Compact and Informative Representation of Functional Connectivity for Predictive Modeling

Raif M. Rustamov[1], David Romano[2], Allan L. Reiss[2], and Leonidas J. Guibas[1]

[1] Computer Science Department, Stanford University, Stanford, CA
[2] Center for Interdisciplinary Brain Sciences Research, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA

Abstract. Resting state functional connectivity holds great potential for diagnostic prediction of neurological and psychiatric illness. This paper introduces a compact and information-rich representation of connectivity that is geared directly towards predictive modeling. Our representation does not require a priori identification of localized regions of interest, yet provides a mechanism for interpretation of classifier weights. Experiments confirm increased accuracy associated with our representation and yield interpretations consistent with known physiology.

1 Introduction

Resting state functional magnetic resonance imaging (fMRI), in conjunction with multivariate pattern analyses, holds great promise for diagnostic prediction of neurological and psychiatric illness [2]. For accurate predictive modeling, it is necessary to have compact representations of functional connectivity. Such representations are usually obtained by a judicious choice of nodes for assembling the correlation matrix/connectivity network. Parcellation schemes based on anatomical and/or functional features are used to yield regions of interest (ROIs) that are identified as nodes of the network [10]. For network-based analyses of anatomical and functional connectivity it is crucial to use nodes derived from spatially localized ROIs that are functionally and biologically meaningful [5]. On the other hand, for predictive modeling it is desirable to maximize the information content of the reduced representation. For example, a priori identification of ROIs may obscure more subtle and complex phenomena that cross the boundaries of ROIs and thus may lead to suboptimal prediction accuracy.

The goal of our work is to introduce a compact and informative representation of functional connectivity that is geared directly towards predictive modeling and does not require a priori identification of localized ROIs. Our approach is based on the observation that a localized ROI can be fully captured by its indicator function, which is simply a particular type of spatial map (i.e., real-valued functions on the collection of voxels). The set of all possible spatial maps is a very high-dimensional vector space, but since fMRI data is already subject to smoothing, both in its acquisition and preprocessing, we choose to restrict to the subspace of spatial maps exhibiting some smoothness. This subspace can

be approximated by the span of low-frequency eigenvectors of an appropriate graph Laplacian (in the same way as Fourier basis provides building blocks for time signals, the Laplacian eigenvectors provide a basis for spatial signals). Therefore, we build our representation from the spatial maps associated to low-frequency eigenvectors, which replace ROIs as nodes of network. Namely, these distributed and overlapping spatial maps are combined with blood oxygenation level dependent (BOLD) signal to obtain the corresponding representative time series. The matrix of correlations between these time series are then computed, giving the sought representation of connectivity.

The proposed approach has a number of advantages. First, it leads to a compact representation of functional connectivity that is hierarchical. Indeed, the Laplacian eigenvectors are naturally ordered by their smoothness and a number of smoother eigenvectors can be retained to obtain a connectivity matrix of desired size. Second, the proposed representation is informative in that it allows approximate reconstruction of correlations between any pair of traditional ROIs, even if these ROIs were not specified a priori. This explains why our approach is effective: in a sense, traditional ROIs are subsumed by our approach and a classifier applied to our representation is able to learn any information that could be extracted from traditional ROIs. Finally, our representation is interpretable—it provides a simple mechanism for mapping the weights learned by linear classifiers back to the brain, allowing a detailed understanding of the predictive model.

2 Methods

Let X be the $n_{\mathrm{TR}} \times n_{\mathrm{vox}}$ matrix of the voxel-wise z-scored corrected BOLD signal acquired in n_{vox} grey matter voxels over n_{TR} time points. The matrix $C = X^\top X$ captures the correlations between all pairs of voxels, but is impractically large. Our goal is to obtain a computationally tractable reduced representation of C.

Motivation. To motivate our representation, let us analyze the commonly used ROI-based approaches. For i-th ROI, consider an $n_{\mathrm{vox}} \times 1$ column vector whose entries correspond to voxels, with an entry of 1 for voxels that belongs to the ROI, and 0 otherwise. Let us scale this vector so that its entries sum to 1, and denote this normalized vector by ϕ^i; this is the indicator vector of the ROI. Note that the product $X\phi^i$ gives the averaged BOLD time series over the i-th ROI, and so the product $(X\phi^i)^\top X\phi^j = \phi^{i\top} X^\top X \phi^j = \phi^{i\top} C \phi^j$ gives the un-normalized correlation between i-th and j-th ROIs. Denote by Φ the $n_{\mathrm{vox}} \times n_{\mathrm{ROI}}$ matrix whose i-th column is the vector ϕ^i. Now the matrix $\Phi^\top C \Phi$ can be seen to be an $n_{\mathrm{ROI}} \times n_{\mathrm{ROI}}$ matrix of un-normalized correlations between all the ROIs.

This process can be interpreted in the following manner. The large matrix C defines a dot product (namely, $\langle f, g \rangle = f^\top C g$) over the space of all spatial maps. The goal is to capture information about this dot product, and so about connectivity, in a compact manner. The ROI based approach achieves this by restricting the dot product to a much smaller subspace of spatial maps. Namely, the subspace in question is $\mathrm{span}\{\phi^1, \phi^2, ..., \phi^{n_{\mathrm{ROI}}}\}$, and the matrix of the restricted dot product is precisely the matrix $\Phi^\top C \Phi$ above.

Proposed Approach. In the light of the discussion above, the gist of our approach is to choose a different subspace of spatial maps for restricting the dot product. For multivariate pattern analyses it is desirable to maximize the information content of the reduced representation and this can be achieved by choosing a subspace that offers strong approximation properties. Since spatial smoothing is applied to BOLD signal during preprocessing, it is natural to restrict the dot product to a subspace of smooth spatial maps.

We propose to use the subspace spanned by the low spatial frequency eigenvectors of an appropriate graph Laplacian. Here, we consider the grey matter template voxels as nodes of a graph, and introduce an edge between voxels that share a face. By solving the eigenvalue problem $L\psi^i = \lambda_i \psi^i$, where L is the $n_{\text{vox}} \times n_{\text{vox}}$ graph Laplacian [4], we obtain an orthonormal basis $\{\psi^i\}_{i=1}^{n_{\text{vox}}}$. Assuming the eigenvectors are ordered by increasing eigenvalue, it can be shown (c.f. [1] Sec. 5.2) that in some precise mathematical sense the eigenvector ψ^1 is the smoothest spatial map, ψ^2 the next smoothest, and so on. Since we would like to obtain low-dimensional subspace of smooth spatial maps, we consider the subspace spanned by the first n_{ev} eigenvectors. Denoting by Ψ an $n_{\text{vox}} \times n_{\text{ev}}$ matrix whose i-th column is ψ^i, our reduced representation of the voxel-wise connectivity matrix C is given by $D = \Psi^\top C \Psi$ which is a matrix of size $n_{\text{ev}} \times n_{\text{ev}}$. In practice, we first compute $X\psi^i$ for $i = 1, 2, ..., n_{\text{ev}}$, which gives the BOLD time series weighted by the spatial map ψ^i, and then we set the entries $D_{ij} = (X\psi^i)^\top X\psi^j$.

Informativeness. The proposed representation is informative in the sense that it allows approximate reconstruction of correlations between two given ROIs. Indeed, let ϕ^1 and ϕ^2 be the indicator vectors of the two ROIs. As discussed in *Motivation*, the un-normalized correlation between the ROIs is given by $\phi^{1\top} C \phi^2$. Let us write the vectors ϕ^1 and ϕ^2 in terms of the eigenvector basis: $\phi^1 = \alpha_1\psi^1 + \alpha_2\psi^2 + \cdots = \Psi\alpha$ and $\phi^2 = \beta_1\psi^1 + \beta_2\psi^2 + \cdots = \Psi\beta$; here α (resp. β) is a column-vector with entries α_i (resp. β_i). Now we have, $\phi^{1\top} C \phi^2 = (\Psi\alpha)^\top C(\Psi\beta) = \alpha^\top \Psi^\top C \Psi \beta = \alpha^\top D \beta$, where we used $D = \Psi^\top C \Psi$. We can compute the correlations between the ROIs by appropriate normalization, namely:

$$\text{corr}(ROI_1, ROI_2) = \frac{\alpha^\top D \beta}{(\alpha^\top D \alpha)^{1/2} (\beta^\top D \beta)^{1/2}} \tag{1}$$

Note that since we have actually truncated the eigenvector basis, this equality holds only approximately. This approximate reconstruction property shows that despite its compactness, our representation is able to capture connectivity in a richly informative manner, which is crucial for accurate predictive modeling.

Interpretability. The proposed representation is interpretable in the sense that the feature weightings obtained from linear classification algorithms can be mapped back to locations in the brain. Let $\hat{y} = \text{sign}(b + \sum_{i,j} W_{ij} D_{ij})$ be the prediction model where W_{ij} are the weights and b is the bias; this form can capture commonly used linear classifiers such as Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM). Using $D = \Psi^\top C \Psi$, we can rewrite $\sum_{i,j} W_{ij} D_{ij} = \sum W_{ij} \sum_{p,q} \Psi_{ip} C_{pq} \Psi_{jq} = \sum_{p,q} C_{pq} \sum_{i,j} W_{ij} \Psi_{ip} \Psi_{jq}$; here i, j run over eigenvector indices, and p, q run over voxels. Thus, we can say that the

coefficient of C_{pq} given by $R_{pq} = \sum_{i,j} W_{ij}\Psi_{ip}\Psi_{jq}$ captures the contribution of connectivity between voxels p and q. It is easy to see that $R = \Psi W \Psi^\top$.

The matrix R is impractically big, but from it, we can derive a quantity that measures the importance of the connectivity of a given voxel for classification. For a voxel p, consider quantity $\omega_p = \sum_q R_{pq}^2$; note that squaring the coefficients is appropriate here, as we would like to amplify importance of voxels whose connectivity coefficients are larger. Letting ω be the vector with entries ω_p, we have:

$$\omega = \mathrm{diag}(R^\top R) = \mathrm{diag}(\Psi W \Psi^\top \Psi W \Psi^\top) = \mathrm{diag}(\Psi W W \Psi^\top)\,, \tag{2}$$

where the last expression is nothing but the vector of row sums of squared entries of $W\Psi^\top$. Here we used that $\Psi^\top\Psi$ is the identity matrix since the full eigenvector basis is orthonormal; in practice since we truncate the basis, the equality above is only an approximation, but it still gives a measure of how much each voxel's connectivity contributes to the classification.

Further Interpretability. Deeper insight into the learned prediction model can be obtained if one uses so-called bilinear classifiers. To motivate their use, recall that a quantity of the form $u^\top D v$ captures the un-normalized correlation between BOLD signals weighted by spatial maps Ψu and Ψv (c.f. *Informativeness*). Assuming that a limited number of such pair-wise correlations should suffice for prediction, one can seek a classifier of the form $\hat{y} = \mathrm{sign}(\sum_k \sigma_k u^{k\top} D v^k)$. Here, we let k index the pair-wise terms; without loss of generality, we can assume that vectors u^k and v^k have unit lengths, and to compensate, we introduce a scalar factor σ_k. With this notation, one can think of σ_k as the importance of each pair-wise connectivity $u^{k\top} D v^k$ to classification task. It is easy to show that we can rewrite $\hat{y} = \mathrm{sign}(\sum_k \sigma_k u^{k\top} D v^k) = \mathrm{sign}(\sum_{i,j} W_{ij} D_{ij})$. Here $W = \sum_k \sigma_k u^{k\top} D v^k$, which is exactly the singular value decomposition (SVD) of W, and now it becomes clear that our assumption is equivalent to W being a low-rank matrix. Classifiers satisfying low-rank constraints have been studied in the machine learning literature, and we will use Bilinear SVM [8] in our experiments. In practice, we find that the learned W is symmetric up to a small error (presumably since D is symmetric), so we apply SVD to the symmetrized matrix $W + W^\top$. Due to symmetry, we get $u^k = v^k$, and for visualization we depict the spatial maps Ψu^k ordered by the decreasing singular value σ_k.

3 Experiments and Results

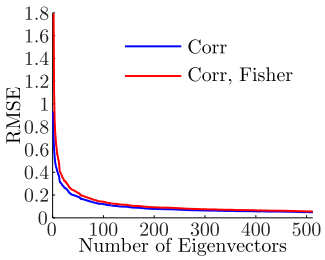
Data and Preprocessing. We utilize the public dataset[3] from Beijing Eyes Open/Closed study [6]. The study included 48 healthy subjects who underwent three resting state scanning sessions. In the first session, all of the subjects were scanned with their eyes closed; the remaining two sessions were scanned one with eyes open (EO) and the other with eyes closed (EC) in a random order counterbalanced across subjects. Since our goal is to conduct analyses for distinguishing EO and EC, we use the data from the second and third sessions only.

[3] `http://fcon_1000.projects.nitrc.org/indi/retro/BeijingEOEC.html`

The data is processed in the following manner. As in [6], we start by discarding the first 10 volumes of each session. Next, we use SPM8 (The Wellcome Trust Centre for Neuroimaging) for spatial preprocessing. We then use the CONN Toolbox [11] with default settings to obtain corrected BOLD signal. The corrected BOLD signal is z-scored at every voxel, giving the final normalized BOLD signal. This pipeline failed to process two of the subjects, and thus all of our experiments are based on the remaining 46 subjects.

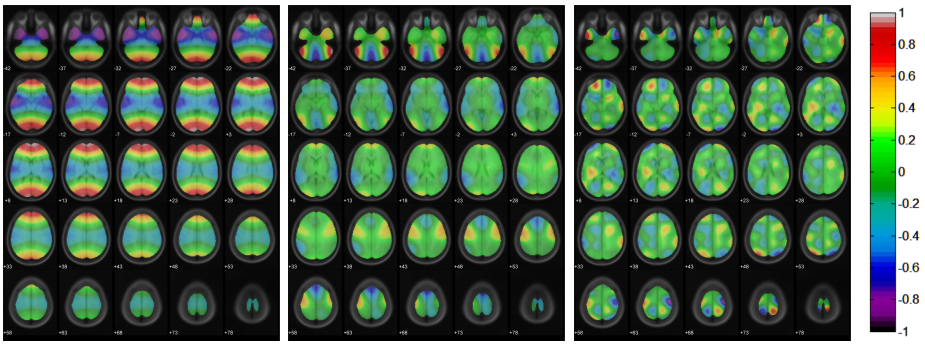
To obtain the proposed representation of resting state connectivity, we first compute a number of low-frequency eigenvectors of the grey matter voxel grid Laplacian. Next, we normalize the eigenvectors to have unit ℓ_1-norm. The normalized BOLD signal is then weighted by the eigenvectors, and the resulting time series are used to compute the dot product and correlation matrices. For comparison, we also compute connectivity matrices based on the commonly used Automated Anatomical Labeling (AAL) atlas [9] with 90 ROIs.

Experimental Results. Fig. 1 depicts the eigenvectors of the grey matter template graph Laplacian that correspond to the indices $5, 25$ and 75. For ease of visualization, the eigenvectors are normalized to have maximum absolute value of 1. The corresponding spatial maps can be seen to have varying weights, both positive and negative, and are distributed across the entire grey matter volume. We stress that no physiological meaning is associated with these eigenvectors.



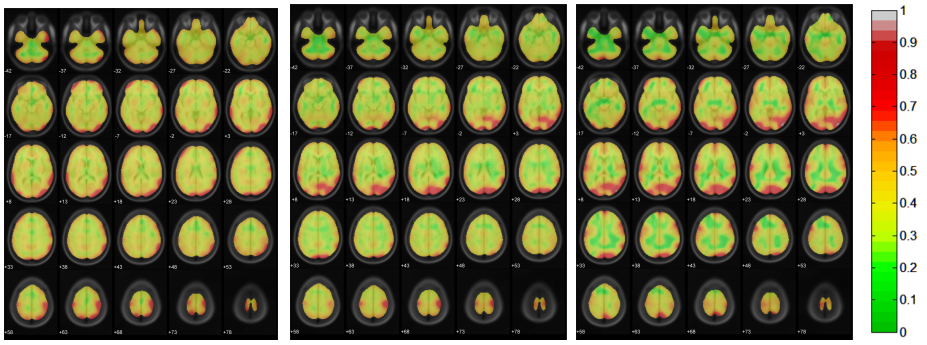
The inset plot on the right depicts the error incurred when reconstructing correlations between AAL ROIs from our representation. In this experiment, we compute all unique non-diagonal correlations between AAL ROIs for one subject. Next, these ground truth values are approximated using Eq. (1) from our representation. We varied the number of eigenvectors used in our representation from 1 to 512. The error is measured by the root-mean-square error (RMSE) over pairs of ROIs. The plot shows the RMSE both for correlations and their Fisher z-transformed versions. As expected, the reconstruction error decreases with the growing number of eigenvectors.

Next, we exemplify our representation in a multivariate analysis task, where our goal is to train classifiers that use resting state connectivity for predicting whether the subject had their eyes open or closed for a particular scan. We train three different classifiers: linear discriminant analysis (LDA) as implemented in MATLAB, with default settings; a linear support vector machine (SVM) as implemented in libsvm [3] with default parameters; and Bilinear SVM (maximum rank set to $d = 8$) with our in-house implementation that directly follows the algorithm description in [8]. LDA and SVM input a vector containing all of the unique entries in connectivity matrices; Bilinear SVM inputs the connectivity matrices directly. Each feature input to the linear SVM is z-scored over all subjects as this improves the performance; LDA is completely insensitive to z-scoring; the inputs to Bilinear SVM are not z-scored as this would potentially destroy the low rank structure of the coefficient matrix.



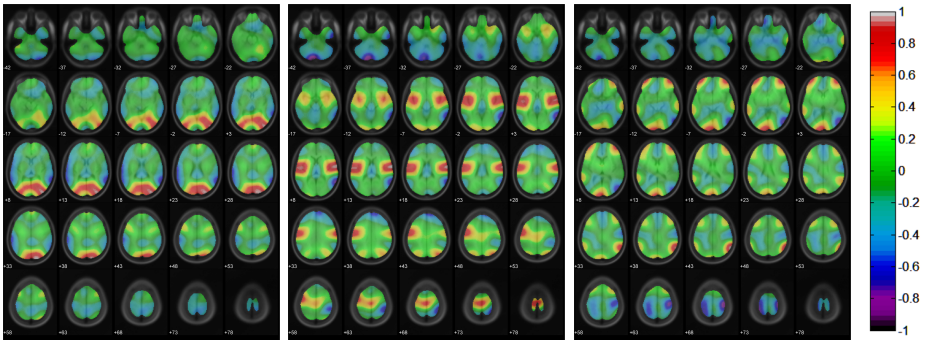
(a) (b) (c)

Fig. 1. Brain images corresponding to the eigenvectors of index (a) 5, (b) 25, and (c) 75 of the grey matter template graph Laplacian. Eigenvectors with higher indices have higher frequencies; i.e., oscillate more frequently in the space.



(a) (b) (c)

Fig. 2. Images depicting importance of voxel connectivity (i.e., "importance" maps), computed from (a) LDA, (b) SVM, and (c) Bilinear SVM classifier weights, respectively, using Eq. (2). Please refer to the text for details.



(a) (b) (c)

Fig. 3. Images corresponding to the largest three singular values of the Bilinear SVM weights, appearing left to right. Please refer to the text for details.

Table 1. Average accuracy (%) of leave-one-out cross-validation over the two sessions.

Method	LDA		SVM		Bilinear SVM	
	corr	dot	corr	dot	corr	dot
AAL	78.26	78.26	80.44	76.09	76.08	79.35
AALrz	80.44	73.92	79.35	72.83	77.18	69.57
Proposed	90.22	89.13	91.30	86.96	85.87	86.96

The classification experiments are conducted as follows. To avoid confounding, we train and test the classifiers on each session (i.e., the second and third sessions) separately. Due to the small sample size and standard practice in the neuroimaging literature, classification performance is measured via leave-one-out cross-validation accuracy. We report average accuracy over the two sessions.

Table 1 shows the performance of these classifiers in detecting EO vs. EC for various representations of functional connectivity. We compare against two versions of AAL ROI based approach—one using the voxel-wise z-scored corrected BOLD signal (AAL) and the other using the corrected BOLD signal that is z-scored region-wise (AALrz). Since there are 90 AAL ROIs, we build our representation using 90 eigenvectors; thus, all connectivity representations in this experiment have the same size (90×90). Columns marked with "dot" use as features the un-normalized correlations (e.g. entries D_{ij} directly) and columns marked with "corr" use the usual correlation (e.g. the normalized entries $D_{ij}/\sqrt{D_{ii}D_{jj}}$). For all combinations, our proposed representation is seen to improve classification accuracy over the AAL by a non-negligible amount.

Fig. 2 depicts the overall importance maps from Eq. (2), computed using the weights from classifiers trained on the entire dataset. For visualization purposes, the maps are scaled to have maximum absolute value of 1. The regions highlighted in red correspond to the voxels whose connectivity is most helpful for classification. For example, the highlighted posterior regions coincide roughly with the visual cortex, where visual sensory data is processed. Fig. 3 depicts the spatial maps corresponding to the first three singular values of the Bilinear SVM's symmetrized weight matrix, and provides an illustration of the information content that can be extracted by our methods. The image corresponding to the first singular value, which represents the component of largest influence, shows prominently (in red) the contribution of the visual cortex to the task of discriminating between EO and EC. Furthermore, the motor cortex and orbitofrontal cortex—regions which are implicated in voluntary movement and the conscious control of attention (often referred to as executive function), respectively—are prominent (red) in the images corresponding to the second and third singular values, respectively, consistent with their secondary roles in the eyes open condition. These results are anatomically consistent with existing analyses that interrogate network differences between these two states [7,12]. Note that classifier weights should be interpreted with caution, as they can reflect properties of the methods used rather than the data they are applied to; yet the fact that they are consistent with known physiology is encouraging.

4 Conclusion

We have introduced a novel information-rich compact representation of functional connectivity based on using low-frequency Laplacian eigenvectors as the spatial maps. The resulting maps only depend on the grey matter template and so enjoy the kind of data-independent universality usually associated with atlas-based methods. Experiments confirm increased classifier accuracy with our representation and lead to plausible interpretations.

Acknowledgments. This work was supported by NIH grant T32 MH019908, ONR MURI grant N00014-13-1-0341, AFOSR grant FA9550-12-1-0372, NSF grant DMS 1228304, and the Max Planck Center for Visual Computing and Communications. Support for the publicly available Beijing EO/EC data was provided by a grant from the National Natural Science Foundation of China: 30770594 and a grant from the National High Technology Program of China (863): 2008AA02Z405.

References

1. Belkin, M., Niyogi, P.: Semi-supervised learning on riemannian manifolds. Machine Learning 56(1-3), 209 – 239 (2004)
2. Castellanos, F.X., Martino, A.D., Craddock, R.C., Mehta, A.D., Milham, M.P.: Clinical applications of the functional connectome. NeuroImage 80, 527 – 540 (2013)
3. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, 27:1 – 27:27 (2011)
4. Chung, F.R.K.: Spectral Graph Theory. American Mathematical Society (1997)
5. Fornito, A., Zalesky, A., Breakspear, M.: Graph analysis of the human connectome: Promise, progress, and pitfalls. NeuroImage 80, 426 – 444 (2013)
6. Liu, D., Dong, Z., Zuo, X., Wang, J., Zang, Y.: Eyes-open/eyes-closed dataset sharing for reproducibility evaluation of resting state fMRI data analysis methods. Neuroinformatics 11(4), 469 – 476 (2013)
7. Patriat, R., Molloy, E.K., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V., Birn, R.M.: The effect of resting condition on resting-state fMRI reliability and consistency: A comparison between resting with eyes open, closed, and fixated. NeuroImage 78, 463 – 473 (2013)
8. Pirsiavash, H., Ramanan, D., Fowlkes, C.: Bilinear classifiers for visual recognition. In: NIPS. pp. 1482 – 1490 (2009)
9. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15(1), 273 – 289 (2002)
10. Varoquaux, G., Craddock, R.C.: Learning and comparing functional connectomes across subjects. NeuroImage 80, 405 – 415 (2013)
11. Whitfield-Gabrieli, S.L., Nieto-Castanon, A.: Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. Brain Connectivity 2(3), 125 – 141 (2012)
12. Xu, P., Huang, R., Wang, J., Dam, N.T.V., Xie, T., Dong, Z., Chen, C., Gu, R., Zang, Y.F., He, Y., Fan, J., Luo, Y.: Different topological organization of human brain functional networks with eyes open versus eyes closed. NeuroImage 90 (2014)