimization. *Algorithms*, *10*(3), 1–21. http://doi.org/10.3390/a10030093

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. *PLoS ONE*, *8*(7). http://doi.org/10.1371/journal.pone.0068910

Yeh, F. C., & Tseng, W. Y. I. (2011). NTU-90: A high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. *NeuroImage*, *58*(1), 91–99. http://doi.org/10.1016/j.neuroimage.2011.06.021

Young, M. P. (1993). The organization of neural systems in the primate cerebral cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *252*(1333), 13–18. http://doi.org/10.1098/rspb.1993.0040

Young, M. P. (1993). The organization of neural systems in the primate cerebral cortex. *Proc.R.Soc.Lond B Biol.Sci.*, *252*(1333), 13–18.

Zachary, W. W. (1977). An Information Flow Model for Conflict and Fission in Small Groups. *Journal of Anthropological Research*, *33*(4), 452–473. http://doi.org/10.1086/jar.33.4.3629752

Zhang, Z., Shan, T., & Chen, G. (2013). Random walks on weighted networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *87*(1), 1–6. http://doi.org/10.1103/PhysRevE.87.012112

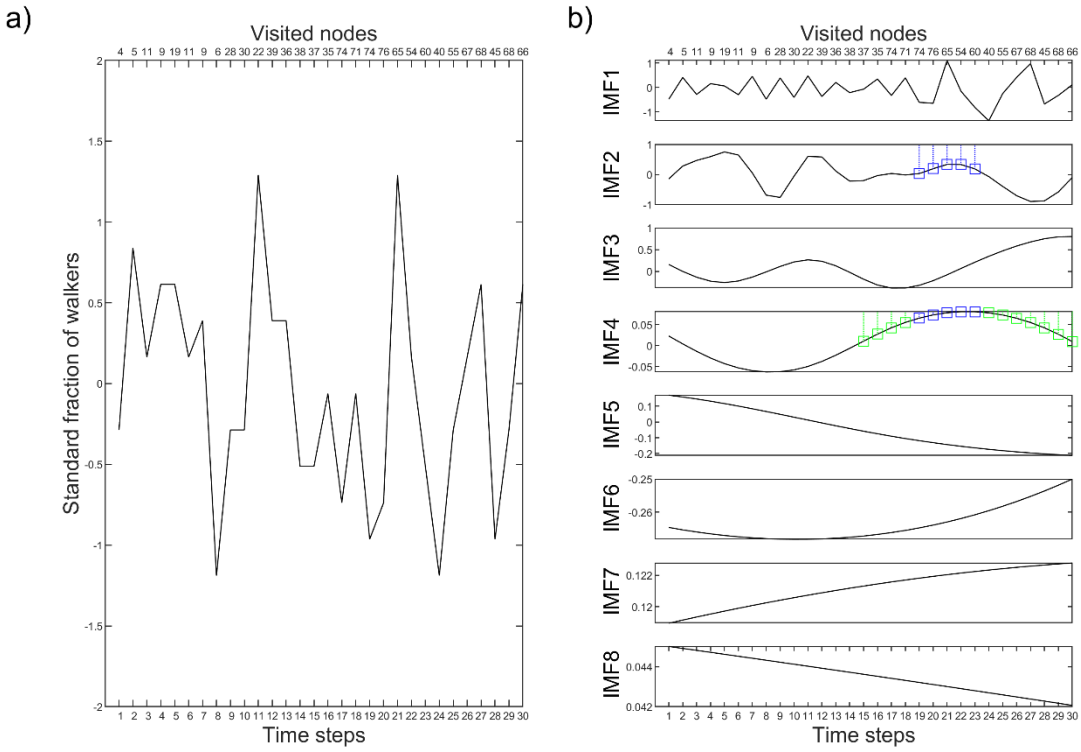
a)
Visited nodes
Standard fraction of walkers
Time steps
b)
Visited nodes
IMF1
IMF2
IMF3
IMF4
IMF5
IMF6
IMF7
IMF8
Time steps

Fig. 1.

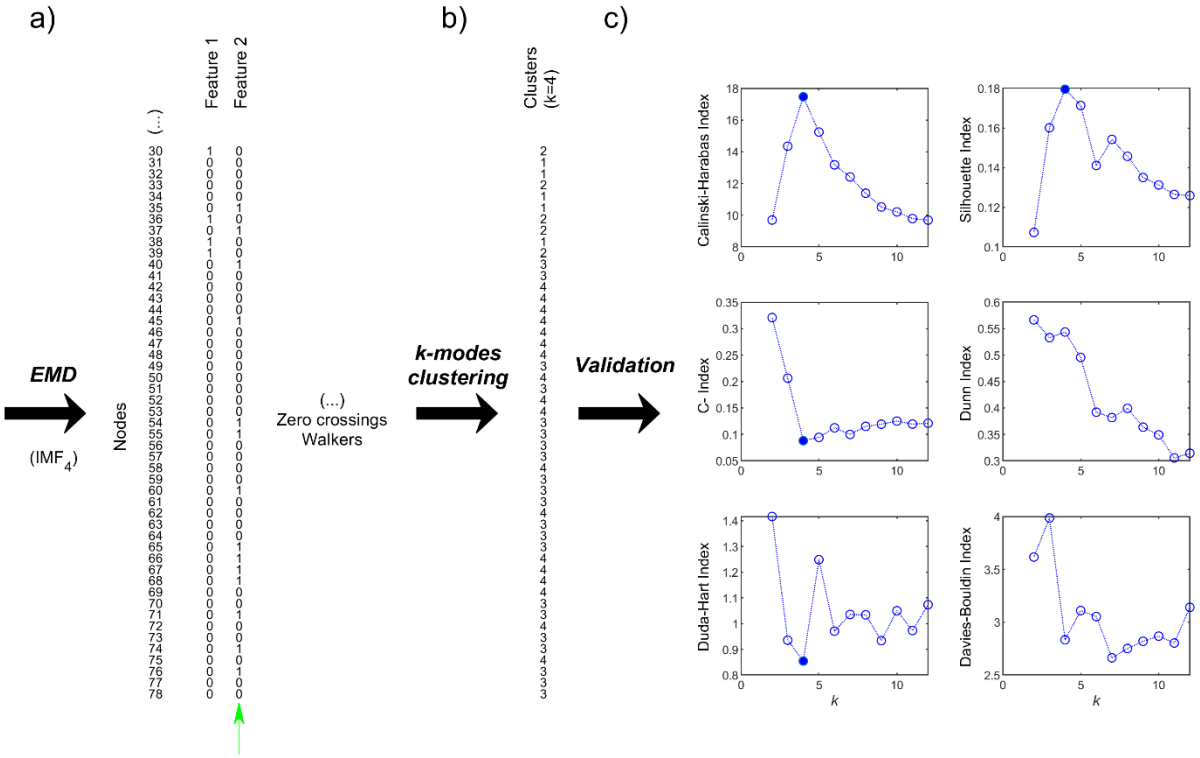
a)
b)
c)
EMD
(IMF4)
Nodes
(...)
Feature 1
Feature 2
(...)
Zero crossings
Walkers
k-modes
clustering
Clusters
(k=4)
Validation
Calinski-Harabas Index
Silhouette Index
C- Index
Dunn Index
Duda-Hart Index
Davies-Bouldin Index
k
k

Fig. 2.

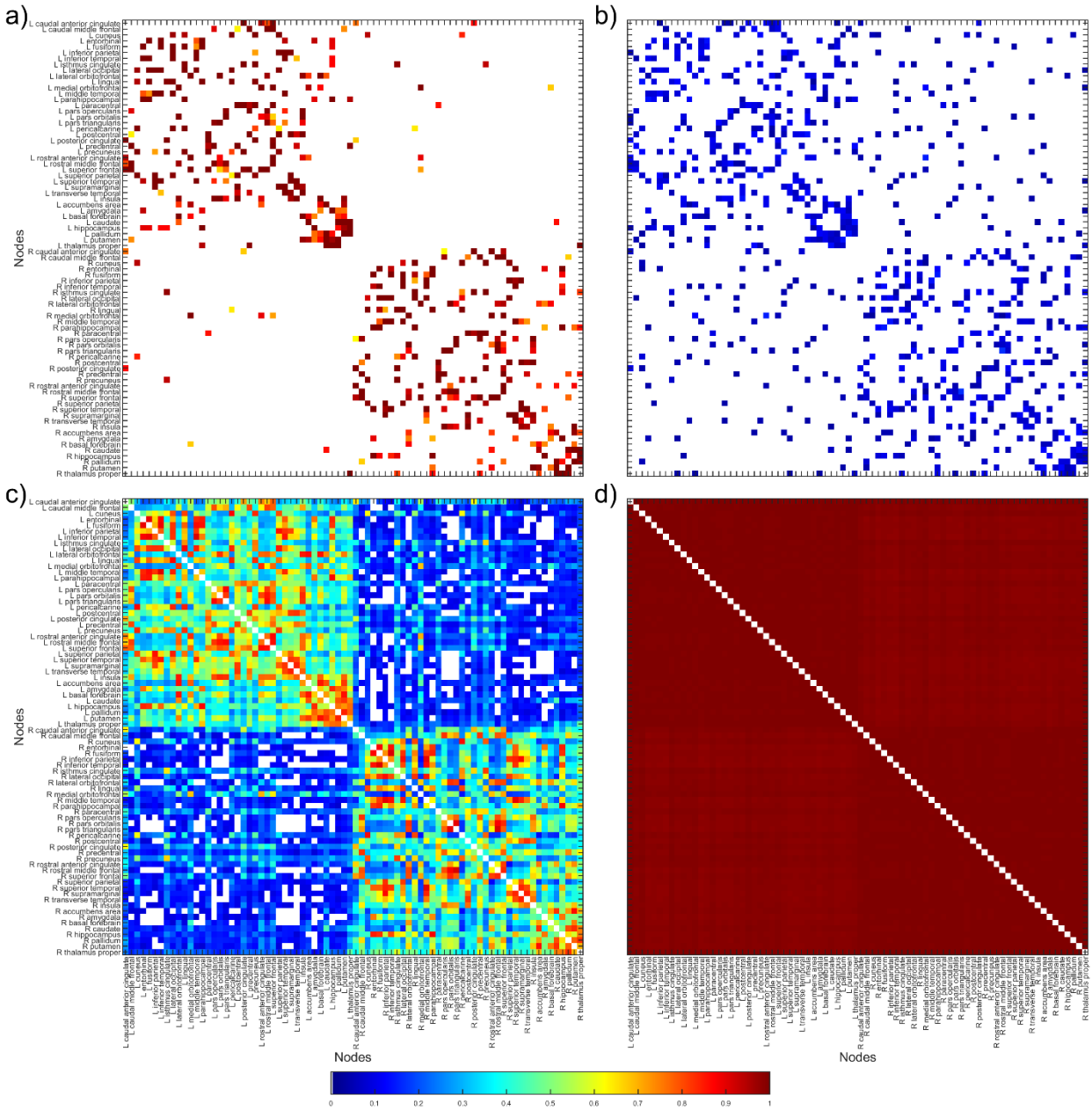

Fig. 3.

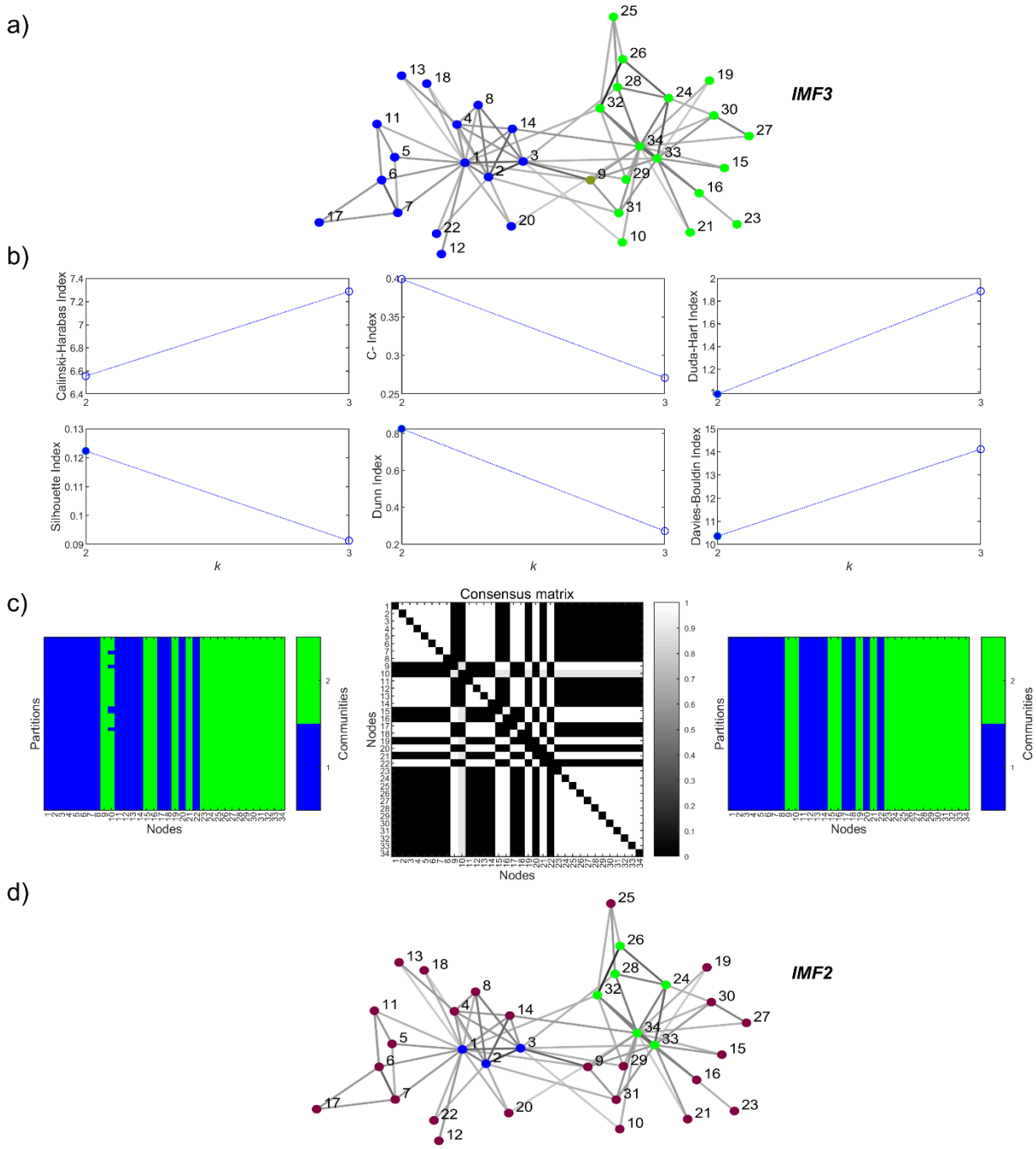
a)
IMF3
b)
Calinski-Harabas Index
C- Index
Duda-Hart Index
Silhouette Index
Dunn Index
Davies-Bouldin Index
k
c)
Consensus matrix
Partitions
Nodes
Communities
d)
IMF2

Fig. 4.

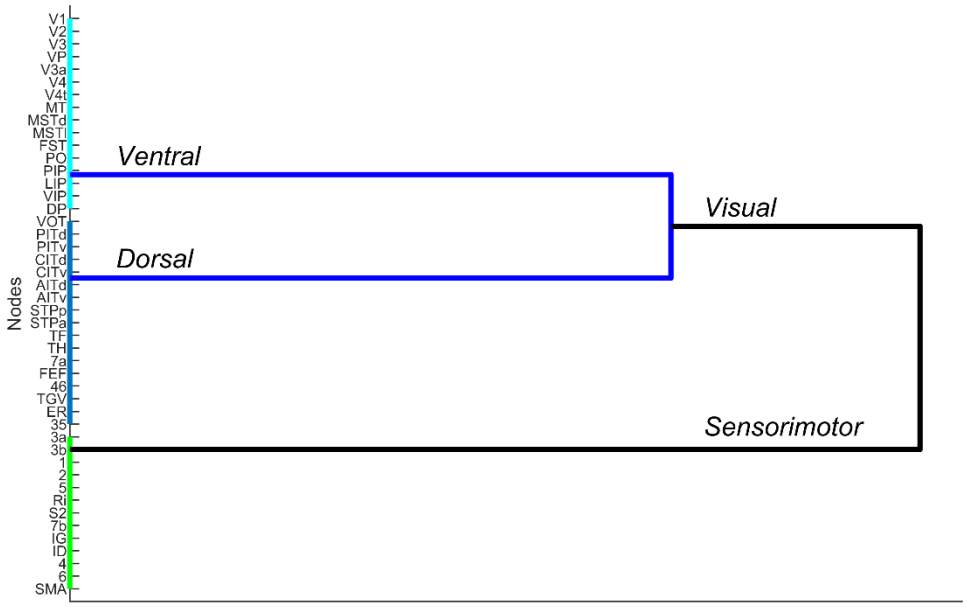
Nodes
V1
V2
V3
VP
V3a
V4
V4t
MT
MSTd
MSTl
FST
PO
PIP
LIP
VIP
DP
VOT
PITd
PITv
CITd
CITv
AITd
AITv
STPp
STPa
TF
TH
7a
FEF
46
TGV
ER
35
3a
3b
1
2
5
Ri
S2
7b
IG
ID
4
6
SMA
Ventral
Dorsal
Visual
Sensorimotor

Fig. 5.

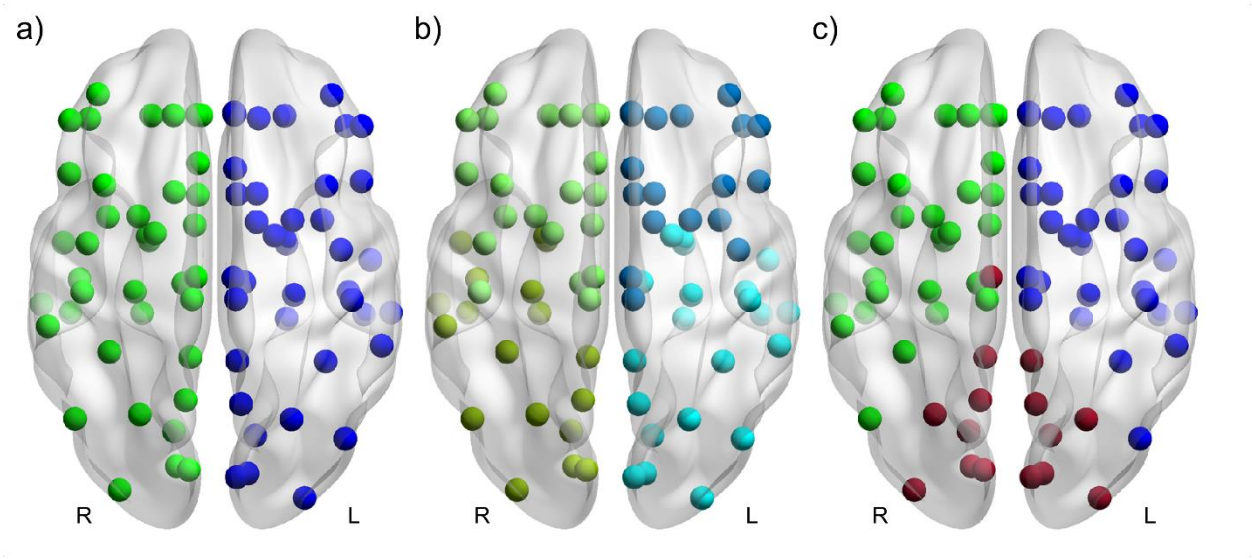

Fig. 6.

Appendix A. Supplementary material

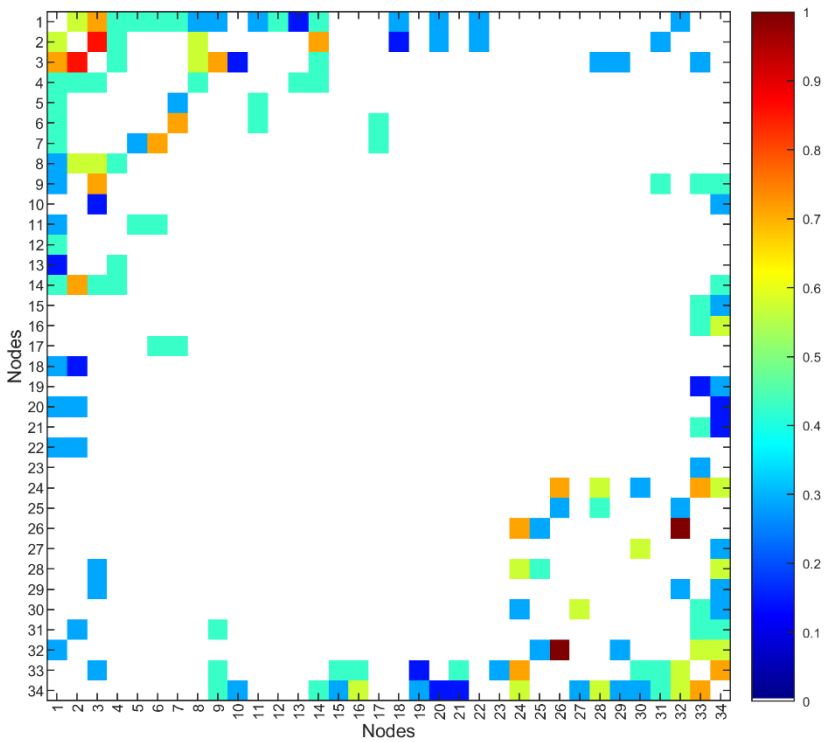

Fig. S1. Matrix of interactions in the karate club. The values indicate the number of contexts in which two members of the club were seen interacting, normalized to the interval [0, 1] (dividing by the maximum value, 7). A number of connections present inconsistencies in Zachary's original paper. When two values in symmetric positions differed, we chose the lowest one. See also: Zachary, W. W. (1977). An Information Flow Model for Conflict and Fission in Small Groups (https://www.jstor.org/stable/3629752).

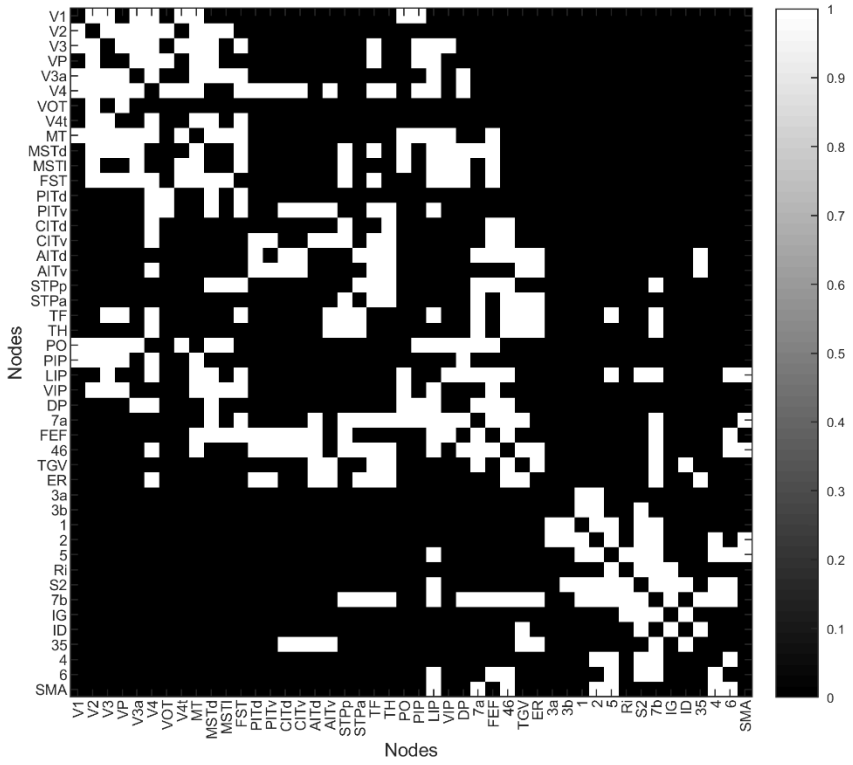

Fig. S2. Connectivity matrix of the macaque neocortex. Pathways that have been identified in tracing studies are marked by 1's. See also: Young, M. P. (1993). The organization of neural systems in the primate cerebral cortex (http://doi.org/10.1098/rspb.1993.0040) and Honey, C. J., Kotter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales (http://doi.org/10.1073/pnas.0701519104).

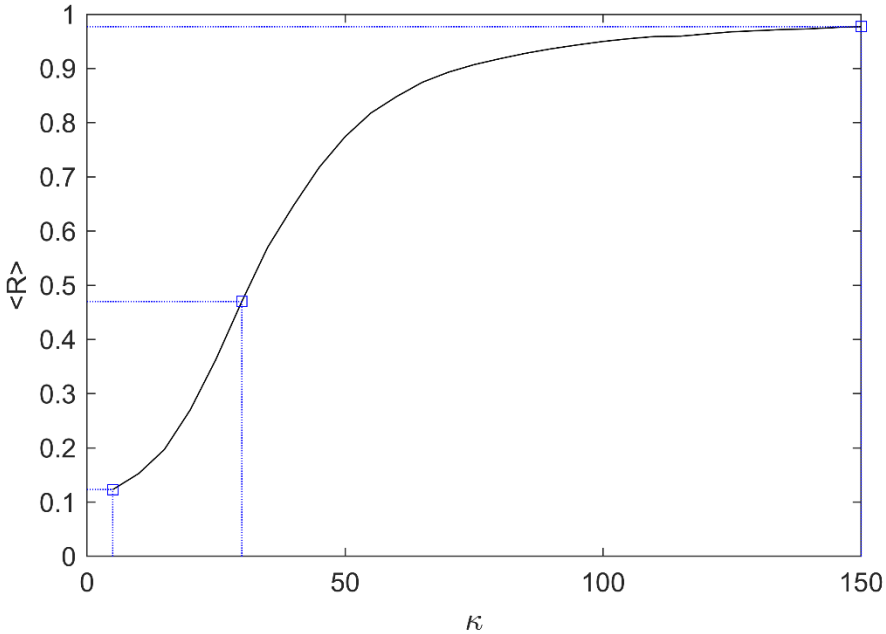

Fig. S3. Grand average of the phase uniformity value, R vs the coupling strength, $\kappa = 0{:}\,5{:}\,150$, of the Kuramoto oscillators. The blue squares represent the values of κ for which the correlation matrices shown in Fig. 3 were calculated.

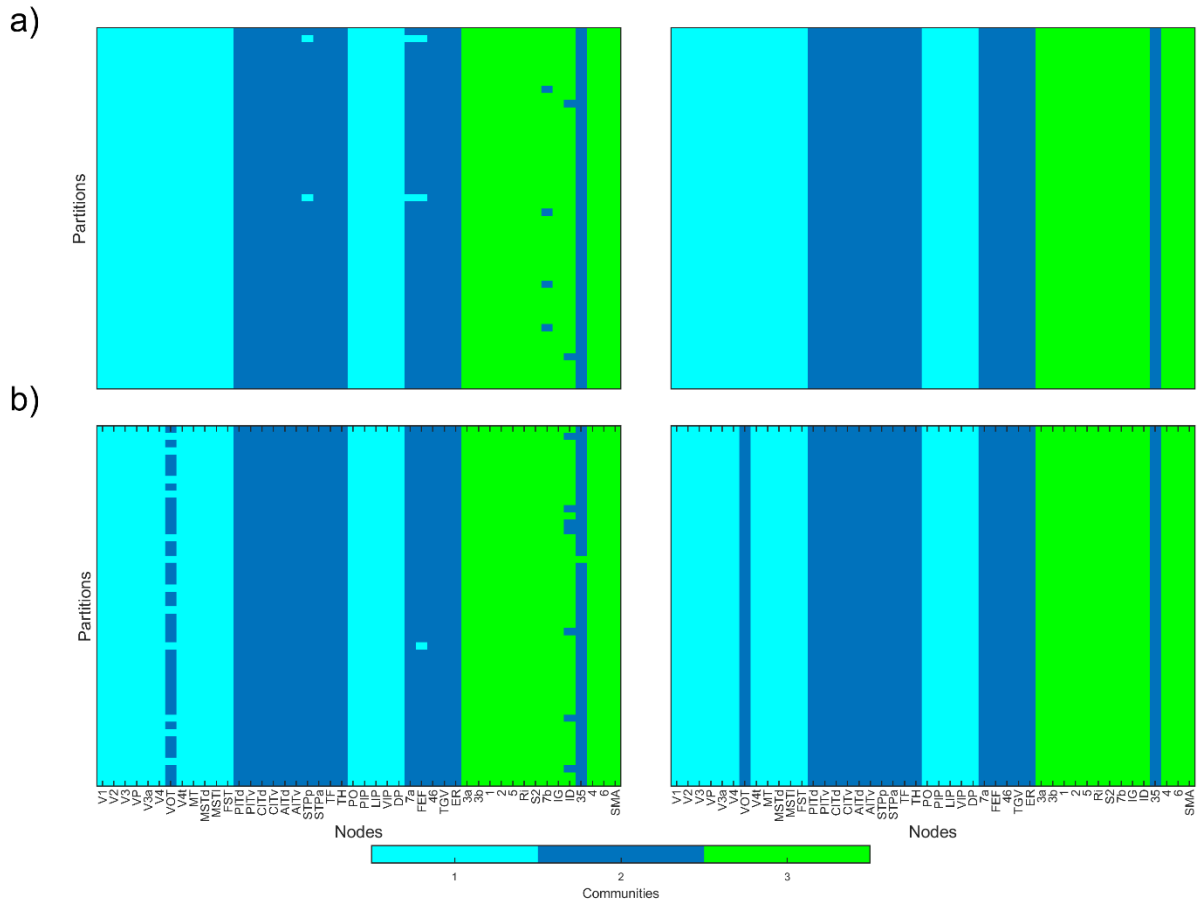

Fig. S4. Consensus clustering for the three communities-partitions of the macaque visual and sensorimotor cortex. a) Partitions with the highest similarity (in terms of adjusted mutual information) obtained through runs of the Louvain community detection algorithm (left). This is for $\gamma = 0.9$. Consensus partitions (right). The consensus partition for the more conventional tuning of the resolution parameter, $\gamma = 1.0$, is identical to the one shown. b) Partitions returned by our clustering algorithm with 50 different sets of random features (left) and their consensus clusters (right).

Left frontal-cingulate-basal	Right frontal-cingulate-basal	Left parietal-occipital-temporal	Right parietal-occipital-temporal	Mixed functional
caudal anterior cingulate	caudal anterior cingulate	cuneus	cuneus	L cuneus
caudal middle frontal	caudal middle frontal	entorhinal	entorhinal	L isthmus cingulate
lateral orbitofrontal	lateral orbitofrontal	fusiform	fusiform	L lateral occipital
medial orbitofrontal	medial orbitofrontal	inferior parietal	inferior parietal	L lingual
paracentral	paracentral	inferior temporal	inferior temporal	L pericalcarine
pars opercularis	pars opercularis	isthmus cingulate	isthmus cingulate	L precuneus
pars orbitalis	pars orbitalis	lateral occipital	lateral occipital	L superior parietal
pars triangularis	pars triangularis	lingual	lingual	R cuneus
posterior cingulate	posterior cingulate	middle temporal	middle temporal	R isthmus cingulate
precentral	precentral	parahippocampal	parahippocampal	R lateral occipital
rostral anterior cingulate	rostral anterior cingulate	pericalcarine	pericalcarine	R lingual
rostral middle frontal	rostral middle frontal	precuneus	precuneus	R pericalcarine
superior frontal	superior frontal	superior parietal	superior parietal	R precuneus
insula	insula	superior temporal	superior temporal	R superior parietal
accumbens area	accumbens area	supramarginal	supramarginal	R posterior cingulate
basal forebrain	basal forebrain	transverse temporal	transverse temporal	
caudate	caudate	amygdala	amygdala	
putamen	putamen	hippocampus	hippocampus	
	postcentral	postcentral		
	pallidum	pallidum		
	thalamus proper	thalamus proper		

Table S1. List of the brain areas per communities in the organization of the human brain network. The first four columns present the arrangement of four-clusters found in the anatomical network (Fig. 6b). The last column lists all the areas that group inter-hemispherically in the functional network with global coupling strength of the Kuramoto model $\kappa = 5$ (Fig. 6c). The table cells of areas spoiling a completely symmetrical organization are shaded. See also: Klein, A., & Tourville, J. (2012). 101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol (http://doi.org/10.3389/fnins.2012.00171) and Iturria-Medina, Y., Sotero, R. C., Toussaint, P. J., Mateos-Perez, J. M., Evans, A. C. (2016). Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis (http://doi.org/10.1038/ncomms11934).

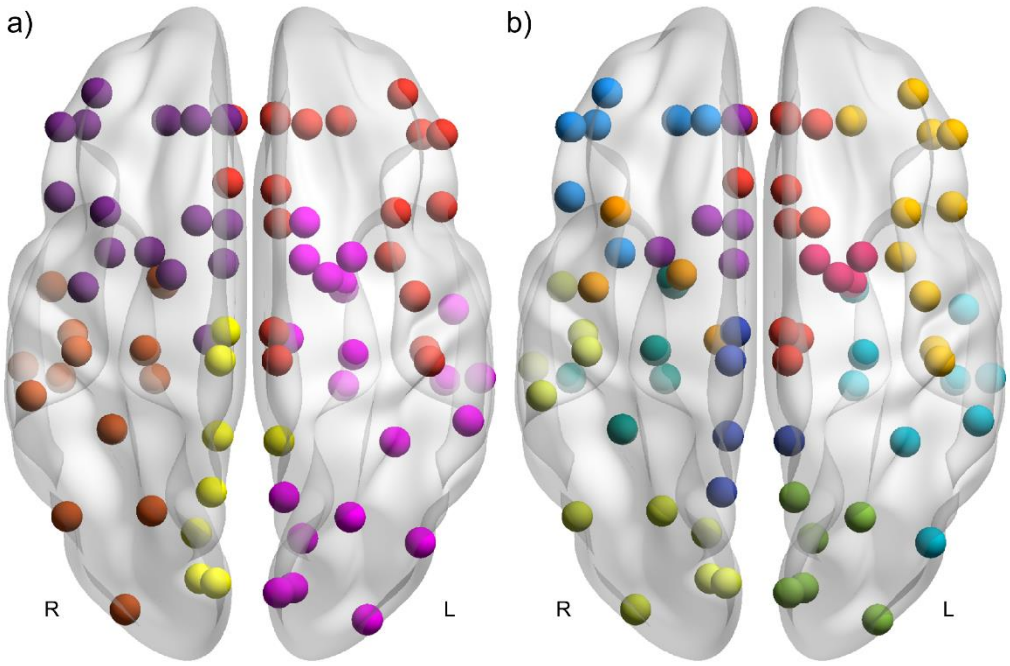

Fig. S5. Communities of the human brain functional network ($\kappa = 5$) obtained through Louvain-modularity maximization. a) Partition resulting from taking the resolution parameter $\gamma = 1.0$. b) Partition with the highest similarity (in terms of adjusted mutual information), existing at $\gamma = 3.0$. Colored nodes correspond to communities and their location, to average coordinates of the brain regions in MNI space.

Visualization of the community structures was achieved by means of BrainNet Viewer (http://doi.org/10.1371/journal.pone.0068910).