

The Optimal Setting for Multilayer Modularity Optimization in Multilayer Brain Networks*

M.G. Puxeddu, *Member, IEEE*, M. Petti, D. Mattia and L. Astolfi, *Member, IEEE*

Abstract— Community detection plays a key role in the study of brain networks, as mechanisms of modular integration and segregation are known to characterize the brain functioning. Moreover, brain networks are intrinsically multilayer: they can vary across time, frequency, subjects, conditions, and meaning, according to different definitions of connectivity. Several algorithms for the multilayer community detection were defined to identify communities in time-varying networks. The most used one is based on the optimization of a multilayer formulation of the modularity, in which two parameters linked to the spatial (γ) and temporal (ω) resolution of the uncovered communities can be set. While the effect of different γ -values has been largely explored, which ω -values are most suitable to different purposes and conditions is still an open issue. In this work, we test the algorithm performances under different values of ω by means of ad hoc implemented benchmark graphs that cover a wide range of conditions. Results provide a guide to the choice of the ω -values according to the network features. Finally, we show an application of the algorithm to real functional brain networks estimated from electro-encephalographic (EEG) signals collected at rest with closed and open eyes. The application to real data agrees with the results of the simulation study and confirms the conclusion drawn from it.

I. INTRODUCTION

Modeling the brain as a complex network made of areas anatomically and functionally interconnected [1] allows the use of graph theory to derive information underlying the brain functioning. In this sense, network science becomes a powerful instrument to reveal non-trivial brain properties such as its modular organization [2]. Modules (communities, clusters) are groups of nodes strongly connected, usually related to specific functions of the system. Investigating the brain community structure is important due to the role it plays in regulating segregation and integration mechanisms between subsystems. Previous works have already shown that functional networks obtained from EEG sensor recordings have modular properties that can be related to the condition of the subject [3]. Moreover, while most of the studies consider nodes linked by static edges, this assumption

is simplistic. In fact, brain interactions might vary in time (e.g. during the execution of a task, in different conditions or at different recording sessions), in frequency (in M/EEG acquisitions) or across subjects. Hence, brain networks can be studied without either throwing away or combining different information, by using a multilayer framework: several classical linked networks each one encoding specific attributes of the system. Specifically, tracking the network modules across the layers can reveal changes in time, as well as a stability of the network structure, which could be both physiologically meaningful. At the state of the art, several methods have been proposed to detect communities in multilayer networks. The most used, especially in the field of network neuroscience, is based on the optimization of a multilayer formulation of the modularity [4]. Modularity is an objective function that estimates the quality of a partition with respect to a null model. The algorithm that realizes this optimization is called genLouvain [5], and is a generalized version of the well-known Louvain [6], used for the single layer modularity optimization. Such algorithm depends on two parameters, γ and ω , which, respectively, determine the spatial and temporal resolution of the found communities. The effect of γ on (gen)Louvain has been largely explored: γ -values lower than 1 (default value) entail larger communities and vice-versa. Concerning ω , it represents the connection between a node and itself in consecutive layers, so that higher values mean that the nodes show the same behavior across layers. However, an agreement about how to set ω according to specific network's features is missing. With this work we aim to provide guidelines about how to correctly set the temporal properties of genLouvain to study time-varying brain networks. To this purpose, we propose a toolbox to generate benchmark graphs whose main advantage is its ability to simulate a wide range of conditions. The tools already available, such as in [7]-[8] (multilayer version of the Girvan and Newman model [9]) show the limitation of constraining most of the parameters characterizing the network (e.g. number of nodes, number of clusters, etc). In [10] a better tool is proposed in which users can set some parameters (e.g. number of nodes and clusters). Nevertheless, here we present an even more flexible tool, based on the one proposed in [11], in which graph density, number of clusters, noise level (modeled as a random permutation of a percentage of links) and the percentage of nodes changing module at a given layer are features that can be set. We then propose a comparative analysis between ω s, exploiting the introduced tool. The graphs we generated account for a wide class of features systematically varied in a range typical of EEG-based brain networks. Moreover, we consider two cases: one in which the community structure is stationary across the layers and one in which it changes dynamically. In fact, getting a single partition out from a multilayer network can be useful, for example if we are interested in phenomena

* Research partially supported by European Union's Horizon 2020 program under the Marie Skłodowska-Curie grant agreement No 778234 (DOCMA), by Sapienza University of Rome - Progetti di Ateneo 2016 (MIME-BCI- P1161550696379A), Progetti di Ateneo 2017 (EMBRACING- RM11715C82606455) and Progetti di Avvio alla Ricerca 2018 (AR1181643680C682).

M.G. Puxeddu (corresponding author: Tel. +39 0677274034, e-mail: mariagrazia.puxeddu@uniroma1.it), M. Petti and L. Astolfi are with the Department of Computers, Control and Management Engineering, Sapienza University of Rome, Via Ariosto 25, Rome, Italy, and with the IRCCS Fondazione Santa Lucia, Via Ardeatina 306, Rome, Italy. D. Mattia is with the IRCCS Fondazione Santa Lucia, Via Ardeatina 306, Rome, Italy.

supposed stationary, or the layers represent subjects of the same category and we aim to extract common features. At the same time, it is also crucial to be able to track the clusters evolution in networks representing non-stationary phenomena. In both conditions, we statistically evaluated the accuracy of the algorithm under different factors by the analysis of variance (ANOVA). Finally, we applied genLouvain to brain functional networks estimated from EEG signals, obtained in a healthy subject during resting state at close-eyes and open-eyes. We report here the differences obtained between the community structure subtending the two phases by using different ω -values, in accordance with the guidelines provided by the simulation study.

II. METHODS

A. Multilayer Networks Generation

The tool we present consists of an algorithm implemented in Matlab environment (release 2017b). It produces directed and unweighted multilayer (ML) adjacency matrices (3D) with stationary or evolving community structure. The parameters that can be set are: number of nodes (**N**), graph density (**D**), number of layers (**nL**), clusters number (**CN**), ratio between intra-cluster and inter-cluster density (**dr**), noise level (**no**), percentage of nodes shifting community across layers (**pn**).

To get a stationary community structure, the algorithm proceeds through two main steps: (i) creation of a single layer network exploiting the algorithm described in [11], used as base for every layer; (ii) addition of the **no** level, set as an input, to each layer. With these two steps we obtain a ML network in which each slice has the same community structure obtained in (i), while the variability between layers is due to noise, applied differently to each slice through (ii). To simulate ML networks with evolving community structure the algorithm still goes through (i), but then it generates the following slice so that a certain percentage of nodes (**p**) changes its allegiances to modules. The algorithm acts only on the links involving the nodes that change membership, maintaining the rest of the network as it was created in (i). Similarly, it can also increase or decrease the **CN**, moving some nodes into a new module, or moving all the nodes of one module into the others. Thus, one can get controlled variation of the community structure, according to **p** and to the possible increase or decrease of the number of communities.

B. Simulation study

To test and compare the effect of different ω -values on the genLouvain performances, we exploited the tool described in the previous paragraph, systematically varying the networks features according to those experimentally detectable in real EEG-based functional brain networks. Thus, we generated networks with:

- **N** = 60
- **D** = [0.1, 0.3]
- **CN** = [2, 4, 6]
- **dr** = 2
- **no** = [10%, 25%, 50%]
- **nL** = [2, 10, 50, 100]

Then, we run the algorithm genLouvain with different ω -values in the range 0.1-10:

- ω = [0.1, 0.2, 0.5, 1, 2, 5, 10]

We performed a repeated measures ANOVA to evaluate the effects of the factors ω , **CN**, **no** and **nL** on the accuracy of genLouvain, assessed by the Normalized Mutual Information (NMI). The NMI is an index of similarity, borrowed from information theory, used in this context to compare the output of the algorithm with the known community structure. The mean NMI over all the layers played the role of dependent variable in the statistical analysis.

To test the algorithm on networks with dynamic community structure we performed a second simulation study in which we also included the parameter **pn** (previously described):

- **pn** = [10%, 30%, 50%, 70%, 100%]

The structure was varied for the second half of the layers, to simulate the condition in which we want to track modules over two different tasks, or two classes of subjects (e.g., healthy/patients), when we want a unique partition associated to all the layers of the first task/healthy people and a second partition associated to the second task/patients. Again, we run genLouvain with the same ω -values as before, and we performed an ANOVA for repeated measures using the NMI as dependent variable and index of accuracy.

C. Application to EEG data

EEG data were recorded using 61 electrodes (according to the extended 10-20 International System) in a healthy subject (female, 33 years old) during rest at close-eyes (CE) and open-eyes (OE). The subject gave informed consent prior to her participation and the experiments were approved by the local Ethics Committee before data acquisition started. The session was composed by 26 trials of 200 seconds each. In the first 100 seconds, the subject was asked to keep her eyes closed (task1) and in the last 100 seconds to keep her eyes open (task2). Pre-processing of the data included band pass filtering (1-45 Hz), artefact rejection and 2-seconds epoch segmentation. For each segment the corresponding functional brain network was estimated through Partial Directed Coherence (PDC) [12] and mediated in four EEG frequency bands, defined according to the Individual Alpha Frequency (IAF) [13] (IAF = 10 Hz): theta [IAF-6, IAF-3], alpha [IAF-2, IAF+2], beta [IAF+3, IAF+14] and gamma [IAF+15, IAF+30]. We assessed the significance of the connections through asymptotic statistics [14]. For each frequency band and for each task we obtained 50 binary networks of dimension 61x61. We used those relative to the alpha band for our analysis, as of interest for resting state [15]. Then we concatenated them so as to obtain 4 multilayer networks, sized $61ch*61ch*[2\ 10\ 50\ 100]nL$, in which the first half layers derive from task1 (CE) and the second half from task2 (OE), similarly to the simulation study. Finally, we run genLouvain with all the ω -values under exam to the 4 multilayer networks.

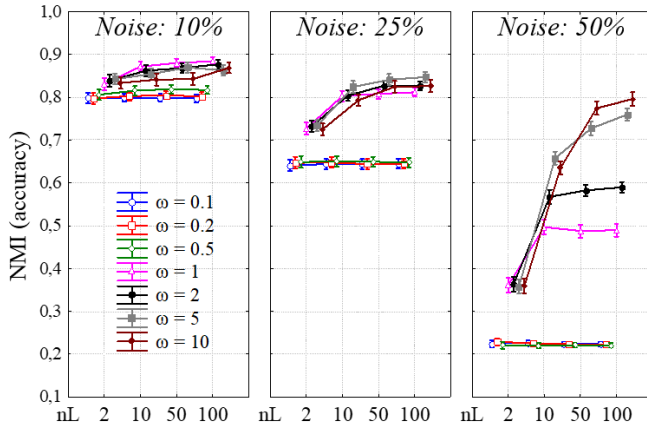
III. RESULTS

A. Simulation study

In the stationary analysis, the ANOVA performed for different ω -values on networks with stationary community

a) Stationary Community Structure

$F(36, 10692)=159.68, p<0.0001$



b) Evolving Community Structure

$F(36, 48492)=2015.5, p<0.0001$

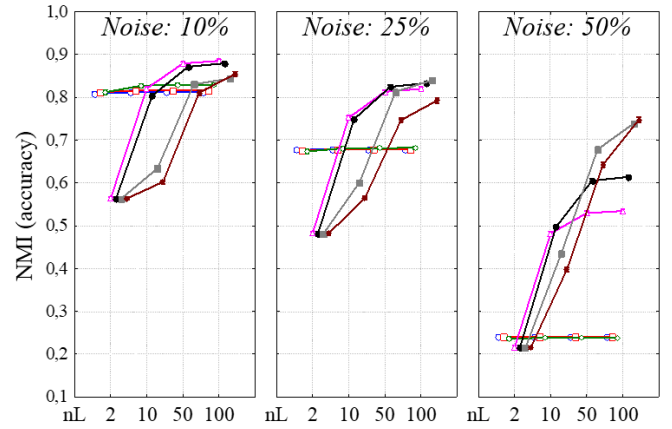


Figure 1. Plot of means of the ANOVA performed on networks with stationary community structure, panel (a), and evolving community structure, panel (b). The effect of the factors ω , number of layers (nL), and noise are shown in each panel in the case in which $D=0.3$ (similar results were obtained with $D=0.1$).

structure, together with Tuckey's post-hoc tests, show that ω -values ≥ 1 globally outperform the lower ones (Fig. 1a). In

fact, while with low noise level they all perform acceptably (NMI between 0.8 and 0.9), as the noise increases values of ω s greater than 1 show statistically higher levels of accuracy. This is even more true with many layers: genLouvain applied with high ω s can recover almost the exact community structure, even with very high percentage of noise, if the network is composed by several layers.

In the dynamic analysis, ANOVA results suggest that values of ω up to 0.5 are suitable in to detect communities that change rapidly (in few layers), while higher ω -values are preferable to discriminate two different community organizations when they persist across several layers before the change happens (Fig 1b). It is worth of note that low ω -values are insensitive to the factor nL . In very noisy networks (no=50%) neither low nor high ω -values can recover the exact community structure if it changes rapidly ($nL=2$), but $\omega \geq 1$ give better performance as nL increases. We found that the factor pn does not affect the global performances of the algorithm, meaning that it can detect small as well as big changes in community structure. Only the case in which $pn=10\%$ causes a loss of accuracy in networks of $D=30\%$, low level of noise and small CN.

While in figure we show results for $D=0.3$, this is true also for $D=0.1$, for each value of CN, and for the three cases in which CN does not change, increases and decreases. Finally, in both analysis, we observed all the ω s performing globally better in denser networks and with more clusters.

B. Application to EEG data

The results of the application to resting state EEG are shown in Fig. 2. In the network composed by 2 layers (CE and OE) the two conditions are better discriminated applying genLouvain with $\omega = 0.2$ (fig 2a). In fact, higher values force the algorithm to find a unique partition for the two conditions (fig 2d). In the network composed by more layers ($nL = [50, 100]$) the discrimination in the community structure between CE and OE is better identified by $\omega = [1, 2, 5]$ (fig 2b). With these values we obtain a pretty much stable partition until $nL/2$ (CE) and a different partition from $nL/2$ on (OE). On

the contrary, with lower values the variability of the found communities is too high even within the same task (fig 2e). These findings confirm the results of the simulation study.

Running genLouvain with the right ω -values we observed (fig2c) that during CE there is a cluster that involves the occipital electrodes and two clusters composed by electrodes from left and right hemispheres, respectively. During OE the first cluster is dismembered between left and right hemispheres, and one can observe modules becoming more hemispheric-specific. Such results are observed both in the EEG network made of $nL=2$ when $\omega=0.2$ and in the ones with $nL>2$ with values of $\omega>1$, suggesting that the choice of ω -values is correct.

IV. DISCUSSION AND CONCLUSION

This work aimed to provide guidelines for the use of the genLouvain algorithm, suggesting a proper setting of its temporal resolution parameter ω . To this purpose, we propose a tool able to generate multilayer networks with fully settable parameters, that helped us in testing the algorithm under a variety of conditions. We assessed its performance, by systematically varying ω , by an ANOVA in which we evaluated the effect of noise level, number of layers, number of clusters and percentage of nodes that change modules. We also distinguished the case in which communities are stationary across layers (a) and the one in which they evolve in time (b). For all ω s we tested, the algorithm performs better in networks with low noise level, i.e. with clearly defined communities, and with high number of clusters. In the case (a) ω -values greater or equal to 1 outperform the others, especially in noisy networks with many layers. It is important to notice that even in network with an almost destroyed community structure (no=50%), the algorithm with high ω can recover the imposed communities if it has enough layers to interpolate. In the case (b) ω -values lower than 0.5 make genLouvain perform better in networks with few layers, because they allow it to recognize fast changes. On the other hand, if we expect the network to be noisy or we want to recover different but persistent partitions, higher ω -values (>0.5) are eligible. The factor representing the number of nodes that change allegiance to their modules does not

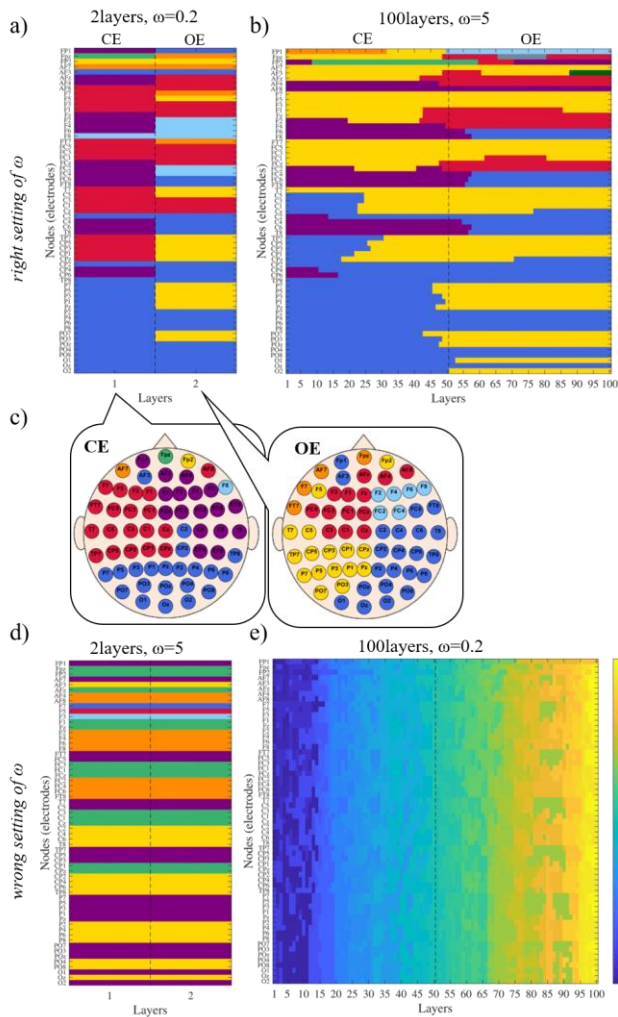


Figure 2. Results of the application of genLouvain to functional brain networks estimated from EEG signals during closed and open eyes. In panel (a)-(b) communities obtained with the right setting of ω , according to the simulation study, are shown in the two cases in which $nL=2$ ($\omega=0.2$) and $nL=100$ ($\omega=5$). X-axis and y-axis represent time and nodes, or channels of EEG, and clusters are identified by colours. Panel (c) reports a projection of the detected communities of panel (a), ($nL=2$, $\omega=0.2$) on a 2D model of scalp (seen from above, with triangles representing noses). The dots are the 61 electrodes, grouped into clusters displayed with the same colours as in panel (a). This topographical representation shows a big cluster involving occipital electrodes during CE, that disappears during OE, where the clusters are all more hemispheric-specific. Panels (d)-(e) report results obtained with a wrong setting of ω , according to the simulation study, in the same two cases in which $nL=2$ ($\omega=5$) and $nL=100$ ($\omega=0.2$). In the first case the algorithm cannot detect a change between CE and OE, while in the second case it shows a huge variability (75 total communities) even among layers relative to the same condition (CE or OE).

seem to impact the algorithm's performances, suggesting that it can detect small as well as big changes in community organization.

We then validated these results applying such algorithm with the same ω -values on EEG-based brain functional multilayer networks estimated from a healthy subject during resting state at CE and OE. To reproduce the conditions of the previous study we estimated 4 multilayer networks composed respectively of [2, 10, 50, 100] layers: the first half relative to CE, and the second half to OE. Results agree with those of the simulation study. In fact, low ω (0.1-0.5) are preferable in networks with few layers in which the algorithm must recover rapid change from CE to OE. Higher

values of ω (>1) are better in networks with many layers and two conditions, as they allow to recover a unique partition for the first one (CE) and another one for the second (OE).

From the topological comparison between the community organization in the two tasks we can say that the closed eyes condition gives rise to a cluster involving the occipital electrodes/brain areas, while in the open eyes condition the clusters are more hemispheric-specific. This result is consistent with what one would expect. In fact, it is well known that during resting state at closed eyes there is an increase of alpha-rhythm coming from the occipital regions, that disappears as soon as the subject open his/her eyes. In conclusion, we show how an appropriate selection of ω -parameter can allow to use genLouvain algorithm to accurately detect the stationary as well as the dynamic properties of brain functional networks.

REFERENCES

- [1] O. Sporns, «The human connectome: a complex network», *Ann. N. Y. Acad. Sci.*, vol. 1224, n. 1, pagg. 109–125, 2011.
- [2] O. Sporns e R. F. Betzel, «Modular Brain Networks», *Annu. Rev. Psychol.*, vol. 67, pagg. 613–640, 2016.
- [3] M. Chavez, M. Valencia, V. Navarro, V. Latora, e J. Martinerie, «Functional Modularity of Background Activities in Normal and Epileptic Brain Networks», *Phys. Rev. Lett.*, vol. 104, n. 11, mar. 2010.
- [4] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, e J.-P. Onnela, «Community Structure in Time-Dependent, Multiscale, and Multiplex Networks», *Science*, vol. 328, n. 5980, pagg. 876–878, mag. 2010.
- [5] «NetWiki : Gen Louvain browse». [In linea]. Available at: <http://netwiki.amath.unc.edu/GenLouvain/GenLouvain>. [Consultato: 29-gen-2019].
- [6] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, e E. Lefebvre, «Fast unfolding of communities in large networks», *J. Stat. Mech. Theory Exp.*, vol. 2008, n. 10, pag. P10008, ott. 2008.
- [7] Y.-R. Lin, Y. Chi, S. Zhu, H. Sundaram, e B. L. Tseng, «Facetnet: a framework for analyzing communities and their evolutions in dynamic networks», in *Proceeding of the 17th international conference on World Wide Web - WWW '08*, Beijing, China, 2008, pag. 685.
- [8] M.-S. Kim e J. Han, «A particle-and-density based evolutionary clustering method for dynamic networks», *Proc. VLDB Endow.*, vol. 2, n. 1, pagg. 622–633, ago. 2009.
- [9] M. Girvan e M. E. J. Newman, «Community structure in social and biological networks», *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, n. 12, pagg. 7821–7826, giu. 2002.
- [10] C. Granell, R. K. Darst, A. Arenas, S. Fortunato, e S. Gómez, «Benchmark model to assess community structure in evolving networks», *Phys. Rev. E*, vol. 92, n. 1, pag. 012805, lug. 2015.
- [11] M. G. Puxeddu, M. Petti, F. Pichiorri, F. Cincotti, D. Mattia, e L. Astolfi, «Community detection: Comparison among clustering algorithms and application to EEG-based brain networks», in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pagg. 3965–3968.
- [12] L. A. Baccalá e K. Sameshima, «Partial directed coherence: a new concept in neural structure determination», *Biol. Cybern.*, vol. 84, n. 6, pagg. 463–474, mag. 2001.
- [13] W. Klimesch, «EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis», *Brain Res. Rev.*, vol. 29, n. 2, pagg. 169–195, apr. 1999.
- [14] D. Yasumasa Takahashi, L. Antonio Baccal, e K. Sameshima, «Connectivity Inference between Neural Structures via Partial Directed Coherence», *J. Appl. Stat.*, vol. 34, n. 10, pagg. 1259–1273, dic. 2007.
- [15] E. Niedermeyer, «Alpha rhythms as physiological and abnormal phenomena», *Int. J. Psychophysiol.*, vol. 26, n. 1, pagg. 31–49, giu. 1997.