

MULTIMODAL GRAPH THEORETICAL ANALYSIS OF FUNCTIONAL BRAIN CONNECTIVITY USING ADAPTIVE TWO-STEP STRATEGY

Djalel-Eddine Meskaldji^{*†}

Dimitri Van De Ville^{*†}

^{*} Medical Image Processing Lab (MIPLAB), Institute of Bioengineering, EPFL, CH-1015 Lausanne, Switzerland

[†] Department of Radiology and Medical Informatics, University of Geneva, CH-1211 Geneva, Switzerland

ABSTRACT

Recently, we proposed a two-step adaptive strategy for the statistical analysis of brain connectivity that is based on a first screening at the subnetwork level and a filtering at the connection/node level. The method was shown to guarantee strong control of type-I error through rigorous statistical proofs. In addition, the gain of power obtained by this method is considerable especially with an appropriate decomposition of the global network. Here, we discuss the extension of the two-step methods to multivariate statistics and we compare its performance against both standard methods and univariate two-step methods. We present as well a practical example of detecting topological nodal differences between functional connectivity matrices of resting state and movie-watching, respectively.

Index Terms— Neuroimaging, brain networks, functional connectivity, type-I error control, graph theory.

1. INTRODUCTION

Statistical inference and graph theory are two important elements of the emerging field of neuroscience called *connectomics*, in which brain connectivity is represented by a network (a weighted graph), where nodes represent brain regions and edge weights represent a certain measure of the strength of connectedness between pairs of nodes [2, 3, 4]. When comparing brain networks at the node/edge level, a large number of tests has to be performed and the type-I error rate, such as the family-wise error rate or the false discovery rate, has to be controlled by applying a multiple testing procedure. One important aspect that should be considered when comparing brain networks is the data structure and the positive dependence that might be present between tests. Connected brain regions (nodes) are reported to behave coherently [5, 6]. Furthermore, prior information about the dependence structure between tests could be available in some situations. In recent work, we developed two-step procedures that exploit the data structure and prior information of positive dependence between tests [7, 8, 9, 1]. The proposed procedures work as the following. First, we group the global set of tests into subsets either by respecting the data structure or according

to prior information, and we apply a first screening at a pre-defined threshold using the mean as a summary statistic for each subset. This screening results in two types of subsets; positive subsets and negative subsets. According to this typing, the original p-values are modified such that they can be used with any multiple comparison procedure to control the desired type-I error rate. We investigated the performances of the new procedures using different screening thresholds in different simulation scenarios in which the new procedures showed great improvements in terms of power.

In this paper, we discuss the extension of the two-step methods to the multivariate case, that is, in situations where p-values are computed on the basis of a vector of values for each node instead of a single value. The multivariate tests exploit the hidden dependence between the different univariate measures. The extension is proposed for both paired and non-paired multivariate tests. We also present an application of univariate and multivariate two-step methods to compare functional brain networks computed within subjects in two different mental states; i.e., resting state and movie-watching.

2. METHODS

2.1. Description of the study

The study consists in detecting topological differences in functional connectivity between two mental states; i.e., resting state (RS) and movie-watching (MW). The study is carried out on functional brain connectivity matrices of healthy subjects [10]. A functional connectivity matrix is a square matrix where rows/columns correspond to brain regions (i.e., 90 regions from the AAL atlas [11]) and each element represents a value of the functional connectivity between the pair of brain regions, typically Pearson correlation between the associated timecourses [12]. This value is situated between -1 and $+1$, and is estimated by the correlation between functional time series measured on each brain region. In our study, we only considered values that are larger than a pre-defined positive threshold $U = 0.1$ or $U = 0.2$. Although the choice of these values is somehow arbitrary but it permits to highlight the influence of thresholding on the inference. For more discussion see [9].

2.2. Nodal graphical measures

Here, we consider multivariate nodal measures, that is, for each node in the brain network, a vector of q nodal topological measures is estimated. A quite different nodal approach based on classification has been proposed in [13]. One could also

This work was supported in part by the Swiss National Foundation (grant number PP00P2-146318), and in part by the Center for Biomedical Imaging (CIBM) of the Geneva-Lausanne Universities and the EPFL, as well as the Leenaards and Louis-Jeantet foundations.

use multivariate connection measures.

In our study, we used two nodal measures ($q = 2$). The first measure is the nodal strength (NS), which is defined as the sum of the weights of the edges connected to that node. It is computed by the sum of row/column values corresponding to that node in the connectivity matrix. The second measure is the nodal efficiency (NE), which is the inverse of the average minimum weighted path length from that specific node to the other nodes.

2.3. Inference for paired multivariate inference

Suppose that the global set of nodes is indexed by $j \in J = \{1, \dots, M\}$. When comparing two populations of sizes n_1 and n_2 respectively, we modal the original data in the following form. Let \mathbf{x}_{jk} be a vector of length q , which represents the nodal measures of node j and for subject $k = 1, \dots, n_1$ in the first group, and similarly, let \mathbf{y}_{jk} be a vector of size q , that represents the nodal measures of node j , for subject $k = 1, \dots, n_2$ in the second group.

For both paired and non-paired multivariate tests, the hypothesis testing p-values are computed with the Fisher distribution \mathcal{F} , according to the Hotelling theorem.

In the paired case, we have $n_1 = n_2 = n$ and we consider the differences

$$\mathbf{D}_{jk} = \mathbf{x}_{jk} - \mathbf{y}_{jk}, \text{ for } j = 1, \dots, M \text{ and } k = 1, \dots, n.$$

Suppose that $\mathbb{E}(\mathbf{D}_{jk}) = \delta_j$ and $\text{Cov}(\mathbf{D}_{jk}) = \Sigma_j$. The j^{th} null hypothesis is that the measures have the same mean, that is, $\delta_j = 0$ and the alternative is that there is a difference for at least one measure among the q measures. In addition, if we assume for each $j = 1, \dots, M$ that $\mathbf{D}_{j1}, \dots, \mathbf{D}_{jn}$ are independent multivariate normal random variables with mean δ_j and covariance Σ_j , then

$$T_j^2 = n(\bar{\mathbf{D}}_j - \delta_j)^T \mathbf{S}_j^{-1} (\bar{\mathbf{D}}_j - \delta_j), \quad (1)$$

where $\bar{\mathbf{D}}_j = \frac{1}{n} \sum_{k=1}^n \mathbf{D}_{jk}$, and $\mathbf{S}_j = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{D}_{jk} - \bar{\mathbf{D}}_j)(\mathbf{D}_{jk} - \bar{\mathbf{D}}_j)^T$, is distributed as an $((n-1)q/(n-q))\mathcal{F}_{q, n-q}$ random variable. Hence, the j^{th} null hypothesis is rejected at level α if the observed score \hat{T}_j^2 satisfies

$$\hat{T}_j^2 > \frac{(n-1)q}{n-q} \mathcal{F}_{q, n-q}(\alpha), \quad (2)$$

or equivalently, if the associated p-value

$$p_j = \mathbb{P}\left(T^2 > \hat{T}_j^2\right) \quad (3)$$

is less than α , where T^2 is distributed as an $((n-1)q/(n-q))\mathcal{F}_{q, n-q}$ random variable.

We compute the z-score Z_j , associated to the p-value p_j using the transformation $Z_j = \Phi^{-1}(1 - p_j)$, where Φ denotes the standard normal cumulative distribution function.

The p-values corresponding to the null hypotheses are distributed as (or stochastically larger than) the uniform distribution $\mathcal{U}(0, 1)$ and those corresponding to the alternative are stochastically smaller than $\mathcal{U}(0, 1)$. The null hypotheses become $\mu_j \leq 0$ and the alternative hypotheses become $\mu_j > 0$.

2.4. Two-step methods

We group the M tests (or nodes) into m subsets J_1, \dots, J_m such that $\bigcup_{i=1}^m J_i = J$. The decomposition is based on prior information about the positive dependence between tests or is obtained by using a data-driven decomposition method.

The two-step methods work as follows. The information in each subset is summarised by a summary statistic; the standardised mean $S_i = \sum_{j \in J_i} Z_j / \sigma_i$, where σ_i is the standard deviation of the subset test scores $\{Z_j : j \in J_i\}$. We compute the p-values $P_i (i = 1, \dots, m)$ corresponding to the summary statistics by $P_i = 1 - \Phi(S_i)$, $i = 1, \dots, m$. The central limit theorem allows us to make this approximation especially when the subset sizes become quite large.

The first step consists in applying a multiple testing procedure to the subset p-values $P_i (i = 1, \dots, m)$ at a predefined level. This screening results in two classes of subsets: positive subsets and negative subsets. Based on the results of the first step, we perform a multiple testing procedure to control the type I error rate at the level of single hypotheses. This consists of the second step of two-step methods.

2.5. Relaxed two-step methods

The two-step procedures proposed in [8] do not require the estimation of the conditional p-values. These two-step procedures work as follows: we divide the original p-values inside the positive (negative) subsets by a positive number called the *relaxation (tightening) coefficient* and then perform a new multiple testing procedure to the modified p-values. For a fixed value of the tightening coefficient, the relaxation coefficient is chosen sufficiently small such that the type-I error rate is controlled at a predefined level α .

Concerning the screening in the first step of the relaxed methods, we used either a multiplicity correction or no correction. We call these methods Relaxed Method With Correction (RMWC) and Relaxed Method with No Correction (RMNC) respectively. In these two cases, we consider only positive subsets in the second step. In a third case, we also investigate inside the negative subsets. We call this the Relaxed Method In/Out (RMIO). In this case, the tightening coefficient was set to be 0.5 and we used a multiplicity correction in the first step.

2.6. Decomposition of the global set of nodes

We use two kinds of decompositions. The first one is a prior decomposition based on brain lobes. The second is data-driven decomposition in which the global set of nodes is partitioned into communities (subsets of nodes) by applying two adaptive algorithms based on modularity maximisation. The first one is the leading eigenvector (LEV) in which the modularity function is rewritten in matrix terms which leads to express the optimisation task as a spectral problem in linear algebra [14]. The second algorithm is called walktrap (WT) which is based on the fact that random walks on graph tend to get trapped into densely connected subgraphs corresponding to communities [15]. The data-driven algorithms were only applied to the control group (resting state group in our case) average matrix to avoid any influence on the statistical inference.

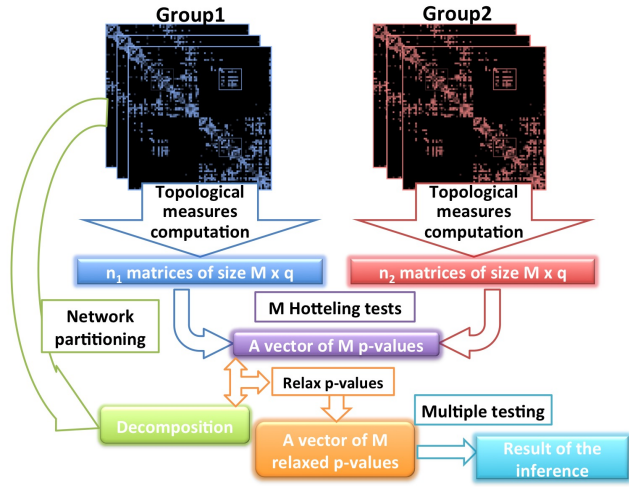


Fig. 1. Illustration of the pipeline of the proposed strategy.

2.7. Correction for multiplicity

We use two different type-I error metrics; the expected number of false positives $\mathbb{E}(\text{FP})$ and the false discovery rate $\text{FDR} = \mathbb{E}(\text{FP}/R)$, where R is the number of rejections. These two type-I error metrics are particular cases of the scaled error rate $\mathbb{E}(\text{FP}/s(R))$, where s is any non-decreasing function [16]. To control the expected number of false positives, we used the Bonferroni procedure which performs each single test at level α/M . The false discovery rate is controlled by applying the linear step-up procedure (LSU) proposed in [17]. The procedure consists in ordering the p-values from smallest to largest and choose as a significance threshold the largest p-value situated under the increasing line with slope α/M .

The different steps of the proposed strategy are illustrated in Fig. 1.

3. RESULTS AND DISCUSSION

We compared the performance of two-step methods (RMWC, RMNC and RMIO) with the standard methods (SM) (that do not consider positive dependence and data structure and performs the multiple testing procedure on the original p-values) in terms of significant results by considering three statistics: nodal strength (NS), nodal efficiency (NE) and a combination of both as a bi-variate statistic (BV). In tables 1 and 2, we report the number of nodes declared to be significantly different between RS and MW brain networks for two different values of the threshold U , using two multiple testing procedures, Bonferroni and LSU. In each table, we report the number of significant nodes either by using the SM or the two-step methods RMWC, RMNC and RMIO. For the relaxed methods we give a value for each decomposition method: Lobes, WT and LEV.

The results show the gain of the power of detecting real differences when using two-step methods. This is because the information of the data structure is exploited and not ignored. The bi-variate statistic gives less significant results in almost all cases. Nevertheless, when the two network measures NS

and NE are tested separately, one should have to correct for multiplicity for $2M$ tests and not only M tests. This is not the case for the bi-variate case in which the number of tests performed is M . In addition, the multivariate statistic catches nodes where the difference is a combined effect between NS and NE.

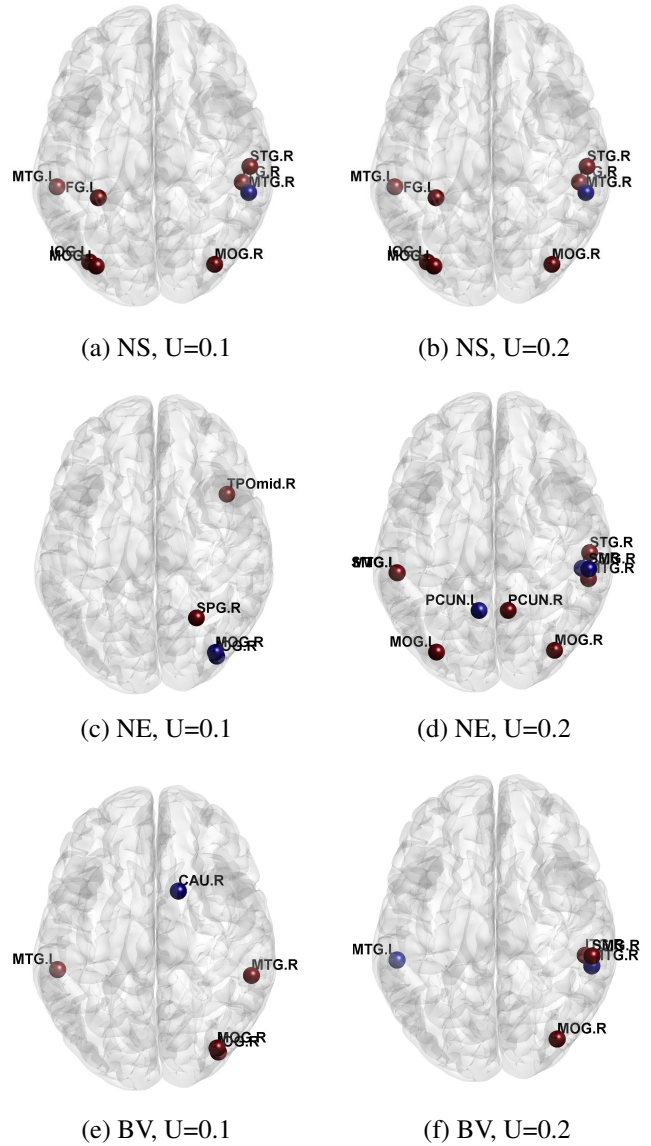


Fig. 2. Nodes declared as significantly different between the two cognitive states RS and MW using the two-step method RMWC with the Bonferroni procedure. Three nodal statistics are used: nodal strength (NS) (a and b), nodal efficiency (NE) (c and d) and a combination of both as a bi-variate statistic (BV) (e and f). Blue nodes are the only ones detected by the standard method. For each statistic we show two plots that correspond to two thresholds $U = 0.1$ and $U = 0.2$. MOG=Medial orbital gyrus, IOG=Inferior orbital gyrus, FFG=Fusiform gyrus, SPG=Superior parietal gyrus, SMG=Supra marginal gyrus, PCUN=Precuneus, CAU=Caudate, STG=Superior temporal gyrus, MTG=Medial temporal gyrus, TPomid=Medial temporal pole, ITG=Inferior temporal gyrus.

| | SM | RMWC | | | RMNC | | | RMIO | | |
|----|-----|-------|-----|-----|-------|-----|-----|-------|-----|-----|
| | | Lobes | LEV | WT | Lobes | LEV | WT | Lobes | LEV | WT |
| NS | 1 | 8 | 7 | 6,5 | 5 | 6,5 | 6,5 | 5 | 6,5 | 6,5 |
| NE | 2,3 | 4,10 | 4,7 | 4,7 | 4,6 | 4,7 | 4,7 | 4,6 | 4,7 | 4,7 |
| BV | 1,2 | 4,5 | 5 | 4,5 | 3,5 | 4,5 | 4,5 | 3,5 | 4,5 | 4,5 |

Table 1. The number of nodes declared to be significantly different using SM, RMWC, RMNC and RMIO under the strong control of false positives (Bonferroni). For the relaxed methods, we give three values that correspond to the different decompositions: Lobes, LEV and WT. Three different nodal measures are used: NS, NE and BV. Two values are reported if there is a difference in the number of significant nodes when using the two different values of the threshold $U = 0.1$ and $U = 0.2$.

| | SM | RMWC | | | RMNC | | | RMIO | | |
|----|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | Lobes | LEV | WT | Lobes | LEV | WT | Lobes | LEV | WT |
| NS | 13 | 21,22 | 40,39 | 35,33 | 22 | 34,33 | 35,33 | 21 | 27 | 25 |
| NE | 4,19 | 28,32 | 51,45 | 52,57 | 34,28 | 51,42 | 52,57 | 23,22 | 32,35 | 32,39 |
| BV | 5 | 33,27 | 41,48 | 41,48 | 33,27 | 41,48 | 41,48 | 21 | 22,25 | 22,25 |

Table 2. The number of nodes declared to be significantly different using SM, RMWC, RMNC and RMIO under false discovery control (LSU). Details are the same as in Table 1.

In Fig. 2, we show the brain regions that are significantly different between RS and MW using the two-step method RMWC with the Bonferroni procedure. We used the Brain Net viewer tool [18] for this visualisation. In panels (a) and (b), we present the significant nodes in terms of NS. In all these nodes, the nodal strength is larger for RS brain networks. In panels (c) and (d), we represent the significant nodes in terms of NE. In all these nodes, the nodal efficiency is larger for MW brain networks. In panels (e) and (f), we represent the significant nodes using the bi-variate statistic (BV). The bi-variate statistic may detect nodes that are not detected by the univariate statistics. One could use $q > 2$ measures to benefit more from the multivariate case. However, the interpretation of the results becomes much harder. Also in previous analysis we found that functional connectivity is strongly increased during RS [12]. In particular, the medial temporal gyrus (MTG) becomes more independently active when processing information during MW. We observe an opposite trend for NE, which might indicate that although nodal strength is increased during RS, the average path length is decreased as some nodes are not part of large RS networks. However, we also found that the results are less robust for NE when changing the threshold value, which might also explain why results do not improve for the bivariate case.

4. REFERENCES

- [1] D.-E. Meskaldji, *Multiple Comparison Procedures for Large Correlated Data with Application to Brain Connectivity Analysis*, Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne, 2013.
- [2] L. Cammoun, X. Gigandet, D. Meskaldji, J. P. Thiran, O. Sporns, K. Q. Do, P. Maeder, R. Meuli, and P. Hagmann, "Mapping the human connectome at multiple scales with diffusion spectrum MRI," *Journal of Neuroscience Methods*, vol. 203, no. 2, pp. 386–397, 2012.
- [3] K. J. Friston, "Functional and effective connectivity in neuroimaging: a review," *Brain connectivity*, vol. 1, no. 1, pp. 13–36, 2011.
- [4] P. Hagmann, L. Cammoun, X. Gigandet, S. Gerhard, P. E. Grant, V. Wedeen, R. Meuli, J.-P. Thiran, C. J. Honey, and O. Sporns, "MR connectomics: Principles and challenges," *Journal of Neuroscience Methods*, vol. 194, no. 1, pp. 34–45, 2010.
- [5] W. Penny and K. Friston, "Mixtures of general linear models for functional neuroimaging.," *IEEE Trans Med Imaging*, vol. 22, no. 4, pp. 504–14, 2003.
- [6] E. T. Bullmore and D. S. Bassett, "Brain graphs: Graphical models of the human brain connectome," *Annual Review of Clinical Psychology*, vol. 7, no. 1, pp. 113–140, 2011, PMID: 21128784.
- [7] D.-E. Meskaldji, M.-C. Ottet, L. Cammoun, P. Hagmann, R. Meuli, S. Eliez, J.-P. Thiran, and S. Morgenthaler, "Adaptive Strategy for the Statistical Analysis of Connectomes," *PLoS ONE*, vol. 6, no. 8, pp. e23009, 08 2011.
- [8] D.-E. Meskaldji, P. Hagmann, J.-P. Thiran, and S. Morgenthaler, "Two step multiple comparison procedures for positively dependent data," *ArXiv e-prints*, 2013.
- [9] D.-E. Meskaldji, E. Fisch-Gomez, A. Griffa, P. Hagmann, S. Morgenthaler, and J.-P. Thiran, "Comparing connectomes across subjects and populations at different scales," *NeuroImage*, vol. 80, no. 0, pp. 416 – 425, 2013, Mapping the Connectome.
- [10] H. Eryilmaz, D. Van De Ville, S. Schwartz, and P. Vuilleumier, "Impact of transient emotions on functional connectivity during subsequent resting state: A wavelet correlation approach," *NeuroImage*, vol. 54, no. 3, pp. 2481–2491, February 2011.
- [11] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in {SPM} using a macroscopic anatomical parcellation of the {MNI} {MRI} single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273 – 289, 2002.
- [12] J. Richiardi, H. Eryilmaz, S. Schwartz, P. Vuilleumier, and D. Van De Ville, "Decoding brain states from fMRI connectivity graphs," *NeuroImage*, vol. 56, no. 2, pp. 616–626, 2011.
- [13] J. Richiardi, S. Achard, E. Bullmore, and D. Van De Ville, "Classifying connectivity graphs using graph and vertex attributes," in *Proceedings of the International Workshop on Pattern Recognition in Neuroimaging*, Seoul, South Korea, 2011, pp. 45–48, IEEE.
- [14] M. E. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Physical review E*, vol. 74, no. 3, pp. 036104, 2006.
- [15] P. Pons and M. Latapy, "Computing communities in large networks using random walks," in *Computer and Information Sciences-ISCIS 2005*, pp. 284–293. Springer, 2005.
- [16] D.-E. Meskaldji, J.-P. Thiran, and S. Morgenthaler, "A comprehensive error rate for multiple testing," *ArXiv e-prints*, dec 2011.
- [17] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *J. Roy. Statist. Soc. Ser. B*, vol. 57, no. 1, pp. 289–300, 1995.
- [18] M. Xia, J. Wang, and Y. He, "Brainnet viewer: A network visualization tool for human brain connectomics," *PLoS ONE*, vol. 8, no. 7, pp. e68910, 07 2013.