E-Health Project
Phase-amplitude coupling events in the EEG

Gulzhanat Dandyyeva - 2016488
Pietro Talli - 2021427

June 29, 2021

Abstract

The analysis of the methods for Phase Amplitude Coupling (PAC) measurement in a dataset of 64-channel electroencephalography (EEG) signals is given in this report. Mean Vector Length Modulation Index (MVL) and Kullback-Leibler Modulation Index (KLmi) were applied to extract coupling measures between nodes and then results from both methods were compared through Robust estimation of phase-amplitude coupling (REPAC). Afterwards, a characterization of the network is given and limitations of the given methods were assessed.

Index Terms—PAC, Mean Vector Length Modulation Index (MVL), Kullback-Leibler Modulation Index (KLmi), Robust estimation of phase-amplitude coupling (REPAC)

1   Introduction

Recent findings show that electrical activity of the brain can be described by the dynamics of multiple oscillations which are initiated by spatially distributed groups of neurons that interact with each other [6]. Generally, the brain activity is represented through neural recordings that show the local network in the time domain. These recordings reveal that interactions between the oscillations of neuronal entities can be translated into cross frequency couplings (CFC) [1]. Besides, intensities of CFC differ in accordance with a given state. This means that CFC can be interpreted as a signal providing information on the physiological level of neuronal activity. In other words, CFC can be used as a tool for detection of brain pathologies, including those recorded through EEG .

The mechanism of the signal processing is based on the fact that the amplitude and the phase of a frequency band interact with other bands in a signal and this leads to CFC events. Although several signal processing algorithms have been developed to identify and quantify CFC from noisy neural signals, there are still challenges related to interpretation of this phenomena. The reason is explained by the analysis of spatially sparse, one dimensional recordings that may not be explicit for understanding complex neural dynamics [6].

There are three commonly applied types of CFC: amplitude-amplitude coupling (AAC), phase-phase coupling (PPC) and phase-amplitude coupling (PAC). The focus of this project is going to be PAC. Based on previous findings, PAC may be involved in information processing and it has a potential to be used as a biomarker for epilepsy detection [5]. For example, PAC patterns including different frequency bands are suggested to use in order to identify the seizure onset zone (SOZ) and to predict the seizure dynamics.

The fundamental mechanism in PAC is described by the istantaneous amplitude of a higher frequency band within a signal is modulated by the immediate phase of a lower-frequency band of the same signal or correlated one. The measures frequently used in the PAC investigations are the Mean Vector Length Modulation Index (MVLmi) [1], the Kullback-Leibler Modulation Index (KLmi) [3], and the General Linear Model Modulation Index (GLMmi). It should be noted that none of the methods have been standardized.

1.1   Mean Vector Length Modulation Index (MVLmi)

Canolty et al.[1] introduced a new method Mean Vector Length Modulation Index (MVL) to measure the phase-amplitude coupling. This method involves an extraction of the phase from the low frequency

filtered analytic signal as well as amplitude - from the high frequency filtered analytic signal. MVL is called a direct PAC estimator as it uses phase angle and magnitude of each complex number of the corresponding signal in order to estimate the degree of coupling. In turn, each complex value in the time series is considered as a vector in the polar plane.

Once the magnitude M of a fragment of all vectors is high at a particular phase the phase-amplitude coupling is considered to be present. Sequentially, averaging operation on vectors gives a mean vector with a specified phase and length. The degree of phase-amplitude coupling is characterized by the length of this vector, while the mean phase with a highest amplitude is shown by direction of the vector. In the case that no coupling occurs, then complex values counteract with each other and the mean vector becomes short and its direction does not give any information [7].

There are constraints applied to the method, such as the value depends on the absolute amplitude of the high-frequency oscillation. Also, outliers of the amplitude may substantially affect the MVL and phase angles are often not distributed in a uniform way [10].

The formula for MVL is given below (Figure 1) and it relies on creation of the composite values $z(t)$ in the complex plane, generated by combining the phase of the lower frequency signal ($\phi(t)$) and amplitude of the higher frequency signal ($A(t)$). The complex vector $z(t)$ is defined as:

$$z(t) = A(t)e^{j\phi(t)} \tag{1}$$

Thus, a distribution of the composite vector $z(t)$ in the complex plane is analyzed as described before.

$$MVLmi = \left| \frac{1}{N} \sum_{t=1}^{N} z(t) \right| \tag{2}$$

1.2   Kullback-Leibler Modulation Index (KLmi)

The disadvantage of the MVL method was emphasized by Tort et al. [3]. The issue with MVL is that this method is based on the assumption that coupling can be presented only at one preferred phase leading to worse performance for more complex coupling relationships. In order to overcome this limitation, a new method based on the Kullback-Leibler Modulation Index (KLmi) has been introduced.

This method relies on the study of the density distributions of the two vectors $A(t)$ and $\phi(t)$. To quantify the coupling, the Mutual information is used. Mutual Information is an information theory measure which is used to asses how much dependant two random variables are. Given two r.v. X and Y, the mutual information is defined as

$$I(X,Y) = E_{X,Y}\left[\log \left(\frac{p_{XY}(x,y)}{p_X(x) \cdot p_Y(y)} \right) \right]. \tag{3}$$

where $p_{XY}(x,y)$ is the joint distribution and $p_X(x) \cdot p_Y(y)$ is the product of the marginals. Another way to express the mutual information is its representation in terms of Kullback Liebler divergence D_{KL},

$$I(X,Y) = D_{KL}(p_{XY(x,y)}||p_X(x) \cdot p_Y(x)), \tag{4}$$

which highlights the fact that the mutual information measures the distance from the joint distribution and the distribution under the assumption of independence between X and Y. In [3] Tort proposes to use mutual information for inferring a relationship between $A(t)$ and $\phi(t)$ and thus quantify the PAC. More specifically he proposed to divide the complex space in a number J of phase bins and obtain the joint probability distribution

$$p_z(j) = \frac{\langle A(t) \rangle_\phi(j)}{\sum_{k=1}^{N} \langle A(t) \rangle_\phi(k)} \tag{5}$$

where $\langle A(t) \rangle_\phi(j)$ is the average of the amplitude values belonging to the j-th phase bin and $\sum_{k=1}^{N} \langle A(t) \rangle_\phi(k)$ is a normalization term so that $\sum_{k=1}^{N} p_z(k) = 1$. In the case of uniform distribution of the vector $z(t)$ among the J phase bins we obtain a probability

$$p_u(j) = \frac{1}{J} \tag{6}$$

The mutual information between the joint distribution of $A(t)$ and $\phi(t)$ and the uniform (independent) case is

$$MI = D_{KL}(p_z(j)||p_u(j)) = \sum_{j=1}^{N} p_z(j) \log\left(\frac{p_z(j)}{p_u(j)}\right) = \log(J) + \sum_{j=1}^{N} p_z(j) \log(p_z(j)) \qquad (7)$$

The actual measure proposed by Tort in [3] is a normalized version of equation (7) so that the actual PAC value is normalized in a range between 0 and 1

$$MI = \frac{\log(J) + \sum_{j=1}^{N} p_z(j) \log(p_z(j))}{\log(J)} \qquad (8)$$

This way to estimate the mutual information tries to estimate a discrete probability distribution starting from the realizations of the vectors $A(t)$ and $\phi(t)$ which are considered stationary during the period of time taken into account [3].

2 Methods

In order to compare the two estimators of the modulation index ($MVLmi$ and $KLmi$), the coupling measures have been extracte through a standard comodulogram pipeline and a more novel technique proposed by Cisotto [4] , which is called Robust estimation of phase-amplitude coupling (REPAC).

2.1 The standard comodulogram method

This procedure to extract the PAc measures is based on the definition of two frequency vectors: one for the high frequency component and one for the low frequency one.

The idea is to estimate the PAC for all the possible combination of high and low frequency and to keep the highest value as the final and most reliable PAC measure. Given F_H containing N_H high frequency values and F_L containing N_L low frequency values, the comodulagram is the $N_H \times N_L$ matrix containing the PAC values computed for all the combination of the frequencies. In the project F_H and F_L vectors used were: $F_H = 13 : 5 : 50$ Hz whereas for the low frequencies $F_L = 4 : 12$ Hz. It is important to note that this method is computationally costly, since to estimate the PAC value $8 \times 9 = 72$ computation of the measure must be obtained. Also, this choice of F_H limits the estimation of the high frequency component with a precision of 5Hz.

2.2 Robust estimation of phase-amplitude coupling (REPAC)

The robust estimation of phase-amplitude coupling (REPAC) is a fully data-driven method to estimate the frequencies at which a PAC event is occurring. The method, proposed by Cisotto [4], is used in this project in order to test for differences in the extraction of PAC coupling measures. This procedure uses the AM (Amplitude Modulation) theory to estimate both the low and high frequencies of a PAC event in a EEG signal. However, for the purpose of this project, the original procedure has been modified to allow to estimate these high and low frequencies starting from two signals and not within the same signal. In this way, the REPAC algorithm is able to estimate the low frequency of the PAC event from one signal, whereas the envelope of the high frequency oscillations and thus the estimation of the high frequency is extracted from the second signal. It is worth noting that, while in the comodulogram procedure firstly all the PAC values are computed and only at the end the modulation index and the center band frequencies are estimated, in the REPAC procedure in the first step the center band frequencies are estimated and then the PAC value is computed only for the selected frequencies. This increases a lot the performance of the extraction of the PAC measures and also allow for better precision in the estimation of the high frequency. The aim of this project is to compare the MVLmi and KLmi PAC estimators, however results for both the comodulogram and REPAC procedure are given and an evaluation of the differences will be provided in the the discussion of the results.

3 A comparison of the two PAC estimators

To provide a complete description of the differences of the two methods, the MVLmi has been used adding a normalization term to equation 1.1
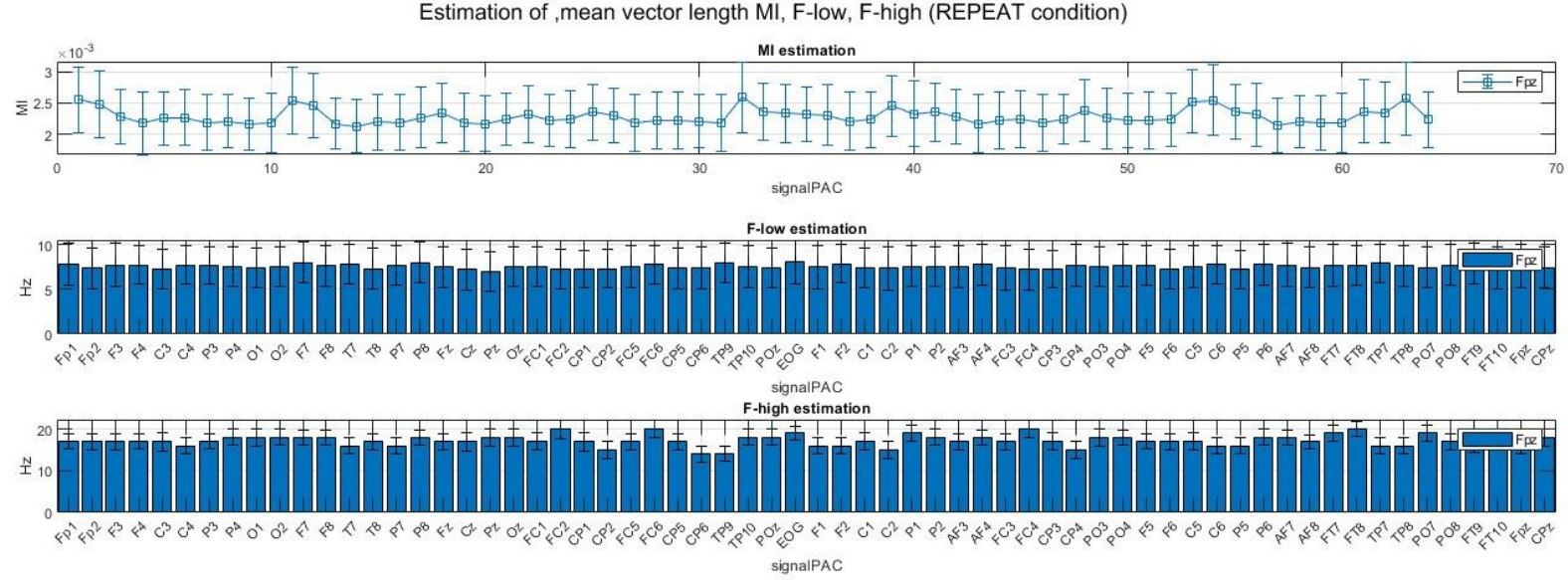
Figure 1: MVLmi values for the 'Fpz' electrode as modulation signal from which the phase vector has been extracted and all the 64 electrodes as modulated signals.

$$MVLmi = \frac{1}{N} \left| \frac{\sum_{t=1}^{N} A(t)e^{j\phi(t)}}{\sqrt{\sum_{t=1}^{N} A(t)^2}} \right|, \tag{9}$$

this way the MVLmi varies in a range between 0 and 1 as it is for the KLmi and an effective comparison of the values is possible. Also this normalization avoid the MVL value to be influenced by a high power in the high frequency band.

3.1 A theoretical comparison of the two methods

In table 3.1 a comparison based on the definition and the mathematical description of the two PAC estimators used in the project has been reported.

Measure	Advantages	Disadvantages
MVLmi	fast to compute	the coupling must happen at a preferred phase value
KLmi	can capture more complex dependancies	some bias introduced by the choice of the number of phase bins J

Table 1: Theoretical comparison of the two PAC estimators.

3.2 Results from the standard comodulogram processing procedure

In this section the results coming from the standard PAC processing procedure are reported.

In figure 1 and 2 the modulation index measures and the high and low center band frequency are reported. For the 64 electrodes of the dataset, the measures corresponding to the condition of the 'Fpz' signal as modulating signal is reported. The purpose of this choice is to compare the results in the case of an electrode which is typically used as a reference. Since the values here are obtained averaging among the epochs of the dataset, all the measures are reported with the standard deviation. It is possible to see
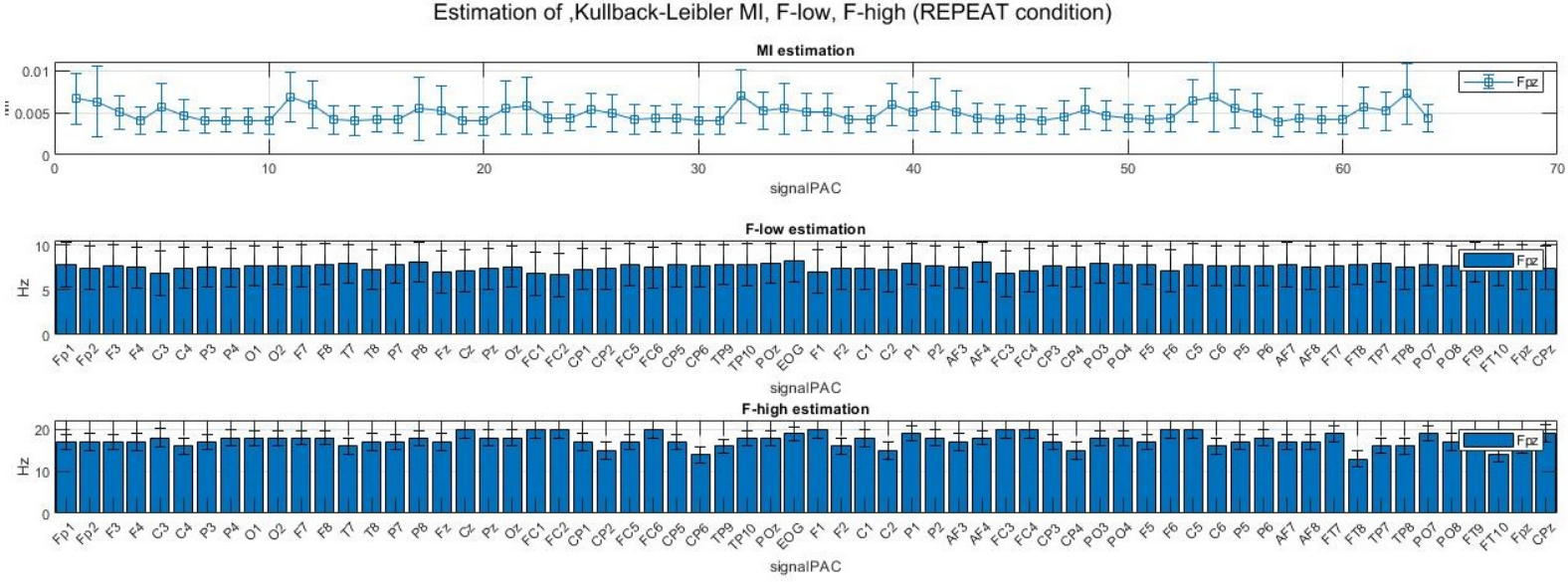
Figure 2: KLmi values for the 'Fpz' electrode as modulation signal from which the phase vector has been extracted and all the 64 electrodes as modulated signals.

that the modulation index (MI) values are highlighting the same behaviour with both the estimators, showing higher values in correspondence of the same electrodes. As for the center band frequencies, there are some differences mostly in the high frequency estimation. In this comparison the case of the 'Fpz' electrode is considered, but these similarities in the MI values occurs for all the 64 electrodes.

This shows that the two methods, despite having different mathematical description are able to capture signals coupling in a similar way.

Another possible comparison which helps to understand the statistics of the measures extracted with the two different methods is to compare the distributions of the values. In figure 3 and 4, the distributions of the average values are reported. In this case it is possible to note some differences between the two PAC estimators. The MVLmi measure (figure 3) shows that most of the values concentrates in the range $[2.1 \times 10^{-3}, 2.4 \times 10^{-3}]$ and only few values higher than 2.4×10^{-3}. On the other side, in figure 4 the distribution seems to have a lot of values concentrated on a low value range of $[3.8 \times 10^{-3}, 4.7 \times 10^{-3}]$.

It is worth noting that the fact that most of the values are concentrated in a small values range is something expected. Since these values are entries of an adjacency matrix we expect to have the majority of edges that are weak connections represented by lower coupling values. In general it is possible to say that the low values set a baseline on the connections between two nodes, but the majority of the information about the network is concentrated on the higher values through which it is possible to discriminate communities and other network parameters. It is possible to say that in the KLmi distribution there is a more clear distinction between the baseline and the stronger connections and this could lead to different results in the network analysis.

3.3 Network parameters

In table 2 the network parameters (average degree and γ value) are reported for both the measures and both PAC extraction procedures. It is worth noting that the MVLmi measure always obtains a higher average degree with respect to the KLmi. This can be due to the fact mentioned in section 3.2 about the shape of the values distributions. This means that in general the MVLmi network results to be more connected and obtains a higher average degree. As for the γ values, in every cases there is a score higher than 3 meaning that every network resembles a random network model. However it is possible to say that in the KLmi case the network better represents the "small-world" network model which is known to be the most suitable model for a brain network [9].
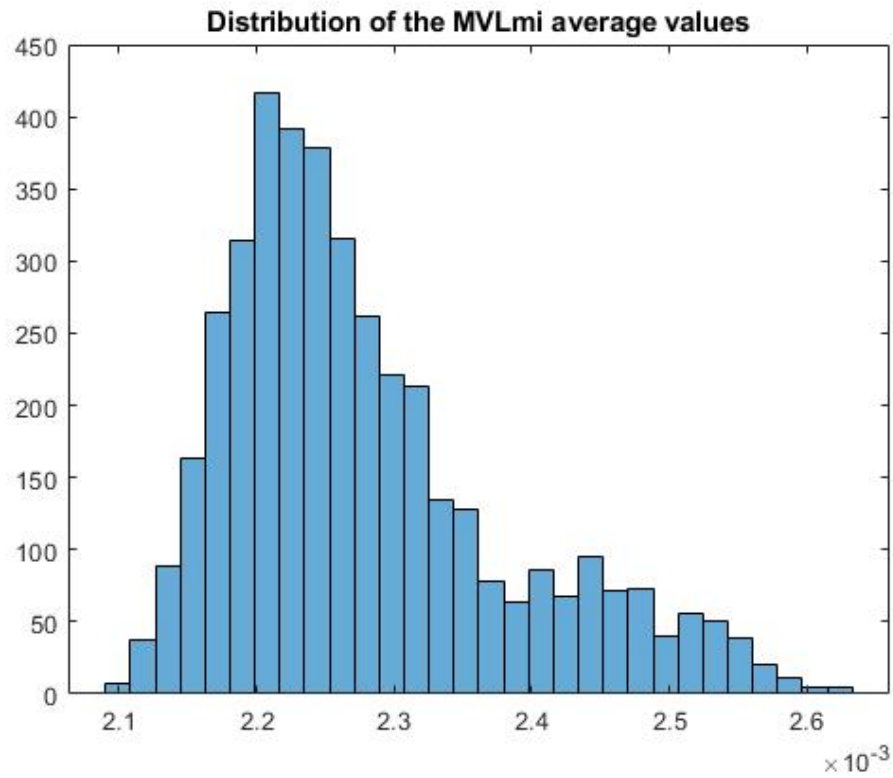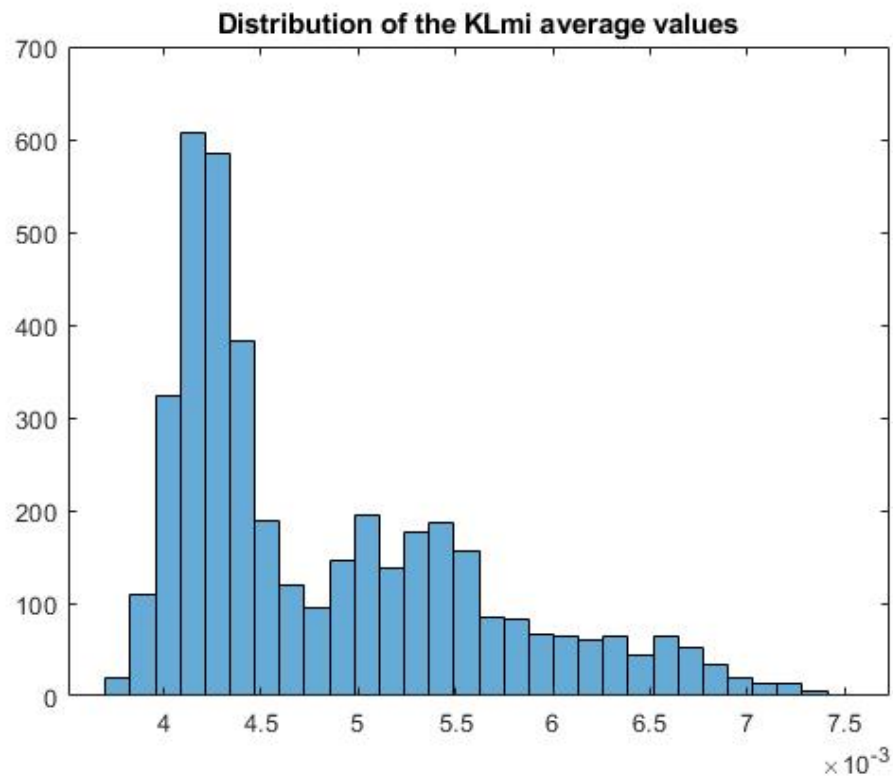
Figure 3: Distribution of the MVLmi values.

Figure 4: Distribution of the KLmi values.

Measure	Average degree	γ parameter
Comodulogram method		
MVLmi	55.40	5.79
KLmi	43.26	3.12
REPAC method		
MVLmi	42.35	5.52
KLmi	34.27	3.34

Table 2: Network parameters.

4 Network Results

The last analysis that was performed is the community detection to characterize and evaluate differences in the networks. The algortihm used to perform community detection is a very fast and common clustering algorithm which is the Louvain algorithm, based on the maximization of the modularity. A partition has been obtained for every epoch of the dataset condition (REPEAT and SWITCH) and finally the consensus clustering among the different epochs partitions' was used to get a single (and stable) community structure. The total number of netowrks obtained after the community detection procedure are 8: 2 measures × 2 extraction methods × 2 conditions in the dataset. The results are shown in figure 5.
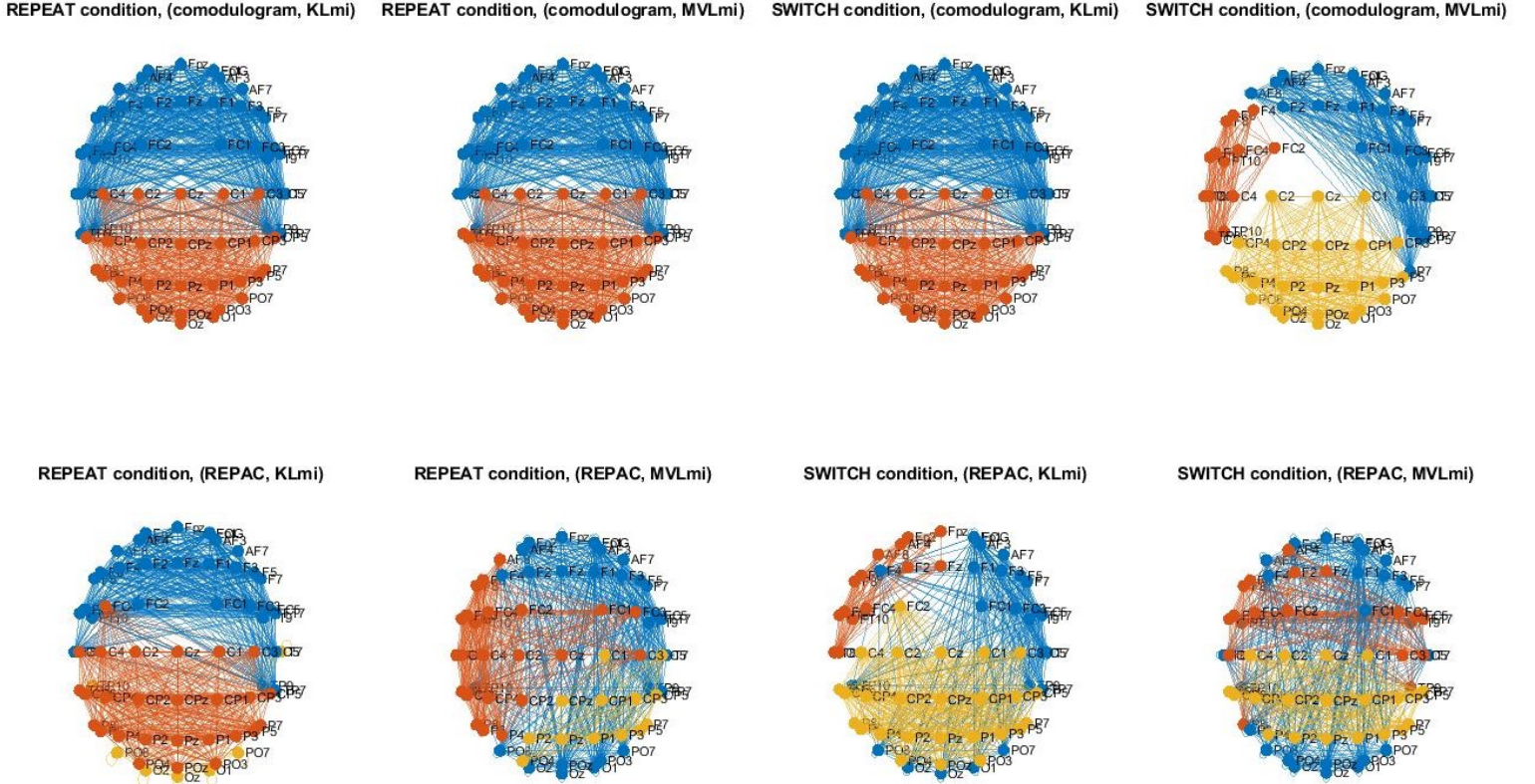

Figure 5: Network results from the community detection procedure.

4.1 Further analysis about networks results

To numerically evaluate the differences of PAC estimators, the normalized variation of information (NVI) was used to measure the distance between the epochs of the same estimator in the same condition. The NVI is an information theory measure which aims at measuring the distance between two clustering solutions. Given two partitions X and Y, the NVI is given by

$$NVI(X,Y) = -\sum_{i,j} r_{ij} \left[\log\left(\frac{r_{ij}}{p_i}\right) + \log\left(\frac{r_{ij}}{q_j}\right) \right], \tag{10}$$

where r_{ij} is the fraction of nodes belonging to i-th community in the partition X and to the j-th community in the partition Y, p_i is the fraction of nodes belonging to the i-th community in partition X and q_j is the fraction of nodes belonging to the j-th community in the partition Y.

The NVI was computed between every couple of partitions inside the same condition and the same measure, thus obtaining $n_{epochs} \times n_{epochs}$ values. Then the sum of these values was normalized by the squared number of epochs in the condition in order to obtain comparable values.

Measure	REPEAT condition	SWITCH condition
Comodulogram method		
MVLmi	0.942	0.931
KLmi	0.923	0.921
REPAC method		
MVLmi	0.962	0.962
KLmi	0.953	0.958

Table 3: Total NVI for the 8 networks.

In table 3, the total NVI for the 8 networks is reported. It is worth noting that the smaller this value is, the more similar the partitions among the various epochs. With this analysis it is possible to test for the ability of the estimator to extract couplings which underlie the same network structure and so the ability to better characterize the condition. In general it is possible to see that the NVI is quite high in all the cases, meaning that the algorithm struggles to find a recurrent pattern and to obtain small NVI values. However it is possible to observe that the KLmi obtains always slightly smaller NVI in both the comodulogram/REPAC procedure and in both REPEAT/SWITCH conditions.

5 Discussion

Based on the results obtained, there are important aspects to be discussed. Firstly, it can be interpreted that for the standard comodulogram processing the methods MVLmi and KLmi generated similar patterns in estimation of mean vector length MI. This means, that despite computational differences, both methods provide similar signal coupling results. Nevertheless, it should be noted that the computational speed for KLmi is slower compared to MVLmi method, because of its formulation. This feature can be considered as the main limitation of the KLmi measurements. The underlying reason for slower execution is a complexity in algorithm, which is given by O(N * log(N)) for KLmi, whereas for MVL the complexity is given by O(N).

The second aspect is related to the distribution levels described in Figures 3 and 4 showing that KLmi gives a distinct concentration of higher value components. This means there is a higher possibility to separate higher values from lower ones and form clusters, leading to more meaningful results. Thus, KLmi is a preferable method for the search of network communities as compared to MVLmi in terms of performance.

The third aspect is related to the results generated through REPAC method, that shows the network parameter for the average degree is significantly higher for MVLmi in comparison with KLmi. This can be explained by a higher number of connections formed between nodes in MVLmi method and it could be associated with its distribution levels. For both methods, another network parameter representing power law γ is higher than 3. Thus, all networks can be considered as random network models. Overall, as for the comparison of the comodulogram and REPAC processing methods, there are slight differences in network parameters, but the trend of the values is conserved giving higher average degree for MVLmi in both cases.

The last aspect is associated with the networks obtained from community detection procedure. The figure 5 demonstrates that REPAC , apparently, can give more insights about the differences in the networks, although there is a significant level of overlapping communities. As for the differences in communities for comodulogram KLmi and MVLmi, the REPEAT condition has not shown any significant differences, while there is a different number of communities is detected for SWITCH condition. However, these results can be considered as insufficient to recognize and differentiate SWITCH and REPEAT conditions in network analysis. In turn, this characteristics is another limitation of the project and further work should be done related to use of the classification algorithms in identifying conditions.

6 Conclusion

In this work a comparative analysis of the Mean Vector Length Modulation Index (MVLmi) and the Kullback-Leibler Modulation Index (KLmi) methods was conducted based on standard comodulogram and REPAC processing procedures. The results show better performance for KLmi method in terms of the possibility to separate clusters for community detection, whereas MVLmi method can be useful in terms of higher computational speed. However, the trade-off between performance and computational complexity for KLmi measurements should be taken into consideration. In conclusion the PAC estimation has shown much slower execution speed while REPAC method turned out to be advantageous in terms of computation time. Although, as for the outcome in the network conditions differentiation, both methods have not shown significant results.

References

[1] Canolty RT, Edwards E, Dalal SS, et al. " High gamma power is phase-locked to theta oscillations in human neocortex". *Science*. 2006;313(5793):1626-1628. doi:10.1126/science.1128115

[2] Onslow AC, Bogacz R, Jones MW. " Quantifying phase-amplitude coupling in neuronal network oscillations". *Prog Biophys Mol Biol*. 2011 Mar;105(1-2):49-57. doi: 10.1016/j.pbiomolbio.2010.09.007. Epub 2010 Sep 30. PMID: 20869387.

[3] Tort, Adriano B. L. et al. " Measuring phase-amplitude coupling between neuronal oscillations of different frequencies." *Journal of neurophysiology* vol. 104,2 (2010): 1195-210. doi:10.1152/jn.00106.2010

[4] Cisotto, Giulia." REPAC: Reliable estimation of phase-amplitude coupling in brain networks." *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021.

[5] De Hemptinne, Coralie, et al. " Exaggerated phase–amplitude coupling in the primary motor cortex in Parkinson disease." *Proceedings of the National Academy of Sciences* 110.12 (2013): 4780-4785.

[6] Dellavalle, Damián, et al. " Two types of ictal phase-amplitude couplings in epilepsy patients revealed by spectral harmonicity of intracerebral EEG recordings." *Clinical Neurophysiology* 131.8 (2020): 1866-1885.

[7] Hülsemann, M. J., Naumann, E., Rasch, B. (2019). " Quantification of phase-amplitude coupling in neuronal oscillations: comparison of phase-locking value, mean vector length, modulation index, and generalized-linear-modeling-cross-frequency-coupling" . *Frontiers in neuroscience*, 13, 573.

[8] Martínez-Cancino, R., Heng, J., Delorme, A., Kreutz-Delgado, K., Sotero, R. C., Makeig, S. (2019). " Measuring transient phase-amplitude coupling using local mutual information". *NeuroImage*, 185, 361-378.

[9] Bullmore, E. and Sporns, O., 2009. "Complex brain networks: graph theoretical analysis of structural and functional systems." *Nature reviews neuroscience*, 10(3), pp.186-198.

[10] Cohen, M. X. (2014). " Analyzing neural time series data: theory and practice". *MIT press*.