

# Multi Cross-correlation Analysis in a Multi-channel EEG applied in Motor Activity (Real/Imaginary)

Fernando Ferraz Ribeiro<sup>1,4,5</sup>, Andréa de Almeida Brito<sup>3</sup>, Florêncio Mendes Oliveira Filho<sup>1,5</sup>, Juan Alberto Leyva Cruz<sup>2</sup>, Gilney Figueira Zebende<sup>2\*</sup>✉

**1** Earth Sciences and Environment Modeling Program, State University of Feira de Santana, Bahia, Brazil

**2** State University of Feira de Santana, Bahia, Brazil

**3** Federal Institute of Education, Science and Technology, Bahia, Brazil

**4** Federal University of Bahia, Brazil

**5** SENAI CIMATEC University Center, Salvador, Bahia, Brazil

✉Current Address: Earth Sciences and Environment Modeling Program, State University of Feira de Santana, Bahia, Brazil

\* gfzebende@uefs.br

## Abstract

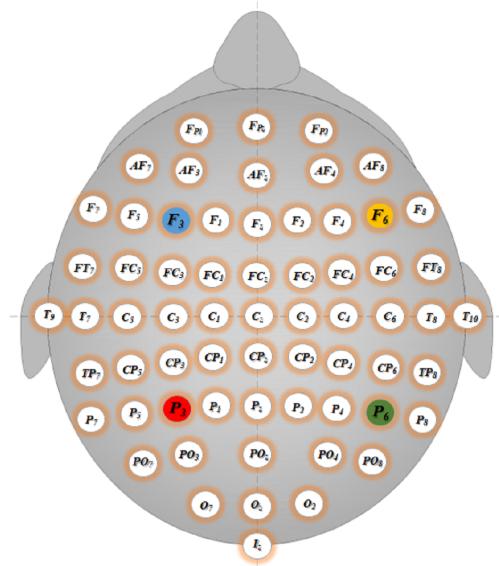
In this paper, we use the detrended multiple cross-correlation coefficient to analyze time-series data recorded in an EEG experiment, where 109 subjects performed 4 Tasks involving real and imaginary motor activities. Four specific channels were selected, two of them on the frontal portion of the scalp and two on the parental region. As a result, the multiple cross-correlation coefficient identifies that a unique signature for each subject appears in all of the performed tasks; there are no significant differences between imaginary and real tasks for each subject neither in the global mean. In the global statistics for all the subjects, the coefficient also identified that de frontal channels have greater correlation values than the parental ones and smaller standard deviations, especially in time scales around 0.42s.

This research presents a statistical tool that has never been applied to EEG signals analysis, and it's also the first time that the  $DMC_x^2$  is used with four time-series (one dependent and three independent). Moreover, the method was applied to a substantial number of subjects and produced quality results (figures, movies and tables). Therefore, this study probably will be the kick-off of a new approach to analyzing multiple cross-correlations in EEG.

## Introduction

The Electroencephalogram, EEG, is a test that evaluates the electrical activity of the brain. It is a normally non-invasive method, where tiny electrodes are placed on the scalp to track brain wave patterns and send those signals to a computer. The EEG equipment measures the electric potential difference (in  $\mu$ V), from each one of the 64 electrodes to a reference electrode, usually placed in the ear lobe. These impulses are amplified and recorded over time, generating time series for each sensor (with high temporal resolution and low spatial resolution). The international 10 – 10 system was used to map the positions where 64 electrodes were fixed in the conduction of the experiments that collected the data used in this analysis. The nomenclature of the

channels in this system is given according to the region in which they are located, namely: *F* (frontal), *T* (temporal), *C* (central), *P* (parietal), and *O* (occipital).  
 11  
 12  
 13  
 14  
 Channels located on the midline are indexed by the letter *z* (zero), the channels located  
 on the left side are represented by odd indexes, while those on the right side are  
 represented by even indexes (see Fig. 1). The EEG analysis is useful for diagnosing



**Fig 1.** Electrodes positions based on international 10 – 10 system for 64 channel. The circles  $F_3$  (blue),  $F_6$  (yellow),  $P_3$  (red), and  $P_6$  (green) identify the channels used for multiple cross-correlation analysis.

problems such as epilepsy, dementia and sleep disorders among others. The diagnosis normally focuses on the spectral content of the EEG, and on the type of neural oscillations (or brain waves) that can be observed in these signals. Most observed signals are between 1 and 20 Hz. Although the EEG is almost a centenary technique, in recent decades EEG has addressed new problems as brain-triggered neuro-rehabilitation treatments, experimental psychology or even computational neuroscience, due to versatility and accessibility, alongside the advances in signal processing [1] that enhance the possibilities of analyses in the field. Thinking about new techniques for analyzing EEG signals, we highlight, the  $F_{DFA}$  root mean square fluctuation function [2], to study the brain activity in the reading task [3], the quantification of long-range correlation of EEG signals [4], and recently in the statistical study of the EEG cross-correlation of two signals using the  $\rho_{DCCA}$  [5].

In this paper, we will analyze the multiple signs of EEG produced in a Motor Activity (Real/Imaginary) for 109 subjects by the Detrended Multiple Cross-Correlation Coefficient ( $DMC_x^2$ ) [6]. Based on symmetry criteria, we choose the channels  $F_3$ ,  $F_6$ ,  $P_3$  and  $P_6$  to apply this new statistical tool. For this purpose, in the sections below we present the Data and Method with dataset, methodology and calculations used to analyze the data, including pre-processing strategies. After we present the Results and Discussion with its statistics and discussion. Finally, the Conclusion is made.

## Data and Method

### Data

Our data was obtained from the *Physionet* databank [7], publicly available at (with Open Data Commons Attribution License (ODC-By) v1.0):

<https://physionet.org/content/eegmmidb/1.0.0/>

This databank presents a large quantity/quality of EEG experiments' recordings, available in EDF (European Data Format) file, containing the series for all the 64 channels of the international 10 – 10 system, using a Brain-Computer Interface Technology called BCI-2000 [8]. The data originated from a study that perform 14 experiments on a population of 109 subjects to record the brain signals under various Tasks, briefly described below.

The first 2 activities are baseline references, the subjects were resting with eyes opened and the second with eyes closed (one minute for each). The other four activities (Tasks) are a combination of two categories with two possible options each. In general, the experiments consist of making the subjects react to visual stimuli, a target that appears on a screen. One category is about the target position, i.e., one option is a target that will appear on the **Left/Right** of the screen and the other, the target will appear on the **Top/Down** of the screen. The second category determines if the subject will actually execute a body movement related to the target position (**Real**) or if the corresponding action will be just imagined (**Imaginary**). All the experiments for all 109 subjects were downloaded and used in this research. These Tasks, with an approximate duration of two minutes, are better described just below:

Every subject executes four Tasks three times (experiments). Table 1 summarizes these Tasks by experiment number.

**Table 1.** The experiment number and the activity executed: Two one-minute baseline runs (one with eyes open, one with eyes closed), and three two-minute of the four Tasks.

| Activity\Experiment      | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|--------------------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|
| Baseline 1 (eyes open)   | 0 |   |   |   |   |   |   |   |   |    |    |    |    |    |
| Baseline 2 (eyes closed) | 0 |   |   |   |   |   |   |   |   |    |    |    |    |    |
| Task 1: Real (L/R)       |   |   | X |   |   |   | X |   |   |    | X  |    |    |    |
| Task 2: Imag (L/R)       |   |   |   | X |   |   |   | X |   |    |    | X  |    |    |
| Task 3: Real (T/D)       |   |   |   |   | X |   |   |   | X |    |    |    | X  |    |
| Task 4: Imag (T/D)       |   |   |   |   |   | X |   |   |   | X  |    |    |    | X  |

After introducing our database, we will proceed to write about the methodology for multiple cross-correlation analyses.

### DCCA multiple cross-correlation methodology [6]

The coefficient  $DMC_x^2$  is part of a new statistical tool to analyze non-stationary time-series in multiple applications. It all starts with the DFA method [9], which was proposed to identify self-affinity in a single time-series. If this time-series has a long-memory behavior, then the fluctuation function,  $F_{DFA}$ , scaling with a power-law as a function of the time scale  $n$  by the auto-correlation exponent  $\alpha_{DFA}$ . This method has already been applied to evaluate the self-affinity in EEG signals in [2], where it is concluded that the amplitude of fluctuation tends to be larger in the frontal channels ( $F_3$  and  $F_6$ ), if compared with the channels located in the parietal region of the brain ( $P_3$  and  $P_6$ ). In terms of brain disorder, its generalization (DCCA method [10]), was

71 applied to analyze: different physiological and pathological states of epilepsy EEG  
 72 signals [11], the change of both synchronization and oscillation in EEG of Alzheimer  
 73 disease patients [12], among other studies [3–5]. Next, we will discuss the multiple  
 74 cross-correlation coefficient, with a brief introduction and some applications.

75 The Detrended Multiple Cross-Correlation Coefficient ( $DMC_x^2$ ) [6] is a  
 76 generalization of the  $\rho_{DCCA}$  [13] (widely known as a robust statistical tool [14]), and  
 77 calculates the multiple cross-correlation of one time-series  $\{Y\}$  (defined as a dependent  
 78 variable) in relation of a number  $k$  of time-series  $\{X_1\}, \{X_2\}, \{X_3\}, \dots, \{X_k\}$  (defined  
 79 as an independent variable). Its mathematical expression is written as:

$$79 DMC_x^2 \equiv \rho_{Y,X_i}(n)^T \times \rho^{-1}(n) \times \rho_{Y,X_i}(n) \quad (1)$$

80 The  $DMC_x^2$  ranges from zero (0), where the time-series do not correlate, to one (1)  
 81 meaning that the time-series have either a perfect correlation or a perfect  
 82 anti-correlation. In this equation,  $\rho^{-1}(n)$  represent the inverse matrix of all possible  
 83 combinations of  $\rho_{DCCA}$  between the independent variables, in other words:

$$\rho^{-1}(n) = \begin{pmatrix} 1 & \rho_{X_1, X_2}(n) & \rho_{X_1, X_3}(n) & \dots & \rho_{X_1, X_k}(n) \\ \rho_{X_2, X_1}(n) & 1 & \rho_{X_2, X_3}(n) & \dots & \rho_{X_2, X_k}(n) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{X_k, X_1}(n) & \rho_{X_k, X_2}(n) & \rho_{X_k, X_3}(n) & \dots & 1 \end{pmatrix}^{-1} \quad (2)$$

84 The main diagonal in eq. 2 is equal 1 (one) because  $\rho_{X_a, X_a}$  between the time series  
 85  $\{X_a\}$  with itself has the maximum value. By definition,  $\rho_{X_a, X_b} = \rho_{X_b, X_a}$ , and this  
 86 makes the matrix symmetric concerning its main diagonal. Also,

$$86 \rho_{Y,X_k}(n)^T = [\rho_{Y,X_1}(n), \rho_{Y,X_2}(n), \dots, \rho_{Y,X_k}(n)] \quad (3)$$

87 represent the transposed vector of the  $\rho_{Y,X_i}(n)$  between the depended variable  $\{Y\}$  and  
 88 each  $\{X_i\}$  independent variable. The DCCA multiple cross-correlation coefficient has  
 89 been applied in many problems, specifically for three variables, i.e, one  $\{Y\}$  (dependent  
 90 variable) and two  $\{X_1\}, \{X_2\}$  (independent variables). Examples of these applications  
 91 can be found applied to meteorological variables [15], in statistical test [16], with the  
 92 implementation of sliding windows [17], to measure the contagion effect on stock market  
 93 indexes [18], to take the statistical analysis between stock market indexes [19], among  
 94 other applications.

95 However, for the particular case of four variables, i.e, one  $\{Y\}$  (dependent variable)  
 96 and three  $\{X_1\}, \{X_2\}, \{X_3\}$  (independent variables), there is still no published study  
 97 and this is what we propose to do in this paper from now on. From eq. 1, the calculus  
 98 of the  $DMC_x^2$  with 3 independent variables can be represented by this equation below:

$$\begin{aligned} 98 DMC_x^2 &= \left( \rho_{X_2, X_3}^2 \times \rho_{Y, X_1}^2 - \rho_{Y, X_1}^2 + \rho_{X_1, X_3}^2 \times \rho_{Y, X_2}^2 - \rho_{Y, X_2}^2 \right. \\ &\quad + 2 \times \rho_{X_1, X_2} \times \rho_{Y, X_1} \times \rho_{Y, X_2} - 2 \times \rho_{X_1, X_3} \times \rho_{X_2, X_3} \times \rho_{Y, X_1} \\ &\quad + \rho_{X_1, X_2}^2 \times \rho_{Y, X_3}^2 - \rho_{Y, X_3}^2 + 2 \times \rho_{X_1, X_3} \times \rho_{Y, X_1} \times \rho_{Y, X_3} \\ &\quad - 2 \times \rho_{X_1, X_2} \times \rho_{X_2, X_3} \times \rho_{Y, X_1} \times \rho_{Y, X_3} \\ &\quad - 2 \times \rho_{X_1, X_2} \times \rho_{X_1, X_3} \times \rho_{Y, X_2} \times \rho_{Y, X_3} \\ &\quad \left. + 2 \times \rho_{X_2, X_3} \times \rho_{Y, X_2} \times \rho_{Y, X_3} \right) / \\ &\quad \left( \rho_{X_1, X_2}^2 + \rho_{X_1, X_3}^2 + \rho_{X_2, X_3}^2 - 2 \times \rho_{X_1, X_2} \times \rho_{X_1, X_3} \times \rho_{X_2, X_3}^{-1} \right) \end{aligned} \quad (4)$$

99 To show the data mining and its calculations, for the aforementioned case, we present  
 100 the next Section.

## Data Mining and Calculations

After we download the EEG files stored in *Physionet* databank, the data mining follows some steps that are presented thereupon. In EEG experiments, usually, the end of the recordings is filled with a sequence of zeros, this corresponding to the time gap between the shooting down of the EEG machine and the recording system. In this pre-processing stage, these sequences of zeros are cut.

To properly apply the  $DMC_x^2$  and make proper comparisons between subjects and experiments, the time-series must have approximately the same length. In this sense, we have decided to delete the EEG relating to S106 because experiment 5 (Top/Down Real) for this subject has only  $N = 5808$  valid points. If we decided to cut all the time-series to this length, only the first 36.3 s will be computed (about 30% of the expected duration of 120 s), resulting in a great amount of data lost. The second smallest series has  $N = 15742$  valid points (S100 in experiment 12). Hence the interval between two consecutive EEG recorded values is  $\Delta t = 0.006250$ s, cutting the size of the series to  $N = 15742$  points, we guarantee a total time of  $T = 98.3875$ s, representing 82% of the expected duration and keeping all the subjects, besides S106. Thus, this data mining results in 108 subjects, each performing four Tasks in three consecutive experiments (see Table 1). In the next step, the  $DMC_x^2$  calculation is performed, with all possible combinations between the four time-series (case of study).

A data visualization approach was used to analyze the calculations due to the great amount of generated information. A total of 763 graphics were plotted to analyze the results. The first set of graphics presents the results of the  $DMC_x^2 \times n$  per subject per Task. In each subplot, four curves for the calculations of the  $DMC_x^2$  using combinations of the four channels as the dependent and independent variables. The subplots present each experiment that applies the same Task and the last one shows the mean values of the other Three.

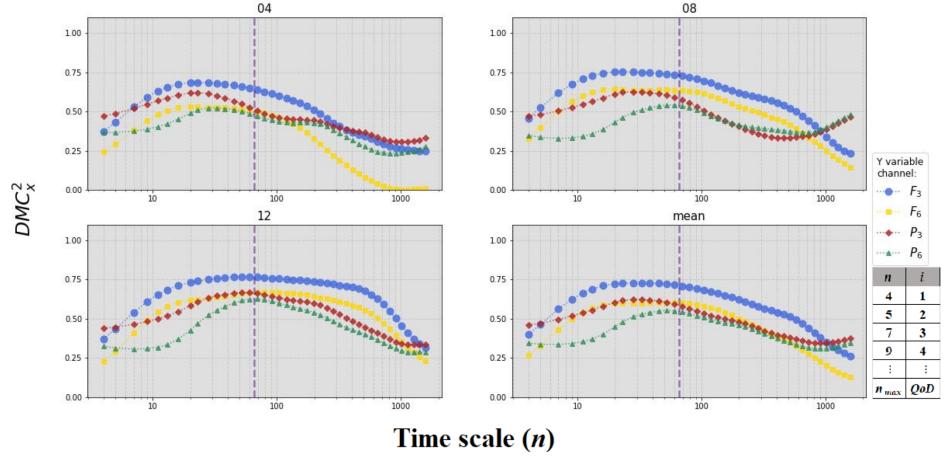
The color code used to denote the dependent variable used in the calculation of each curve is the same as the one used in Fig. 1 to locate the position of the studied channels,  $F_3$  (blue),  $F_6$  (yellow),  $P_3$  (red), and  $P_6$  (green). Table 2 explicates the color scheme relating to the channel used as the dependent variable (Y channel) and the ones used as the independent variables (X channels).

**Table 2.** The Y[Channel] (dependent variable) is represented by a specific color, that will be implemented in all figures with DCCA multiple cross-correlation coefficient.

| Color  | Y[Channel] | X[Channel1,Channel2,Channel3] |
|--------|------------|-------------------------------|
| Blue   | $Y[F_3]$   | $X[F_6,P_3,P_6]$              |
| Yellow | $Y[F_6]$   | $X[F_3,P_3,P_6]$              |
| Red    | $Y[P_3]$   | $X[F_3,F_6,P_6]$              |
| Green  | $Y[P_6]$   | $X[F_3,F_6,P_3]$              |

As an example, we randomly selected the subject S014 to present the first results. Fig. 2 shows the graphics of  $DMC_x^2 \times n$  for the three experiments where S014 executes Task 2, experiments number 04 (upper left subfigure), 08 (upper right) and 12 (lower left). The lower right subfigure presents the mean value for the three experiments. For this experiment and subject, the highest correlation value is in a smaller time scale than the highest value observed in the global mean, but the  $F_3$  channel as Y variable is the function with the highest correlation, following the global mean pattern, corroborating the findings in [2] for auto-correlation analysis.

The second set summarizes the Tasks by subject, using the mean values calculated presented in the last set, aiming to investigate characteristics regarding the different



**Fig 2.**  $DMC_x^2$  as a function of time scale  $n$ . Here is showing the results for subject S014 recordings for Task 2, presenting experiments 04, 08, 12 (Table 1) and the mean values for these experiments. The vertical line represents  $n = 67$  and  $QoD$  is the total amount of time scales involved in  $DMC_x^2$  calculations.

Tasks executed by the same subject.

The third set computes the global statistics of all the subjects by Task. The mean, median and standard deviation for all the subjects, using the mean for each Task by subject as input.

Sets four and five deal with differences. The fourth set of graphics shows, per subject, each Task mean value minus the global mean for the same Task. Set number five presents the differences on the mean value of imaginary Tasks minus the mean of real Tasks, showing a subplot for the Top/Down activities and another for the Left/Right one per subject. The goal is to gauge the differences between imagining a movement and realizing it.

All sets of graphics presented up to this point use the color scheme presented in Table 2.

The last set of graphics presents the Mean square error (MSE) in each graphic, using one of the channels as the dependent variable and each subplot to show the MSE for one of the Tasks. This set of figures aims to summarize the variations between the global mean and the individual means.

This great amount of graphics, although very important to ground the discussions and conclusions, wouldn't fit in the size of a paper, but were published in an online repository, for public scrutiny, in the link below:

[https://github.com/gfzebende/RHO\\_DCCA-Statistics.git](https://github.com/gfzebende/RHO_DCCA-Statistics.git)

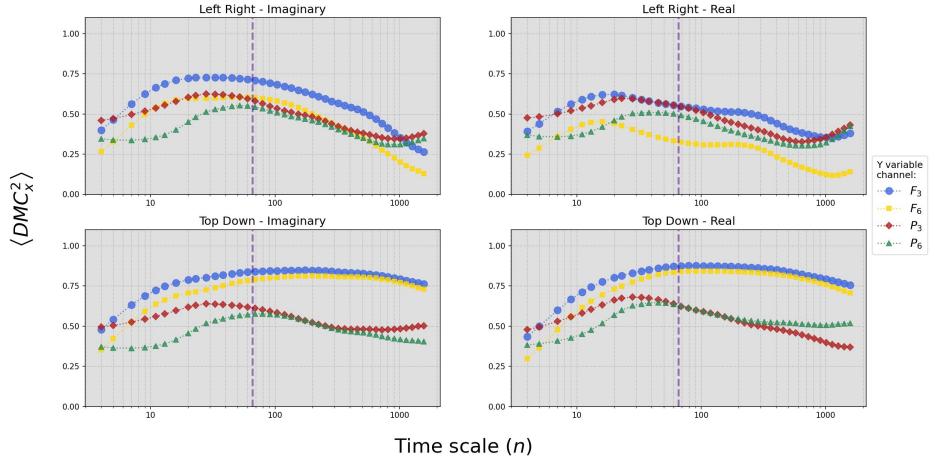
After all preamble regarding to EEG, methodology and data mining, in the next Section, we present the results and the discussion of  $DMC_x^2$  application in motor activity (Real/Imaginary), starting with the second set of graphics.

## Results and Discussion

The second set of graphics presents the mean values for each of the Tasks, as a better way to compare how each subject performs in different questions. This data treatment was implemented to mitigate the effect of outlines and get a better idea of how a given

Task was performed in terms of its DCCA multiple cross-correlation coefficient, represented by  $\langle DMC_x^2 \rangle$ . With a sample of five randomly selected results (S014, S036, S039, S078 and S099) among the 108 subjects, our initial statistical and graphical analysis will be presented, as listed below:

- S014 in Fig. 3;
- S036 in Fig. 4;
- S039 in Fig. 5;
- S078 in Fig. 6;
- S099 in Fig. 7.



**Fig 3.** Mean values of  $DMC_x^2 \times n$  for all Tasks: Left/Right (Imaginary), Left/Right (Real), Top/Down (Imaginary), and Top/Down (Real) for the S014 subject.

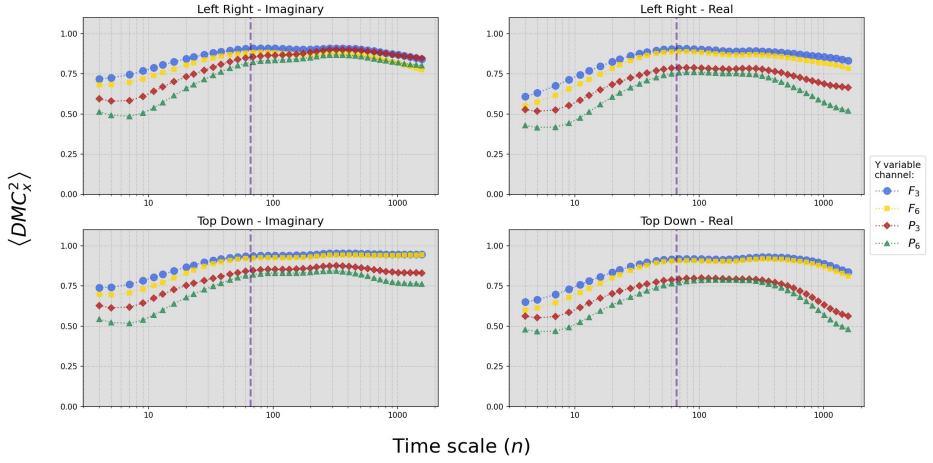
The first (upper left) subplot of Fig 3 is the same graphic presented in the last (lower left) graphic of Fig 02. In the first time scale the  $P_3$  function is higher than all the other channels for all the Tasks, followed by  $F_3$ ,  $P_6$ , and  $F_6$ , respectively. As the time scale grows, channel  $F_3$  becomes the greatest in correlation values. The left-right Tasks are very similar, and also the top-down ones.

Fig 4 shows channel  $F_3$  as the greatest correlation, and channels  $F_6$ ,  $P_3$ , and  $P_6$  in sequence for most of the time scales. The mean of the real Tasks experiments is more similar than any other two. The highest correlation values are found in the time scale on  $n = 67$ . In the top-down imaginary subplot, the correlation values remain very close to the highest value after the  $n = 67$  time scale for function  $F_3$  and  $F_6$ .

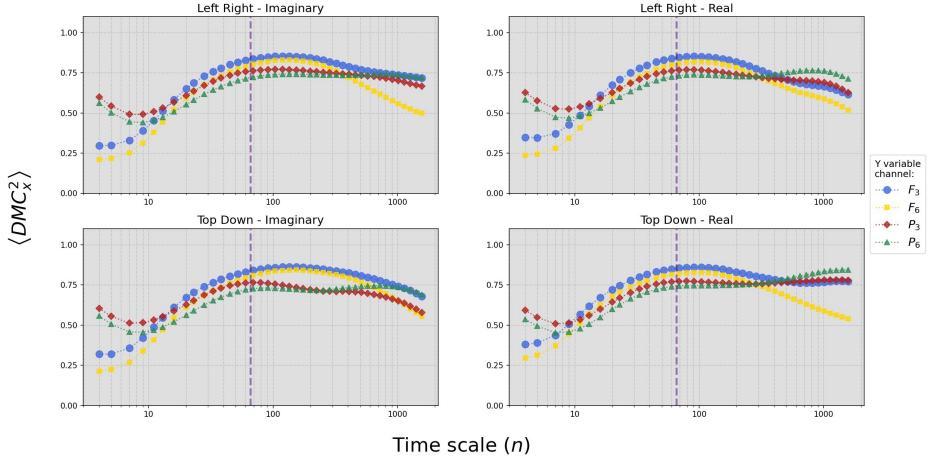
Fig 5 all Tasks starts with channels  $P_3$ , and  $P_6$  as the greatest values. The highest values are found after the time scale of  $n = 67$ . Around the  $n = 67$ , the order of functions are  $F_3$ ,  $F_6$ ,  $P_3$ , and  $P_6$ . All the Tasks presents similar graphics.

In Fig 6, all Tasks are very similar, beginning with  $P_3$ ,  $P_6$ ,  $F_3$ , and  $F_6$  as channel order, but the order are rapidly reordered as  $F_3$ ,  $F_6$ ,  $P_3$ , and  $P_6$ . The correlation values keep getting slightly higher after the  $n = 67$  time scale, the left-right Tasks presents a more pronounced decrease in the greatest time scales than the top-down ones.

Subject S099, in Fig 7, functions  $P_3$ , and  $P_6$  presents an almost linear behavior. In the left-right real Task, function  $F_6$  displays the lowers values, and values higher than the ones in time scale  $n = 67$  after it, for functions  $F_3$ ,  $P_3$ , and  $P_6$ .



**Fig 4.** Mean values of  $DMC_x^2 \times n$  for all Tasks: Left/Right (Imaginary), Left/Right (Real), Top/Down (Imaginary), and Top/Down (Real) for the S036 subject.



**Fig 5.** Mean values of  $DMC_x^2 \times n$  for all Tasks: Left/Right (Imaginary), Left/Right (Real), Top/Down (Imaginary), and Top/Down (Real) for the S039 subject.

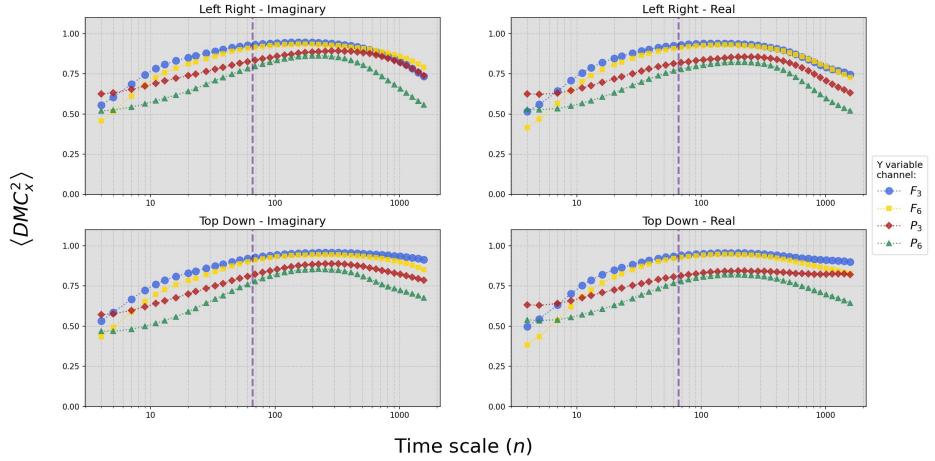
With  $DMC_x^2$  mean obtained, Thus, even for this small sample (five subjects), it can be said that each subject has his own signature when performing these Tasks, and in general, the Tasks (Real/Imaginary or Top/Down) follow the same type of pattern. For a better visualization of this statement, it is possible to watch the video with  $\langle DMC_x^2 \rangle$  for all subjects, at this link below:

### Videos

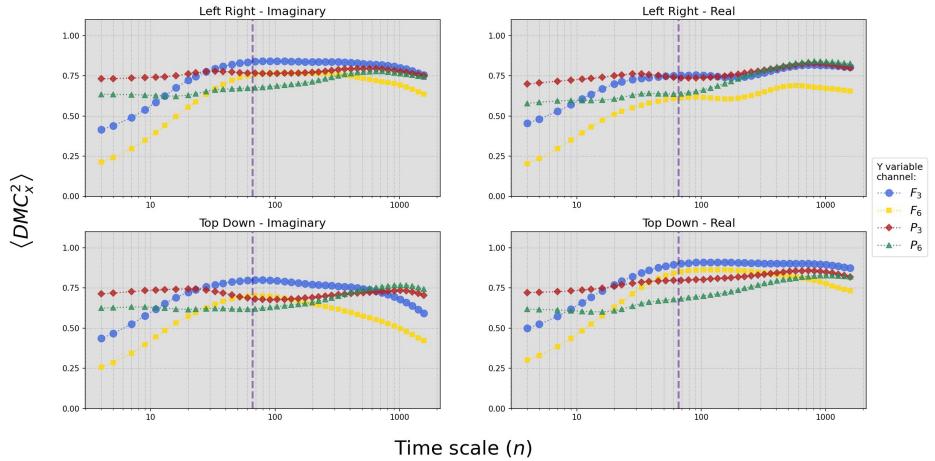
With  $DMC_x^2$  mean for each subject, our next step was to calculate the global mean:

$$\overline{\langle DMC_x^2 \rangle}(n) = \frac{1}{108} \sum_{i=1}^{108} \langle DMC_x^2 \rangle_i(n) \quad (5)$$

And its results can be seen in Fig. 8. In this global mean,  $\overline{\langle DMC_x^2 \rangle}$ , we can see that, in fact, there are no appreciable differences between the 4 Tasks (Left/Right Imaginary,



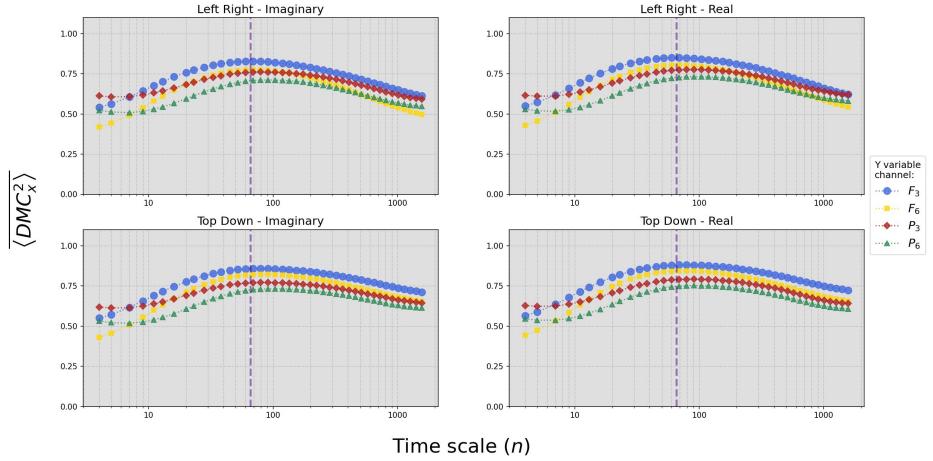
**Fig 6.** Mean values of  $DMC_x^2 \times n$  for all Tasks: Left/Right (Imaginary), Left/Right (Real), Top/Down (Imaginary), and Top/Down (Real) for the S078 subject.



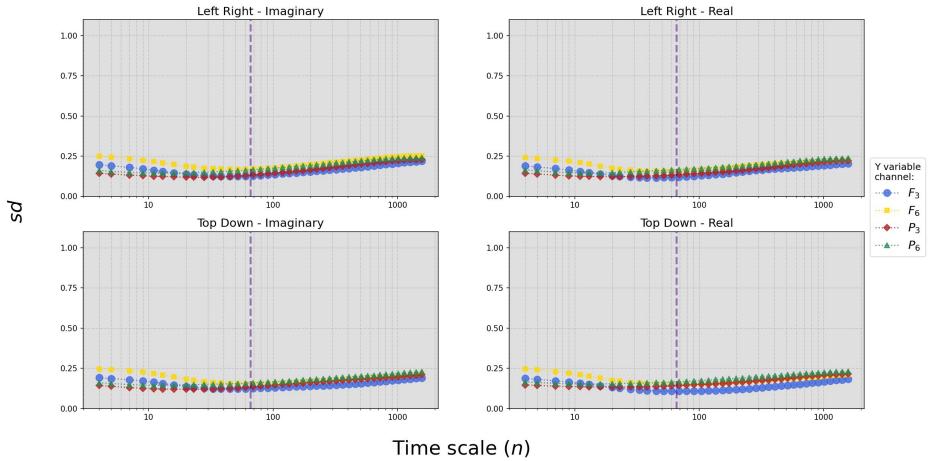
**Fig 7.** Mean values of  $DMC_x^2 \times n$  for all Tasks: Left/Right (Imaginary), Left/Right (Real), Top/Down (Imaginary), and Top/Down (Real) for the S099 subject.

Left/Right Real, Top/Down Imaginary and Top/Down Real). The channel  $F_3$  as a dependent variable,  $Y[F_3]$  (blue circles), in fact it was confirmed as the one with the highest value for DCCA multiple cross-correlation coefficient, mainly for  $n = 67$  or 0.42s (about 2.38Hz). Remembering that, in EEGs the most characteristic signals are observed between 1 and 20 Hz. For small time-scales,  $n \leq 7$ , the parental channels ( $P_3$  e  $P_6$ ), as dependent variables, have  $\langle DMC_x^2 \rangle$  greater than if we look at  $F_6$  channel.

If we are interested in the measure of dispersion around the mean, we must calculate its standard deviation,  $sd$ . Thus, the  $sd$ , of the global mean is shown in the Fig. 9. A low  $sd$  indicates that the data points tend to be close to the mean, whereas a high  $sd$  indicates that the data points are spread over a wide range of values. Here, Fig. 9, we can see the same type of behavior for  $sd$  in all Tasks,  $sd$  with minimum value is around  $n = 67$  and  $sd$  maximum in small time-scales (the largest being for  $Y[F_6]$  (yellow circles)). For Median (not presented here but present in our public repository), its values are very similar to the global mean.



**Fig 8.**  $DMC_x^2 \times n$  mean global for all Subjects and Tasks: Left/Right (Imaginary), Left/Right (Real), Top/Down (Imaginary), and Top/Down (Real).



**Fig 9.** Standard deviation,  $sd$ , of the global mean for all Subjects and Tasks.

## QUAL O MOTIVO DE UTILIZARMOS ISTO AQUI.....????

Looking at each subject in terms of the global mean, we can define the mean square error,  $MSE$ , as

$$MSE(j) = \frac{1}{QoD} \sum_{i=1}^{QoD} \left( DMC_x^2[i] - \overline{\langle DMC_x^2 \rangle}[i] \right)^2 \quad (6)$$

that which will be calculated for each subject  $S_j$  (mean in its three experiments), for each channel and Task. Where  $QoD$  is the total amount of time-scales used (in this paper  $QoD = 42$ , see inset in the Fig. 2 as example) and  $\overline{\langle DMC_x^2 \rangle}[i]$  the global mean value for the time-scale  $i$ . The results for  $MSE$  per channel ( $F_3$ ,  $F_6$ ,  $P_3$ , and  $P_6$ ) and Tasks can be seeing in Appendix A, at the Figs. 10, 11, 12, and 13.

## O QUE SE VÊ AQUI???????

Finally, if it is necessary and important to look at the difference between the

subjects  $j$  in relation to the global mean, we can apply:

$$diff_j(n) \equiv \langle DMC_x^2 \rangle_j(n) - \langle DMC_x^2 \rangle(n) \quad (7)$$

performed by Task and Channel. This statistic can also be seen here by accessing the link below:

### Videos

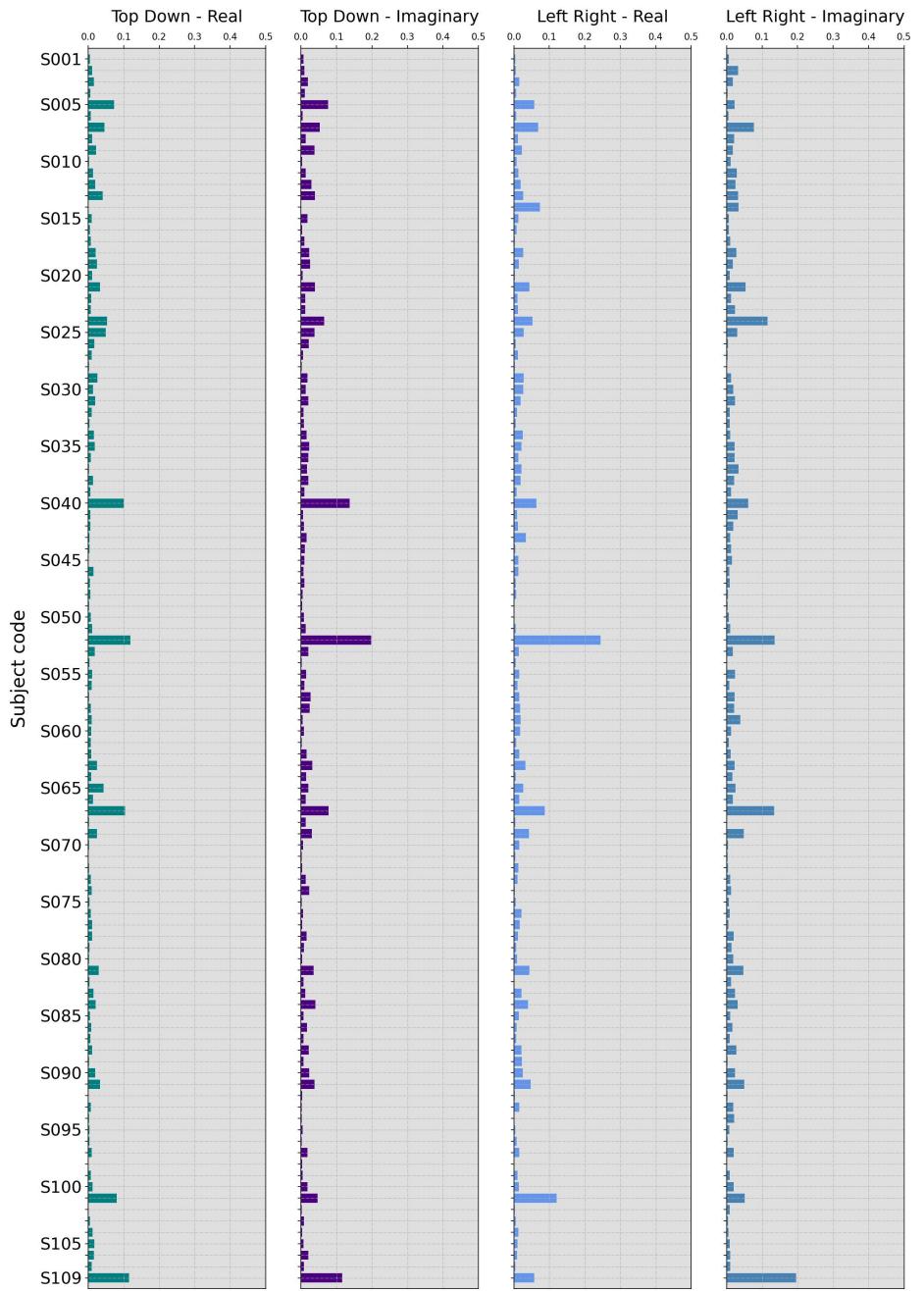
in the file “Difference”.

## Conclusion

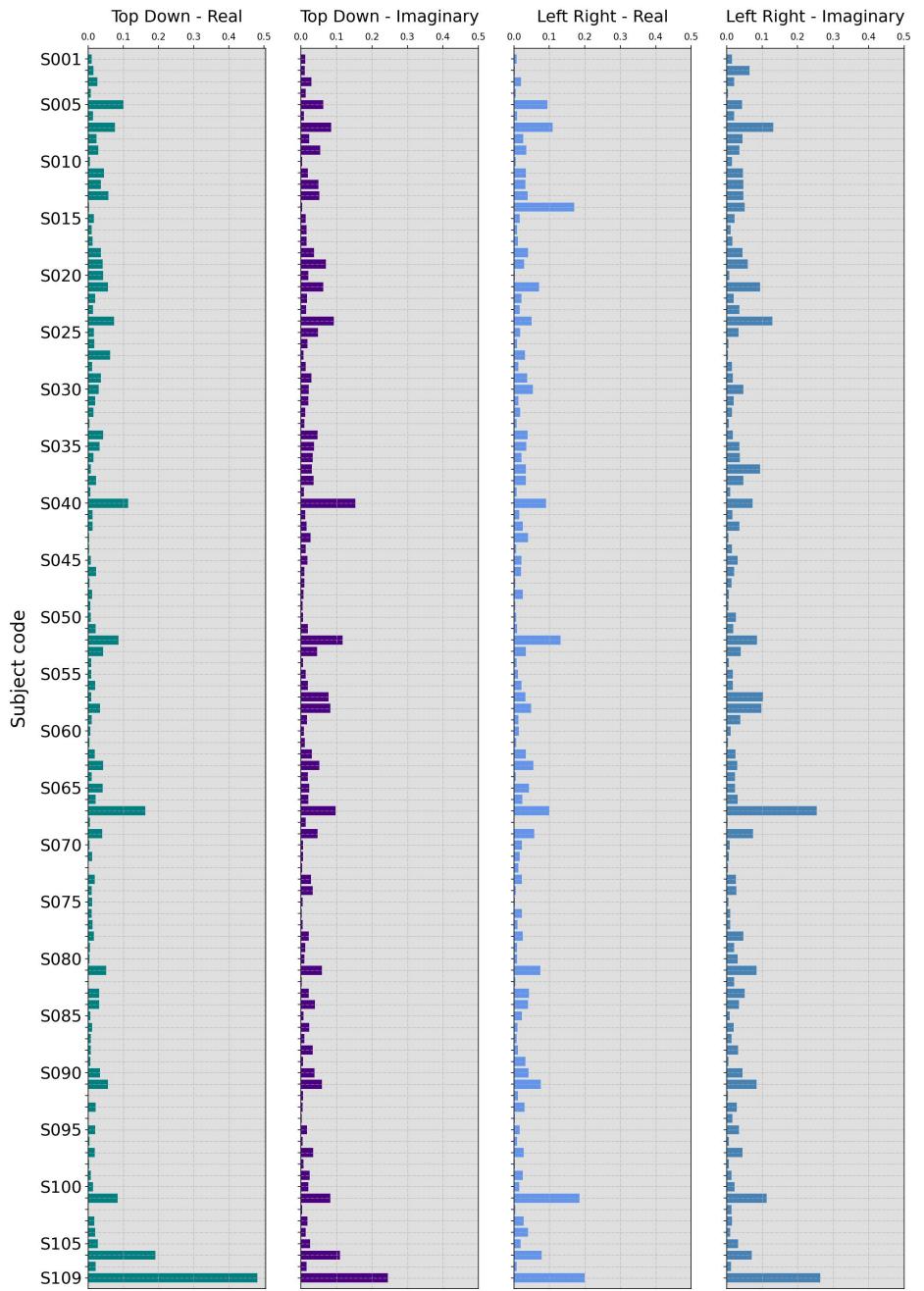
The  $DMC_x^2$  analysis of the EEG data captures particular patterns for each subject. These patterns appear as a unique signature, like a “fingerprint”. Comparing the images, is plausible to assume that a computational algorithm will be able to identify, among the participants of the experiment, a subject that performed an anonymous record of one of the Tasks, based on the previously collected data.

Fig. 8, shows that the greater multiple cross-correlation values for the mean of all studied channel occurs around  $0.42s$ . This time scale is also the maximum value observed in the median, indicating that the mean is not significantly affected by outliers, and the lowest values for the standard deviations, indicate a small dispersion. The value of  $0.42s$  or  $2.38\text{ Hz}$  possibly corresponding to the frequency of a  $\delta$ -wave.

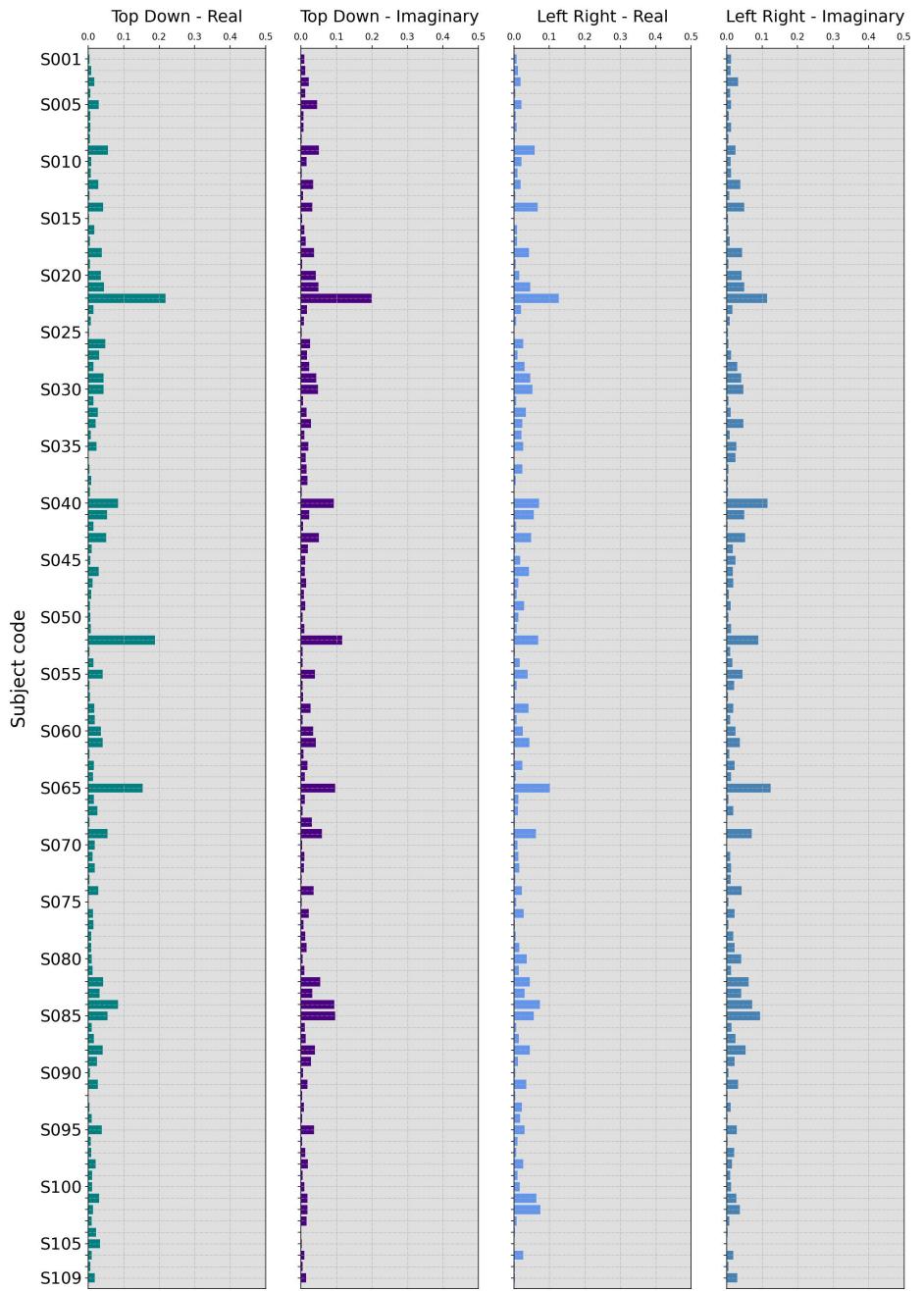
## Appendix A



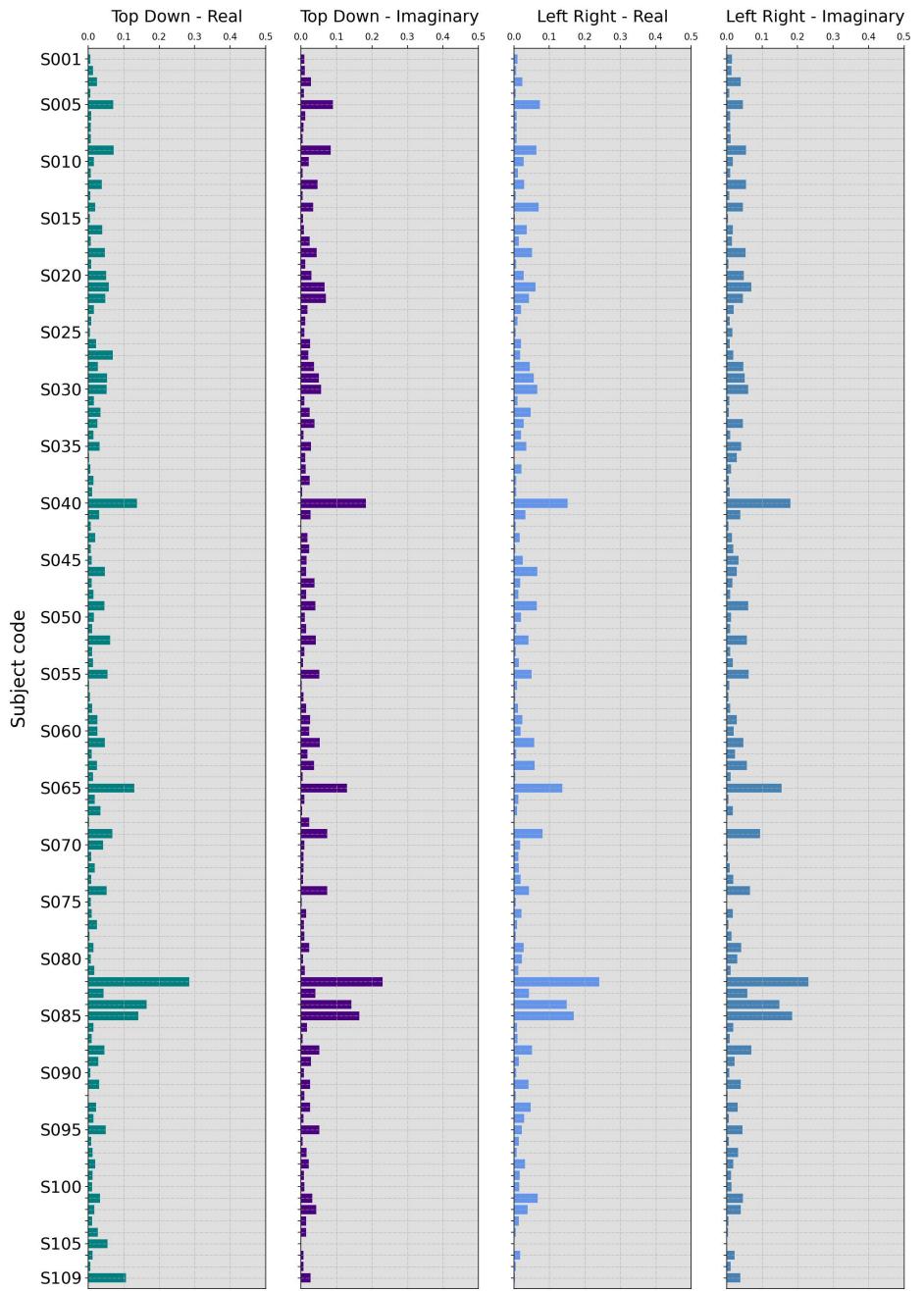
**Fig 10.** MSE for the Channel  $F_3$ .



**Fig 11.** MSE for the Channel  $F_6$ .



**Fig 12.** MSE for the Channel  $P_3$ .



**Fig 13.** MSE for the Channel  $P_6$ .

## Acknowledgments

G. F. Zebende thanks the Brazilian agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) (Grant 310136/2020-2). F. M. Oliveira Filho thanks the Brazilian agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) (Grant 150655/2022-3). We would also like to thank Paulo Murilo Castro de Oliveira, great statistical physicist (*in memoriam*).  
248  
249  
250  
251  
252  
253

## References

1. Biasiucci A, Franceschiello B, Murray MM. Electroencephalography. *Curr Biol.* 2019;29(30):R80–R85.
2. Zebende GF, Oliveira-Filho FM, Leyva-Cruz JA. Auto-correlation in the motor/imaginary human EEG signals: A vision about the FDFA fluctuations. *PLoS One.* 2017;12(9).
3. Oliveira Filho FM, Leyva Cruz JA, Zebende GF. Analysis of the EEG bio-signals during the reading task by DFA method. *Physica A: Statistical Mechanics and its Applications.* 2019;525:664–671.
4. Mesquita VB, Oliveira-Filho FM, Rodrigues PC. Detection of crossover points in detrended fluctuation analysis: An application to EEG signals of patients with epilepsy. *Bioinformatics.* 2020;doi:10.1093/bioinformatics/btaa955.
5. Oliveira Filho FM, Ribeiro FF, Cruz JAL, de Castro APN, Zebende GF. Statistical study of the EEG in motor tasks (real and imaginary). *Physica A: Statistical Mechanics and its Applications.* 2023; p. 128802.
6. Zebende GF, da Silva Filho AM. Detrended Multiple Cross-Correlation Coefficient. *Physica A: Statistical Mechanics and its Applications.* 2018;510:91–97.
7. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet. *Circulation.* 2000;101(23):e215–e220. doi:10.1161/01.CIR.101.23.e215.
8. Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JE. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng.* 2004;51(6):1034–1043.
9. Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL. Mosaic organization of DNA nucleotides. *Phys Rev E.* 1994;49:1685–1689.
10. Podobnik B, Jiang ZQ, Zhou WX, Stanley HE. Statistical tests for power-law cross-correlated processes. *Physical Review E.* 2011;84:1–8.
11. Zhao JC, Dou WH, Ji HD, Wang J. Detrended Cross-Correlation Analysis of Epilepsy Electroencephalogram Signals. *Advanced Materials Research.* 2013;765–767:2664–2667.
12. Chen Y, Cai L, Wang R, Song Z, Deng B, Wang J, et al. DCCA cross-correlation coefficients reveals the change of both synchronization and oscillation in EEG of Alzheimer disease patients. *Physica A: Statistical Mechanics and its Applications.* 2018;490:171–184.

13. Zebende GF. DCCA cross-correlation coefficient: Quantifying level of cross-correlation. *Physica A*. 2011;390(4):614–618.
14. Kristoufek L. Measuring cross-correlation between non-stationary series with DCCA coefficient. *Physica A: Statistical Mechanics and its Applications*. 2014;402:291–298.
15. Brito AA, Araujo HA, Zebende GF. Detrended Multiple Cross-Correlation Coefficient applied to solar radiation, air temperature and relative humidity. *Sci Rep*. 2019;9:19764. doi:10.1038/s41598-019-56114-6.
16. da Silva Filho AM, Zebende GF, de Castro APN, Guedes EF. Statistical test for multiple detrended cross-correlation coefficient. *Physica A: Statistical Mechanics and its Applications*. 2021;562:125285.
17. Guedes EF, da Silva Filho AM, Zebende GF. Detrended multiple cross-correlation coefficient with sliding windows approach. *Physica A: Statistical Mechanics and its Applications*. 2021;574:125990.
18. Guedes EF, de Castro APN, da Silva Filho AM, Zebende GF.  $\Delta$  DMC x 2: A New Approach to Measure Contagion Effect on Financial Crisis. *Fluctuation and Noise Letters*. 2022;21(04):2250026.
19. Zebende GF, Aguiar LC, Ferreira P, Guedes EF. Statistical Approach to Study the Relationship Between Stock Market Indexes by Multiple DCCA Cross-Correlation Coefficient. *Fluctuation and Noise Letters*. 2022;21(5):2250045–1314.