Topological Modules of Human Brain Networks Are Anatomically Embedded: Evidence from Modularity Analysis at Multiple Scales

Anvar Kurmukov, Yulia Dodonova, Margarita Burova, Ayagoz Mussabayeva, Dmitry Petrov, Joshua Faskowitz and Leonid E. Zhukov

Abstract Human brain networks show modular organization: cortical regions tend to form densely connected modules with only weak inter-modular connections. However, little is known on whether modular structure of brain networks is reliable in terms of test–retest reproducibility and, most importantly, to what extent these topological modules are anatomically embedded. To address these questions, we use MRI data of the same individuals scanned with an interval of several weeks, reconstruct structural brain networks at multiple scales, and partition them into communities and evaluate similarity of partitions (i) stemming from the test–retest data of the same versus different individuals and (ii) implied by network topology versus anatomy-based grouping of neighboring regions. First, our results demonstrate that modular structure of brain networks is well reproducible in test–retest settings. Second, the results provide evidence of the theoretically well-motivated hypothesis that brain regions neighboring in anatomical space also tend to belong to the same topological modules.

Keywords Brain networks · Modularity · Community structure
Test–retest reliability · Physically embedded networks

A. Kurmukov (✉) · M. Burova · A. Mussabayeva · L. E. Zhukov
Higher School of Economics, National Research University, Moscow, Russia
e-mail: kurmukovai@gmail.com

Y. Dodonova · D. Petrov
Kharkevich Institute for Information Transmission Problems, Moscow, Russia

D. Petrov
Imaging Genetics Center, University of Southern California, Los Angeles, USA

J. Faskowitz
Indiana University, Bloomington, USA

V. A. Kalyagin et al. (eds.), *Computational Aspects and Applications in Large-Scale Networks*, Springer Proceedings in Mathematics & Statistics 247,
https://doi.org/10.1007/978-3-319-96247-4_22

1 Introduction

Brain networks (also called connectomes) are known to have modular organization [1, 2]. This means that cortical brain regions tend to group into modules (communities) with dense anatomical and functional within-module connections and weak inter-module links. Moreover, community structure of anatomical networks were shown to capture important information about brain anatomy and functioning, including that is useful for classifying healthy and pathological brains [3, 4].

However, there are still many important questions to be answered. First, little is known about test–retest reliability of the modular structure of human brain networks. Recently, there has been several papers on reproducibility of brain networks themselves (e.g., [5] and references herein), but test–retest reliability of brain community structures to our knowledge has never become a subject of special evaluation.

Another important question is how topological modules are anatomically co-localized in the brain. Theoretically, it can be expected that brain regions belonging to the same topological module should also be neighbors in physical space. This should be advantageous in terms of minimizing the connection distance and wiring cost within a module. However, empirical work is still strongly needed to understand how modular structure of connectomes is embedded in brain anatomy [1, 2].

In this study, we address both of the above questions. We first question whether community structure of anatomical brain networks is reproducible in the sense that optimal partitions of the connectomes of the same individual scanned with an interval of several weeks show high similarity. To be able to address this question, we analyze appropriate MRI data from the Consortium for Reliability and Reproducibility [6]; we reconstruct connectomes based on different MRI scans of the same subject, find their optimal partitions into communities and evaluate similarity of these partitions. We expect that similarity of partitions stemming from connectomes of the same individual are high, and in particular, higher than that of connectomes of different individuals.

Second, we question whether topological modules of brain networks are at least to some extent anatomically embedded. For each high-resolution brain network, we independently produce two partitions. The first partition is purely topological and only uses information about network connectivity of the respective regions. The second partition use only minor topological information and mainly groups nodes into modules based on their anatomical spacing (and hence their belonging to the same anatomical regions in a low-resolution analysis). We estimate and compare modularity of both types of partitions, and evaluate similarity of the obtained community structures. Details on the procedure are provided in the next section.

2 Data and Experimental Settings

We describe the data and preprocessing steps in this section, and also provide explanation for the terms from network analysis most relevant for our study. Data analysis procedures are described in the next section, each immediately followed by the results obtained on this dataset.

2.1 Dataset

We used a dataset from the Consortium for Reliability and Reproducibility (CoRR, [6]). Data sites within CoRR were chosen due to availability of T1-weighted (T1w) and diffusion-weighted images (DWI) with retest period less than two months as described in [5].

The dataset included 49 participants aged 19–30 years. All subjects were healthy volunteers without any psychiatric or neurological disorders. For each participant, two MRI session were performed with the retest period of about 6 weeks (40.9 ± 4.51 days). Details on the sample and MRI acquisition procedure are available on the website of CoRR project [6].

2.2 Tractography and Brain Parcellation

We followed the same MRI preprocessing steps as described in [5]. We used probabilistic constant solid angle (CSA) approach to reconstruct white matter structures. This method allows for estimating the fiber orientation distribution (FOD) within each voxel. Streamlines reconstructed based on this method gave information about edge weights of brain networks.

To define nodes of brain networks at multiple scales, we followed the procedure similar to that described in [7, 8]. Based on the Lausanne atlas, the cortex of each subject was divided into 68 cortical regions. Then, the obtained regions were subdivided into 1000 small regions so that coverage area of each region was about 1,5 cm^2. To obtain parcellation at multiple scales, these 1000 small regions were additionally combined into 448, 219, and 114 nodes; in each step, neighboring regions of interest were manually combined into larger regions. This successive grouping thus gave us a hierarchical decomposition between 68 and 1000 brain regions; importantly, smaller regions in high-resolution parcellations were strictly embedded into larger low-resolution regions.

We counted the number of streamlines having endpoints in each pair of labels for each parcellation and used them as edge weights in each constructed network. We, thus, obtained sets of 98 brain networks (two for each of 49 individuals) at five different scales.

2.3 Notation

We use the following definitions:

Weighted network. Let $\mathbb{G} = (V, \mathbb{E}, \mathbb{W})$ be an undirected graph (network), where V $=\{1, \ldots, n\}$ is a set of nodes, $\mathbb{E}$ is a set of edges of the graph, $A = \{a_{i,j}\}_{i,j=1}^{n}$ is its adjacency matrix and $\mathbb{W} = \{w_{i,j}\}_{i,j=1}^{n}$ is a connectivity matrix where each element represents a weight of the respective edge.

Modularity is defined as follows:

$$Q = \frac{1}{s} \sum_{i,j=1}^{n} \left(w_{i,j} - \frac{s_i s_j}{s} \right) \delta_{i,j},$$

where $s_i = \sum_{j=1}^{n} w_j$ is the weighted degree of node i, $s = \sum_{i=1}^{n} s_i$ is the sum of weights of the whole graph, and δ_{ij} equals 1, if nodes i and j belongs to the same community and 0 otherwise. Note that, modularity is always defined with respect to some network partition.

Rand Index (**RI**) [9] is used to measure the similarity of two partitions U, V:

$$RI(U, V) = \frac{a + d}{a + b + c + d},$$

where a is a number of pairs with objects placed in the same community in both U and V, b is a number of pairs with objects placed in different communities in U and V, c is a number of pairs with objects placed in the same community in V and in different communities in U, and d is a number of pairs with objects placed in the same community in U and in different communities in V.

Here, we use an adjusted version of RI [10]. Similarly to RI, it takes values between 0 and 1, with values closer to 1 for more similar partitions.

3 Experiments and Results

3.1 Modular Structure of Brain Networks and Its Reliability

We first analyzed a set of connectomes with 68 nodes. This was the only reasonable network size for which globally optimal modularity score and the corresponding community structure could be found.

First, all networks showed strong modularity, with sample average modularity score of 0.568, standard deviation 0.010. Figure 1 shows an example connectome of this size, with nodes located in their physical 3D coordinates and colored according to their best partition into communities.

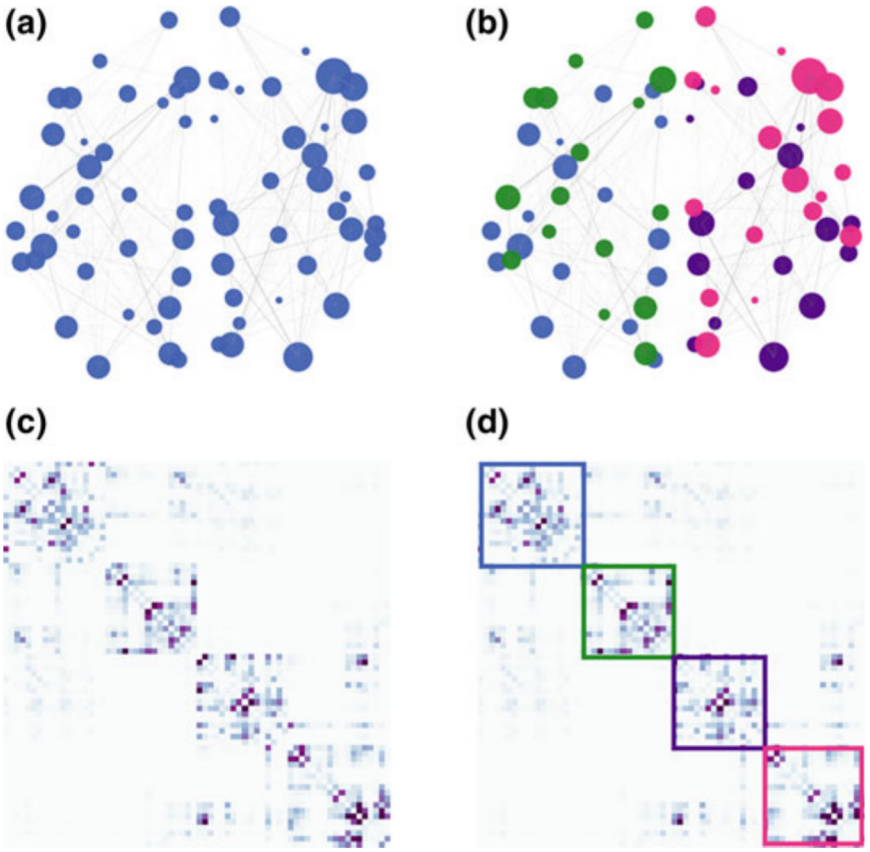

Fig. 1 Example brain network and its optimal modularity structure. **a** Brain network with 68 nodes shown in their physical 3D coordinates (axial view). **b** The same brain network with 68 nodes, node colors show its optimal modular structure. **c** Connectivity matrix of the same network. **d** Connectivity matrix with four optimal modules shown by colored squares

Although the number of modules was not constrained during partitioning, the algorithm stopped at four-module community structure in all networks. All modules were intra-hemispheric, with two of them located in the left hemisphere and two in the right one (there was the only exception: in one network, there was a single node attributed to a module from a "wrong" hemisphere).

We next addressed a question of reliability of the obtained community structures. For each individual, we had two MRI scans and thus two connectomes and two corresponding partitions. We estimated similarity of these two partitions by computing ARI. We also computed ARI for all pairs of partitions from different individuals. For the same-individual pairs, average ARI value was 0.976, with a standard deviation 0.037 across individuals; median value was strictly 1, thus indicating for most connectomes identical community structures in test and retest connectomes. For interindividual pairs, ARI values were still high, but significantly lower than those obtained for the same-individual pairs. Average ARI value for modular structures of different participants was 0.920, with a standard deviation of 0.061 (median value 0.923).

Finally, we performed additional analysis to evaluate whether the same or very close modular structures could be found by an algorithm that does not necessarily stop at globally maximal modularity value (compared to the respective optimal partitioning). We used Louvain algorithm for this analysis. For each network of size 68, we found its best Louvain partition in addition to the optimal one. Modularity values obtained with Louvain algorithm were exactly the same or close to optimal ones in most networks, with the sample average Louvain modularity score of 0.568 and standard deviation 0.010.

However, close modularity scores do not necessarily indicate similarity in the corresponding partitions. To compare optimal modularity structure to that obtained by Louvain algorithm, we again used the ARI. We computed pair-wise similarity score between partitions obtained with two algorithms for each connectome. Mean sample ARI value was 0.895, with a standard deviation of 0.095 (recall that ARI value for two identical partitions is 1). High ARI values indicate that Louvain algorithm revealed modular structures identical or very close to those corresponding to globally maximal modularity. This allows us to use Louvain algorithm instead of global optimization in further analyzes, a step needed to evaluate modular structure of the larger networks.

3.2 Anatomical Embedding of Topological Modules

We now consider sets of 98 connectomes reconstructed at other scales, with 114, 219, 448, and 1000 nodes (note that information on whether connectomes belong to the same or different individuals is not relevant in what follows).

Figure 2 shows an example of 1000-node connectome in its physical 3D coordinates, with the corresponding 68-node connectome put in the same space. Each node in a 1000-node connectome corresponds to some node in a 68-node connectome; this means that for the 1000-node connectome, we already have a partitioning into 68 communities that is strictly anatomically determined.

By addressing a question of whether topological modules are anatomically embedded, we, in fact, aim to check that optimal topological partition of a network is close to that defined based on anatomical neighboring. The most straightforward approach to checking this hypothesis would be to find the best possible topological partition of a 1000-node network into 68 communities and compare this topological partition to that defined a priori from anatomical parcellation. However, this approach does not work in practice because the number of communities is too large and hence the obtained topological partitions are too far from being optimal (to compare with, even the first iteration of the Louvain algorithm reveals on average only 22 communities in a 1000-node network).

We thus propose the following procedure. For a 1000-node network, we find its best topological partition using Louvain algorithm. On the other hand, we consider a corresponding 68-node network (obtained from the same MRI scan) and its optimal partition into 4 modules described in the previous section. We map this four-module

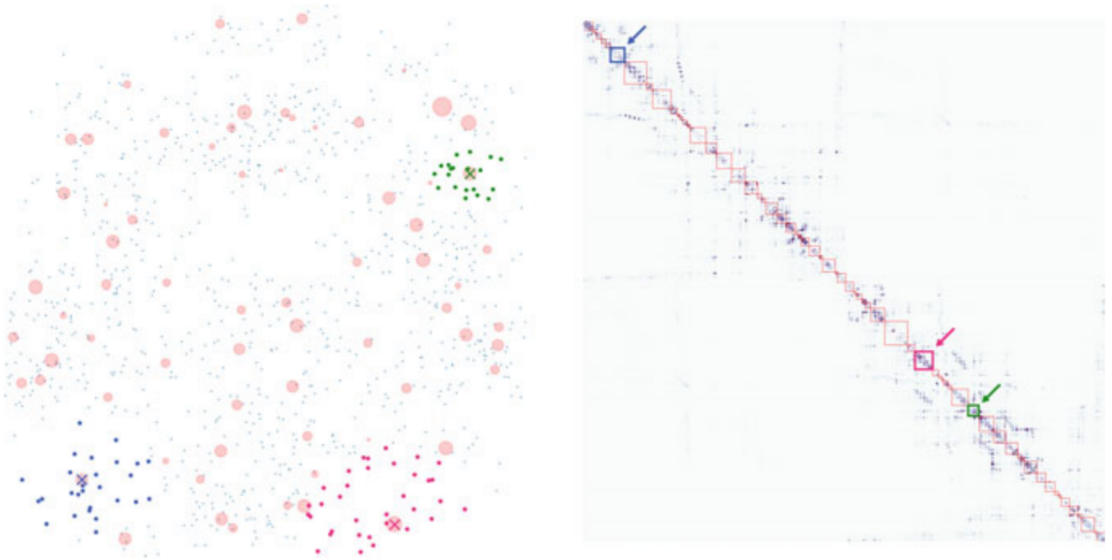

Fig. 2 **Left** Brain networks with 68 (light red circles) and 1000 (black dots) nodes shown in the same physical space (axial view); 3 nodes of the 68-node network are colored with blue, pink, and green, and their corresponding nodes in a 1000-node network are colored accordingly. **Right** Connectivity matrix of the same 1000-node network; light red boxes show how the nodes of this network are grouped to produce nodes of a 68-node network, and the same 3 nodes are additionally colored

optimal partition from a low-resolution network to a high-resolution network so that each node in a 1000-node network inherits its community membership from the respective parental node in the 68-node network.

This mapped partition of the 1000-node network certainly carries some information about network topology because we used topological modules of a low-resolution network for mapping. But still, this mapped partition is substantially determined by the anatomy. It, in fact, ignores topology of a 1000-node network and mainly puts nodes to a module based on whether or not the respective small cortex areas belong to a given larger region in the anatomical parcellation.

For each 1000-node network, we thus know its best Louvain partition and the corresponding modularity score, and its mapped partition and the corresponding modularity score. We finally evaluate similarity between these two partitions by computing ARI.

The above procedure was repeated for networks of size 114, 219, and 448 (with 68-node networks always being a reference for obtaining mapped partitions). Table 1 reports the results.

First, networks of all sizes were again confirmed to be highly modular. Network modularity steadily increased with increasing resolution (network size).

Second, modularity scores of the anatomically mapped partitions were only marginally lower than those obtained for the best topological partitions. In other words, ignoring actual connections between nodes in a high-resolution network and simply coloring them based on anatomical membership of the respective cortex areas did not result in a dramatic drop in modularity values.

Table 1 Mean values (standard deviations) of the modularity scores obtained for each network size by Louvain partitioning algorithm (top row) and by mapping the partitioning of the corresponding low-resolution network (middle row) and the Rand index indicating similarity between these two partitions (bottom row)

	114 nodes	219 nodes	448 nodes	1000 nodes
Modularity, best partition	0.598 (0.011)	0.631 (0.010)	0.665 (0.011)	0.691 (0.011)
Modularity, mapped partition	0.592 (0.010)	0.615 (0.009)	0.629 (0.009)	0.642 (0.009)
ARI for two partitions	0.762 (0.114)	0.652 (0.099)	0.561 (0.068)	0.525 (0.051)

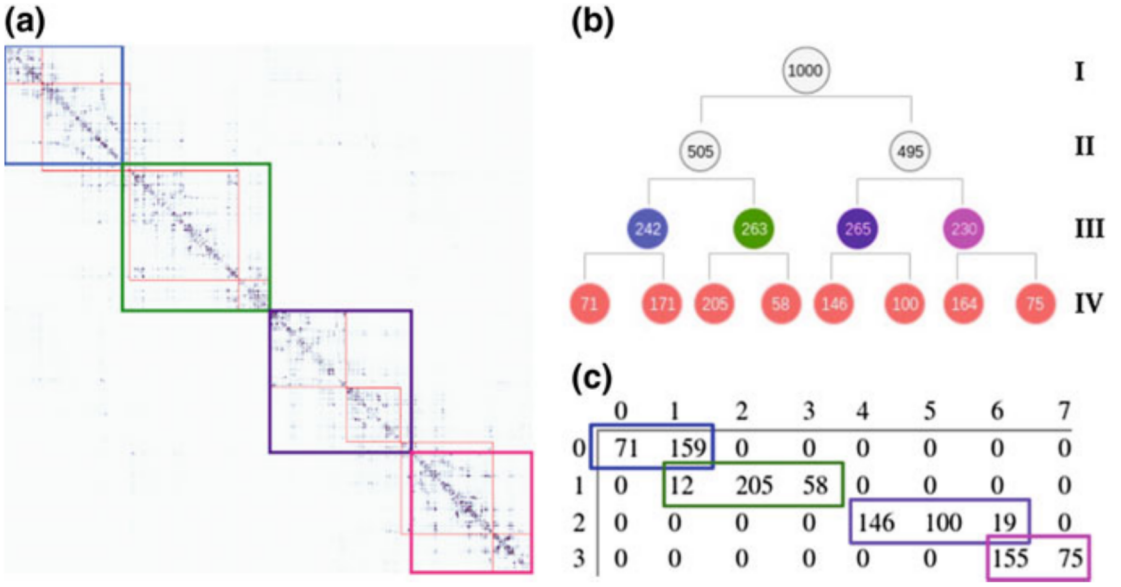

Fig. 3 Different ways to illustrate that topological modules of high-resolution networks are nested within the anatomically mapped network structure inherited from the low-resolution parental network. **a** Connectivity matrix of the high-resolution network with eight best-partitioned topological modules are shown with light red squares and four anatomically mapped modules inherited from the low-resolution network shown with blue, green, violet, and pink squares; the former partition is almost strictly embedded into the latter. **b** Hierarchical tree with four levels: I—whole 1000-node connectome, II—two hemispheres, III—four modules inherited from the anatomically mapped partition of the low-resolution network, and IV—eight modules obtained by Louvain partitioning of the high-resolution network. **c** Cross tabulation (contingency table) of co-occurrence of the high-resolution network labels in eight modules of the best Louvain partition (horizontal) and four modules of the anatomically mapped partition (vertical); again, best-partitioned topological modules are almost strictly submodules of the anatomically mapped partition

Third, these anatomically mapped partitions were highly similar to the best topological partitions of the respective networks, as indicated by moderate to high ARI values.

Importantly, computation of ARI does not require that partitions have the same number of modules. The obtained moderate to high ARI values also not necessarily

indicate that best-partitioned high- resolution networks resembled the four-modular structure of the anatomically mapped partitions and the parental low-resolution networks. High-resolution networks could show larger number of modules, but what was important for ARI not to drop down was that these smaller topological modules were hierarchically nested within the four-module anatomically mapped structure.

Figure 3 illustrates this idea for high-resolution 1000-node networks. Louvain algorithm revealed eight-module community structure of this network; however, these topological modules were almost embedded into four-module structure obtained by anatomical community mapping of the parental low-resolution network.

4 Conclusions

We analyzed modular structure of anatomical brain networks defined at different scales, from low-resolution networks that represented connectivity of atlas cortical regions to high-resolution 1000-node networks. We worked with dataset from the Consortium for Reliability and Reproducibility that included data obtained from the same individuals with a time interval of several weeks, a step that additionally allowed us to evaluate test–retest reliability of the community structure of brain networks.

First, our experiments confirm that structural brain networks are highly modular. Brain regions tend to form small number of densely connected modules, with the number of topologically defined modules varying from four in the low-resolution networks based on the Lausanne anatomical atlas to eight in the highest resolution 1000-node networks.

Second, we proposed to use a measure of similarity between partitions (Rand index) for evaluating test–retest reliability of the community structure of brain networks. Our experiment confirmed that modular organization of anatomical brain networks is highly reliable in the sense that it almost perfectly resembles in networks reconstructed from different MRI scans of the same individual taken with an interval of several weeks.

Third, and most important, for the high-resolution networks, we considered two approaches to partitioning brain regions into communities. The first one was purely topological and revealed optimal modules only based on network connections between the nodes. The second one was largely anatomical because nodes were put into communities based on the fact that they were anatomically neighboring and belong to the same cortical region in the anatomical atlas (and hence to the same parental node in a low-resolution network). We demonstrated that modularity with respect to these latter partitions were still very high, and only slightly lower than modularity estimated with respect to topologically optimal partitions.

Moreover, we demonstrated that similarity between topologically optimal and anatomically implied partitions was very high. Topological modules largely resembled anatomical grouping of neighboring cortical regions. By using multi-scale analysis and network algorithms of partitioning and comparing partitions, we found new

evidence in support of the theoretically well-motivated hypothesis that brain regions neighboring in anatomical space also tend to belong to the same topological modules.

Acknowledgements The publication was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2017 (grant 16-05-0050) and by the Russian Academic Excellence Project "5-100".

References

1. Meunier, D., Renaud, L., Bullmore, E.T.: Modular and hierarchically modular organization of brain networks. Front. Neurosci. **4** (2010)
2. Meunier, D., Lambiotte, R., Fornito, A., Ersche, K.D., Bullmore, E.T.: Hierarchical modularity in human brain functional networks. Front. Neuroinform. **3** (2009)
3. Kurmukov, A., Dodonova, Y., Zhukov, L.: Classification of normal and pathological brain networks based on similarity in graph partitions. In: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)
4. Kurmukov, A., Dodonova, Y., Zhukov, L.: Machine learning application to human brain network studies: a kernel approach. Models, Algorithms, and Technologies for Network Analysis. NET 2016. Springer Proceedings in Mathematics and Statistics, vol. 197 (2017)
5. Petrov, D., Ivanov, A., Faskowitz, J., Gutman, B., Moyer, D., Villalon, J., Thompson, P.: Evaluating 35 methods to generate structural connectomes using pairwise classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (2017)
6. He, Y.: Connectivity-based brain imaging research database (c-bird) at Beijing normal university
7. Cammoun, L., Gigandet, X., Meskaldji, D., Thiran, J.P., Sporns, O., Do, K.Q., Hagmann, P.: Mapping the human connectome at multiple scales with diffusion spectrum MRI. J. Neurosci. Methods **203**(2) (2012)
8. Hagmann, P., Kurant, M., Gigandet, X.: Mapping human whole-brain structural networks with diffusion MRI. PLoS ONE **2**(7) (2007)
9. Rand, W.M.: Objective criteria for the evaluation of clustering methods. J. Am. Stat. Assoc. **66**(336), 846–850 (1971)
10. Vinh, N.X., Julien, E., Bailey, J.: Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. J. Mach. Learn. Res. **11** (2010)