A new method to build multiplex networks using Canonical Correlation Analysis for the characterization of the Alzheimer's disease continuum

Saúl J. Ruiz-Gómez[1,2], Roberto Hornero[1,2,3], Jesús Poza[1,2,3], Eduardo Santamaría-Vázquez[1,2], Víctor Rodríguez-González[1,2], Aarón Maturana-Candelas[1,2], Carlos Gómez[1,2]

[1] Biomedical Engineering Group, E.T.S.I. Telecomunicación, University of Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain
[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), Spain
[3] IMUVA, Mathematics Research Institute, University of Valladolid, Valladolid, Spain

E-mail: saul.ruiz@gib.tel.uva.es

Abstract. *Objective.* The aim of this study was to solve one of the current limitations for the characterization of the brain network in the Alzheimer's disease (AD) continuum. Nowadays, frequency-dependent approaches have reached contradictory results depending on the frequency band under study, tangling the possible clinical interpretations. *Approach.* To overcome this issue, we proposed a new method to build multiplex networks based on canonical correlation analysis (CCA). Our method determines two basis vectors using the source and electrode-level frequency-specific network parameters for a reference group, and then project the results for the rest of the groups into these hyperplanes to make them comparable. It was applied to: (i) synthetic signals generated with a Kuramoto-based model; and (ii) a resting-state EEG database formed by recordings from 51 cognitively healthy controls, 51 mild cognitive impairment subjects, 51 mild AD patients, 50 moderate AD patients, and 50 severe AD patients. *Main Results.* Our results using synthetic signals showed that the interpretation of the proposed CCA-based multiplex parameters (multiplex strength, multiplex characteristic path length and multiplex clustering coefficient) can be analogous to their frequency-specific counterparts, as they displayed similar behaviors in terms of average connectivity, integration, and segregation. Findings using real EEG recordings revealed that dementia due to AD is characterized by a significant increase in average connectivity, and by a loss of integration and segregation. *Significance.* We can conclude that CCA can be used to build multiplex networks based from frequency-specific results, summarizing all the available information and avoiding the limitations of possible frequency-specific conflicts. Additionally, our method supposes a novel approach for the construction and analysis of multiplex networks during AD continuum.

Keywords: canonical correlation analysis, connectivity, multiplex networks, synthetic signals, electroencephalography (EEG), Alzheimer's disease, mild cognitive impairment.

Submitted to: *J. Neural Eng.*

1. Introduction

In terms of graph theory, a network can be defined as a set of elements (nodes) linked by connections or interactions (edges) [1]. In this regard the human brain can be seen as a dynamic, complex network, where brain regions are the nodes and the connections between them are the edges. Cerebral networks are continuously adapting themselves following external and internal stimuli, as well as physiological sculpture changes during brain maturation and throughout the life span [2]. However, normal brain aging can be altered by some physiopathological processes, like those associated to neurodegeneration.

Dementia due to Alzheimer's disease (AD) is the most common cause of dementia, accounting for an estimated 60 to 80 % cases [3]. Based on current research advances, AD is best conceptualized as a biological and clinical continuum, covering from the preclinical to symptomatic phases of AD. Cognition (including episodic memory, executive function, and verbal fluency) and function (basic and complex activities of daily living) appear to decline in a very subtle way at AD early stages. While the concept of a continuum appears to be more appropriate for the characterization of AD course, some degree of staging has been helpful for clinical purposes. Traditionally, two key stages have been considered: mild cognitive impairment (MCI) due to AD, and dementia due to AD [4, 5]. MCI is usually considered a prodromal stage of AD in which patients exhibit cognitive deficits, but do not fully accomplish the criteria for dementia diagnosis [6]. Then, depending on the symptoms, AD dementia can be further divided into mild (AD_{mil}), moderate (AD_{mod}), and severe AD (AD_{sev}) [5]. Firstly, AD_{mil} patients show clear deficits on clinical examination but they may function independently. In the next stage, AD_{mod} patients suffer increasingly poor judgment and deepening confusion, and require a greater level of care. Finally, AD_{sev} patients are completely dependent as they loose the ability to respond to their environment and to control movement.

Nowadays, AD diagnosis is complex and with a high degree of subjectivity [3]. Notwithstanding, whole-brain neuroimaging techniques have shown their potential to study MCI and AD brain alterations, claiming their role as interesting methodologies to find relevant biomarkers and develop new objective diagnostic tools [7, 8, 9, 10]. Among them, neurophysiological techniques provide the ability to study fast changing neural networks. This is the case of electroencephalography (EEG), which is a brain imaging method that measures the electrical activity generated by synchronized cortical neurons pools [9].

EEG has already proven it usefulness to characterize MCI and AD pathophysiological brain dynamics during resting state [11]. Traditionally, three major effects of MCI and AD have been observed from EEG recordings: slowing of the EEG, reduced complexity of EEG signals, and perturbations in EEG connectivity [11]. Based on connectivity results, graphs can be constructed as an abstract representation of the cerebral activity, providing insights into the specific organization of the brain. The organization of these graphs has been quantified using parameters derived from complex network theory in the conventional EEG frequency bands: delta, theta, alpha, beta-1, beta-2, and gamma. Recent functional network studies have reported that AD is mainly associated with a decrease of integration (local information processing) and a increase of segregation (global information processing) [12, 13]. However, previous studies that compared graph properties between AD and healthy control networks in specific frequency bands, using EEG, MEG, and fMRI data, did not report consistent result between them. On one hand, contradictory results depending on the frequency band under analysis were showed. For example, Stam *et al.* reported significant clustering coefficient decreases in alpha, but increases in theta [12]. The opposite trend was reported for the characteristic path length, with significant increases in alpha and decreases in theta [12]. This makes difficult to extract conclusions about the alterations that AD neurodegeneration elicits in the global brain network. On the other hand, previous studies also reported inconsistent results between them when studying the same frequency band. Some works that investigated the average clustering coefficient did not find differences between AD and control subjects [14, 15, 16], others found significant decreases in AD [12, 17] and other studies also found an increased clustering in AD [18, 19, 20]. This issue also happens with the characteristic path length, as higher values as the disease progresses were reported [14, 15, 18, 19, 20], but also lower values [16, 17]. All these results tangle the possible clinical interpretations about the brain changes provoked by the AD neurodegeneration process. Furthermore, analyzing single layers separately could not capture relevant information

of the whole brain network that could be provided by a multiplex approach. Then, analyzing the multiplex network considering the information of all layers together has focused a great attention.

Multiplex network theory has been applied in different real-world multiplex networks, such as social systems [21] and interconnected hyperlink networks [22]. In the case of EEG-based networks, each frequency-specific network (typically: delta, theta, alpha, beta-1, beta-2, and gamma) corresponds to a layer. These layers are interconnected by the links between the same set of electrodes across layers. Different methods have been used to build these multiplex networks: from a simple parameter sum across all frequency-specific layers to more complex approaches [21, 23, 24], such as using a quality function to determine the value of the links across the different layers [25]. However, these methods present some limitations. The estimation of multiplex parameters and the comparison between multiplex networks could be biased if the link weights are different across studies [26]. Furthermore, these approaches could compensate opposite trends between groups for the analyzed frequency bands and blur the differences between groups.

In this study, we proposed a new approach to build multiplex networks based on Cononical Correlation Analysis (CCA). The proposed method shows several benefits over single-layer methodologies, and over other previously proposed multiplex approaches. Regarding the advantages over single-layer methodologies, results obtained using the proposed method can summarize the available information across all layers in one single value, making the global brain network easier to interpret from the clinical point of view. Having only one value that represents a particular network feature instead of one value per frequency-band also avoids the problem of contradictory results depending on the frequency band under study, which are commonly obtained in AD studies [12, 27]. With regard to the advantages over previously proposed multiplex approaches, most of them require having the link values between homologous nodes of different layers to build the multiplex network, including those described in the studies of De Domenico, Boccaletti, and Buldú [22, 23, 28]. However, these interlayer links are not usually available, as in our study. Other authors have proposed multiplex approaches for scenarios where the interlayer links are not available, such as the sum of the different parameters for each layer [21, 23]. However, considering the results of some previous studies [12, 27], where opposite trends were reported depending on the frequency band under analysis (e.g., significant clustering coefficient decreases in alpha, but increases in theta), it could be expected that this approach could compensate the opposite trends for the analyzed frequency bands and blur the differences between groups. In our method, the different weights assigned to each layer during the procedure of computing the multiplex parameters are given by an objective procedure (CCA). These weights define the hyperplane where the correlation between source and electrode-level results is maximized and they could be negative, thus avoiding to compensate the aforementioned AD opposite trends in the different frequency bands. CCA has been previously applied as a feature extraction method in Brain Computer Interface applications [29, 30], to investigate the relationship between synchronous neural interactions and cognitive neuropsychological domains in AD patients [31], and to improve the accuracy in coherence estimation by means of weighting [32]. We propose a new methodology to build multiplex networks based on CCA, which determines the coefficients that maximize the correlation between electrode-level and source-level results. Firstly, in order to evaluate the behavior of the multiplex parameters under controlled coupling conditions, we used our previously published synthetic signals [33], generated by the combination of a surface based computational model of the human head and a model of coupled oscillators. Then, we applied the proposed methodology to real EEG recordings to obtain new multiplex parameters that summarize all the frequency-band information for each subject under study. Therefore, the aim of this study was to analyze how three different CCA multiplex parameters (CCA global strength s_{CCA}, CCA characteristic path length L_{CCA}, and CCA clustering coefficient C_{CCA}) are affected by variations of connectivity using synthetic signals, and if they are able to characterize the brain alterations during AD continuum.

2. Materials

2.1. Synthetic signals

Firstly, in order to analyze the behavior of the new proposed metrics, we used a set of simulated signals using artificial brain sources and a realistic head model previously published by our research group [33]. The surface-based model was constructed from the open-source Visible Human Project® (VHP) dataset [34]. The brain, the cerebrospinal fluid, the skull, and the skin were identified and hand-segmented using ITK-Snap [35]. Then, neuronal activity was propagated from synthetic brain sources to 19 simulated EEG electrodes placed following the international 10-20 system.

Once the model was created, the electrical source activity was simulated using a set of coupled oscillators to generate a Kuramoto model by varying the global

coupling parameter K. For each value of global coupling K, 300 trials were simulated using 200 oscillators randomly placed inside the brain mesh. This procedure results in 300 time series of 19 simulated channels and 2500 samples. It is noteworthy that in all trials, the initial 5000 samples were discarded to eliminate transitory states of the oscillators. Further details about the generation of these synthetic signals can be found in our previous study [33].

2.2. Real EEG signals

The real dataset formed by 253 resting-state EEG recordings. They were recorded from 51 cognitively healthy control (HC) subjects, 51 MCI subjects, 51 AD_{mil} patients, 50 AD_{mod} patients, and 50 AD_{sev} patients. HC group was composed by elderly subjects with no neurological disorders. On the other hand, patients with MCI and dementia due to AD were diagnosed following the criteria of the National Institute on Aging and Alzheimer's Association (NIA-AA) [4, 36]. The socio-demographic and clinical characteristics of each group are provided in table 1.

For each subject, five minutes of resting state EEG activity were recorded by means of a 19-channel EEG system Nihon Kohden Neurofax JE-921A at a sampling frequency of 500 Hz. The electrodes were placed following the specifications of the international 10–20 system at Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, T5, T6, Pz, P3, P4, O1, and O2. Signals were re-referenced by means of common average referencing. Then, they were preprocessed in three steps [33, 37]: (i) band-pass digital filtering to limit the spectral content to the wide frequency band between 1 and 70 Hz (Hamming window, filter order 2000, forward and backward filtering) and band-stop digital filtering to remove the 50-Hz noise power (Hamming window, filter order 2000, forward and backward filtering); (ii) independent component analysis (ICA) to minimize the presence of oculographic, cardiographic, and myographic artifacts; and (iii) selection of 5 s artifact-free epochs by visual inspection.

During the ICA analysis, an average of 2.01 ± 1.88 (mean ± standard deviation) components were discarded per subject from a total number of 19 components. After the visual inspection, an average of 38.81 ± 13.03 5-s artifact-free epochs per subject were considered for further analysis.

3. Methods

Firstly, synthetic signals were used to study the behavior of the CCA multiplex parameters under different connectivity conditions. For this purpose, we performed the following steps:

(i) Estimation of connectivity between pair-wise nodes using Phase Lag Index (PLI) (see Section 3.2) at the electrode-level and source-level synthetic signals.

(ii) Calculation of frequency-specific network parameters: global strength (s), characteristic path length (L), and clustering coefficient (C) were obtained from PLI matrices both at electrode-level and at source-level (see Section 3.3).

(iii) These electrode and source frequency-specific parameters formed the multidimensional variables to compute CCA. Then, two basis vectors were obtained (electrode-level and source-level) using the less connected synthetic signals, which were computed with $K = 2$ in the Kuramoto model (see Sections 3.4 and 3.5).

(iv) Finally, in order to make results comparable, network parameters from all groups (K = 2, 2.5, 3, and 3.5) were projected onto these hyperplanes computing the linear combination of the frequency-dependent results multiplied by the reference basis vectors.

Then, real EEG signals were used to analyze the trends of the CCA multiplex parameters during AD continuum. This procedure is represented in figure 1, and it is divided into the following steps:

(i) Source level signals were reconstructed from real EEG recordings using standardized low resolution brain electromagnetic tomography (sLORETA) (see Section 3.1).

(ii) Connectivity between pair-wise electrodes was computed using PLI for the real EEG recordings, and between pair-wise Regions of Interest (ROIs) for reconstructed source-level signals (see Section 3.2).

(iii) The aforementioned frequency-specific network parameters were also obtained from PLI matrices (see Section 3.3).

(iv) Two basis vectors were obtained (electrode-level and source-level) using only the HC results (see Sections 3.4 and 3.5).

(v) Then, results from all groups (HC, MCI, AD_{mil}, AD_{mod}, and AD_{sev}) were projected into these hyperplanes computing the linear combination of the frequency-dependent results multiplied by the reference basis vectors (see Section 3.5)."

It is noteworthy that, in both cases, we only analyzed the source-level results, since source-reconstructed networks are comparable regardless of the EEG recording system configuration. That is, if the Desikan-Killiany atlas is used to reduce signal dimensionality into ROIs, the derived networks

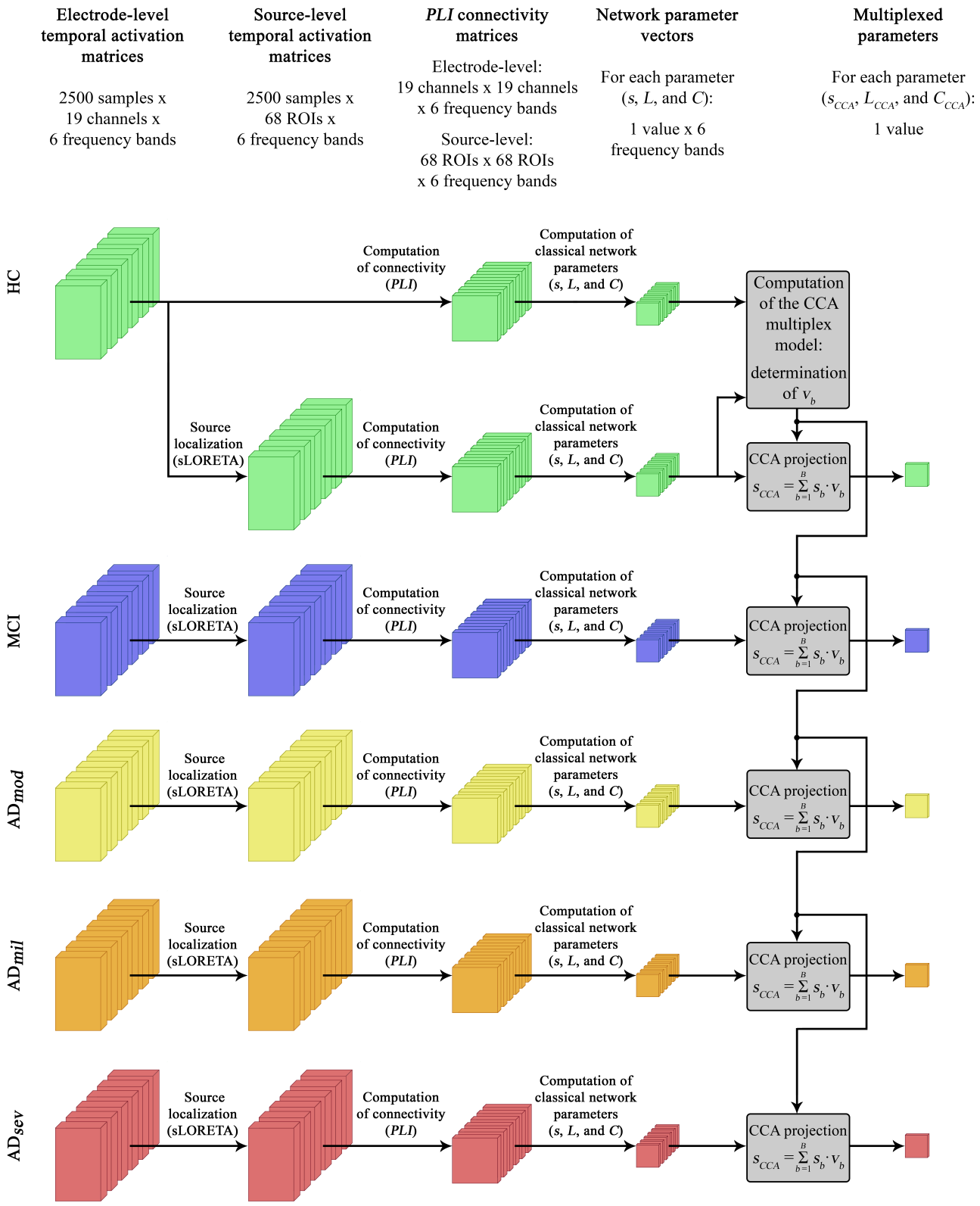

Figure 1. Block diagram of the used methodology with the real EEG signals. It is divided into the following steps: (i) reconstruction of the source-level signals from real EEG recordings (see Section 3.1); (ii) estimation of connectivity between pair-wise nodes (see Section 3.2); (iii) computation of frequency-specific network parameters (see Section 3.3); (iv) Computation of the multiplex model based on CCA: determination of the basis vector (v_b) that defines the reference hyperplane using the healthy control results (see Sections 3.4 and 3.5); (v) CCA projection: projection of the results onto these hyperplanes computing the linear combination of the frequency-dependent results multiplied by the reference basis vectors: $s_{CCA} = \sum_{b=1}^{B} s_b \cdot v_b$ (see Section 3.5).

Table 1. Socio-demographic and clinical data.

	HC	**MCI**	**AD$_{mil}$**	**AD$_{mod}$**	**AD$_{sev}$**
Number of subjects	51	51	51	50	50
Age (years) (mean±SD[a])	80.14±7.09	85.53±7.25	80.69±7.05	81.30±8.04	79.98±7.82
Gender (Male:Female)	26:25	15:36	21:30	7:43	7:43
MMSE[b] (mean±SD[a])	28.82±1.13	23.33±2.84	22.49±2.27	13.60±2.76	2.42±3.70

[a] SD: standard deviation; [b]MMSE, Mini Mental State Examination score; HC: cognitively healthy control subjects; MCI: Mild Cognitive Impairment patients; AD$_{mil}$: mild Alzheimer's disease patients; AD$_{mod}$: moderate Alzheimer's disease patients; AD$_{sev}$: severe Alzheimer's disease patients.

will have always 68 nodes. Otherwise, electrode-level networks size depends on the number of acquisition channels, that can vary depending on the configuration. Furthermore, source-level results will allow us to analyze and compare new groups that could have been recorded with different EEG configurations, or even using MEG, with the healthy control group used in this study. This would not be possible if we use the electrode-level results.

3.1. Source reconstruction: sLORETA

Time series at source level from real EEG recordings were obtained using sLORETA, available in Brainstorm (http://neuroimage.usc.edu/brainstorm) [38]. sLORETA is based on a model of lineal distributed sources and provides a solution maximizing the correlation between neighboring sources, assuming that nearby neurons are synchronized [39]. In addition, sLORETA reduces the errors on the estimated source activity applying some physiological restrictions. sLORETA provides high-dimensionality source-level signals, which were projected in 68 ROIs using the Desikan-Killiany atlas. Technical details of the method are described in [39].

3.2. Connectivity

In order to construct the brain networks, adjacency matrices were computed using PLI. PLI is a connectivity metric proposed by Stam *et al.* [40] that quantifies the phase difference distribution between two time series. The stronger the phase locking is, the larger the PLI will be, except the case of shared signals at zero phase lag (where it is non sensitive). Moreover, our previous work using synthetic signals [33] demonstrated that PLI can detect real changes in synchronization. Additionally, this work also showed that PLI is less affected by the volume conduction problem than other measures, such as magnitude squared coherence, imaginary part of coherence, amplitude envelope correlation, synchronization likelihood, or phase locking value [33].

Mathematically, PLI is defined as [40]:

$$PLI_{X,Y} = |\langle sign \sin(\Delta\phi_{X,Y})\rangle|, \quad (1)$$

where $\langle\cdot\rangle$ indicates the expectation operator and $\Delta\phi_{X,Y}$ is the phase difference or relative phase between signals X and Y.

PLI was computed for the six conventional EEG-frequency bands: delta (δ, 1-4 Hz), theta (θ, 4-8 Hz), alpha (α, 8-13 Hz), beta-1 (β_1, 13-19 Hz), beta-2 (β_2, 19-30 Hz), and gamma (γ, 30-70 Hz).

3.3. Network parameters

Three network parameters derived from graph theory were computed from PLI matrices to summarize diverse aspects of global and local brain connectivity: global strength (s), characteristic path length (L), and clustering coefficient (C).

s is a basic network metric that is commonly used as a measure of density, representing the total 'wiring cost' of the network [41]. It is computed as the mean number of links connected to the nodes:

$$s = \frac{1}{N}\sum_{i\in N} s_i = \frac{1}{N}\sum_{i\in N}\sum_{j\in N} a_{ij}, \quad (2)$$

where a_{ij} is the PLI value between nodes i and j and N is the total number of nodes.

L is the most commonly used measure of functional integration. Brain integration resumes the ability to combine specialized information from distributed brain regions. Anatomically, paths are sequences of nodes and links that represent potential routes of information flow between different brain regions. Specifically, L is defined as the average shortest path length between all pairs of nodes in the network [41]:

$$L = \frac{1}{N}\sum_{i\in N}\frac{\sum_{j\in N, j\neq i} l_{ij}}{N-1}, \quad (3)$$

where l_{ij} is the shortest path length (distance) between nodes i and j.

C is a simple segregation metric based on the number of triangles in the network. Segregation quantifies the ability of the brain to process information in

densely interconnected groups of brain regions. Mathematically, C reflects the prevalence of clustered connectivity around individual nodes and represents the fraction of the node neighbors that are also neighbors of each other [41]:

$$C = \frac{1}{N}\sum_{i\in N}\frac{2t_i}{s_i(s_i-1)}, \tag{4}$$

where t_i is the geometric mean of triangles around node i and s_i is the strength of a node i.

Since PLI results are frequency dependent, one value for every network parameter was obtained for each of the previously defined frequency bands. These frequency-specific values were then used to compute the multiplex parameters by means of CCA.

3.4. Canonical Correlation Analysis (CCA)

CCA is a statistical method for measuring the linear relationship between two datasets of multidimensional variables $X \in \Re^{m\times N}$ and $Y \in \Re^{n\times N}$, where N is the number of observations which must be the same for both datasets, and m and n are the number of variables in each dataset, respectively [42]. CCA can be defined as the problem of finding two sets of basis vectors $w \in \Re^{m\times 1}$ and $v \in \Re^{n\times 1}$, such that the correlations between the projections of the variables onto these basis vectors are mutually maximized [42]:

$$\rho = \arg\max_{w,v}\frac{w^T XY^T v}{\sqrt{w^T XX^T w}\sqrt{v^T YY^T v}} \tag{5}$$

3.5. CCA: A new way to build multiplex networks

In this study, we propose to use the basis vectors obtained with CCA to reduce the frequency band dimensionality of every network parameter and to obtain multiplex metrics. Specifically, electrode and source-level network parameters in every frequency band formed the two datasets of multidimensional variables: $X, Y \in \Re^{B\times N}$ and $w, v \in \Re^{B\times 1}$ where N is the number of observations, and $B = 6$ corresponds to the six conventional EEG frequency bands. Then, we computed new CCA multiplex parameters as the projection of the results onto these hyperplanes, which is equivalent to compute the linear combination of the frequency-dependent results multiplied by the basis vectors. In terms of notation, our new CCA multiplex parameters are defined as s_{CCA}, L_{CCA}, and C_{CCA} corresponding to those obtained from the classical measures in each frequency band s, L, and C, respectively.

Mathematically, s_{CCA} is defined as:

$$s_{CCA} = \sum_{b=1}^{B} s_b \cdot v_b, \tag{6}$$

where s_b is the value of s in the bth frequency band and v_b is the bth value of the basis vectors obtained with CCA. L_{CCA} and C_{CCA} can be defined analogously.

3.6. Statistical analysis

For both analyses, an exploratory analysis was carried out to assess the distribution of the CCA multiplex values with the synthetic and real EEG recordings. As our results did not meet the parametric test assumptions of normality (Shapiro-Wilk test) and homoscedasticity (Levene's test), statistical differences between groups were evaluated with the Mann-Whitney U-test.

4. Results

4.1. Synthetic signals

In order to analyze the behavior of the new CCA multiplex parameters depending on the level of connectivity, they were calculated from the time series generated with the Kuramoto model. Firstly, we calculated the source-level basis vector for each network parameter obtained with the less connected signals and then projected the results from the rest of the groups onto these hyperplanes.

Distribution results for s_{CCA}, L_{CCA}, and C_{CCA} using synthetic signals are shown in figures 2(a)-(c), respectively. Firstly, our s_{CCA} results revealed an increasing trend as connectivity increases. It is also possible to observe that differences between consecutive groups are getting larger as the Kuramoto global coupling parameter (K) increases. This trend is similar to the one expected for the classical network parameter s: as connectivity increases, networks are more connected and the values of s become higher. Secondly, L_{CCA} followed the opposite trend, as L_{CCA} results were smaller for higher values of global coupling. These results are in line with the one expected for L, as connectivity increases will lead to more integrated networks, associated with lower values of L. Finally, our results using C_{CCA} showed a growing trend as connectivity increases. This increasing segregation trend was also supported by the expected values of C for each frequency band. With higher connected nodes, more segregated networks will be obtained, since more node neighbors could be also neighbors of each other.

To sum up, our proposed method provides results that are able to summarize the available information across all layers in one single value, overcoming the existing problems derived from the traditional frequency-dependent network analysis. Additionally, the proposed multiplex parameters followed a similar behavior compared with the classical network parameters. Then, they can be interpreted

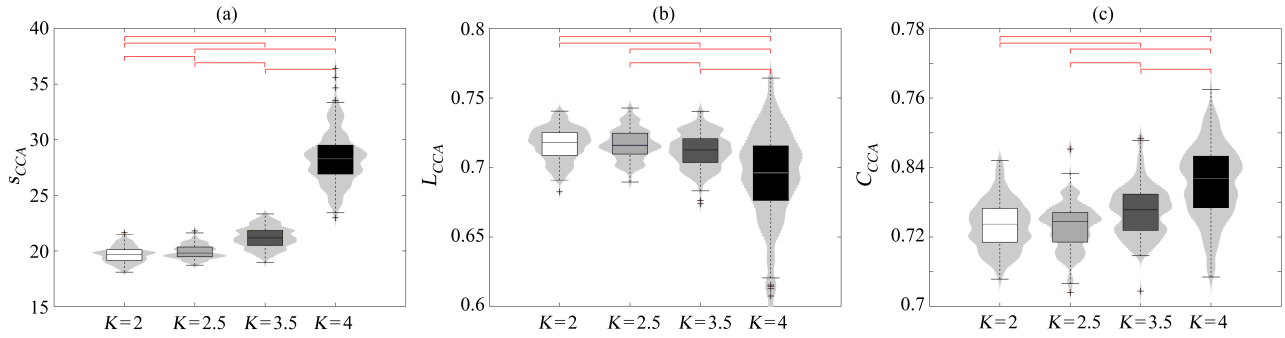

Figure 2. s_{CCA}, L_{CCA}, and C_{CCA} value distributions obtained using synthetic signals as a function of the Kuramoto model global coupling intensity (K). Statistically significant pairwise differences are marked with red brackets ($p < 0.05$, Mann-Whitney U-test).

in similar network terms: multiplex global average connectivity for s_{CCA}, multiplex network integration for L_{CCA}, and multiplex network segregation for C_{CCA}.

4.2. Real EEG recordings

The distribution plots for s_{CCA}, L_{CCA}, and C_{CCA} obtained using the real EEG recordings from HC and the different AD severity groups are displayed in figure 3. Statistically significant differences between groups (p-values < 0.05, Mann–Whitney U-test) are also displayed.

Firstly, in the case of s_{CCA}, our results showed that HC subjects displayed low average connectivity levels and they rise as the disease progresses, excluding the case of AD_{mil} that had the lowest connectivity values. Mann-Whitney U-tests revealed that the AD_{sev} group could be differentiated from the rest of the groups, as statistically significant differences were found for all these comparisons. Significant differences between AD_{mil} and AD_{mod} s_{CCA} values were also found. Secondly, L_{CCA} results showed slightly higher values as dementia severity increases, being only statistically significant for HC vs AD_{sev} comparison. Finally, C_{CCA} results revealed more segregated multiplex networks as the disease progresses, as it showed a decreasing trend from HC subject to AD_{sev} patients. Only AD_{mil} were slightly shifted up from this trend. The Mann-Whitney U-tests revealed statistically significant differences between AD_{sev} patients and all the rest of the groups, HC subjects vs. AD_{mod}, and AD_{mil} vs. AD_{mod} comparisons.

5. Discussion

In the present study, we proposed a new methodology to build multiplex networks based on CCA. Additionally, we evaluated the behavior of these new metrics using synthetic signals varying their connectivity, and we analyzed their ability to characterize the brain network alterations during the different stages of AD continuum.

5.1. CCA multiplex network parameter interpretations

In order to asses the interpretation of the new CCA-based multiplex network parameters, a set of synthetic signals was used. This approach allowed us to study the behavior of the proposed multiplex parameters (s_{CCA}, L_{CCA}, and C_{CCA}) in different connectivity scenarios and relate them with their frequency-specific counterparts. Previous studies have evaluated the performance of various connectivity metrics by simulating different scenarios with the well-studied Kuramoto model [33, 40, 43]. To the best of our knowledge, no previous work has analyzed the performance of frequency-specific network parameters nor multiplex ones.

In the frequency-specific case, the strength represents the sum of the weighted links connected to that node [41]. Then, the global strength s is most commonly used as a measure of density, with higher values for more connected networks [41]. Our results using s_{CCA} followed the same trend, showing bigger increases between consecutive groups as connectivity grows. Thus, the interpretation of the CCA-based multiplex parameter obtained from s can be analogous to the frequency-specific metric.

The characteristic path length L is most commonly used integration metric and measures the average shortest path length between all pair-wise nodes [41, 44]. Theoretically, more disconnected networks lead to higher values of L. In the extreme case, paths between disconnected nodes are defined to have infinite length [41]. This trend is also observed for our multiplex L_{CCA} results, as lower values are obtained as connectivity grows. Consequently, connectivity increases elicits decreases in the values of L_{CCA}, which can be interpreted as more integrated multiplex networks.

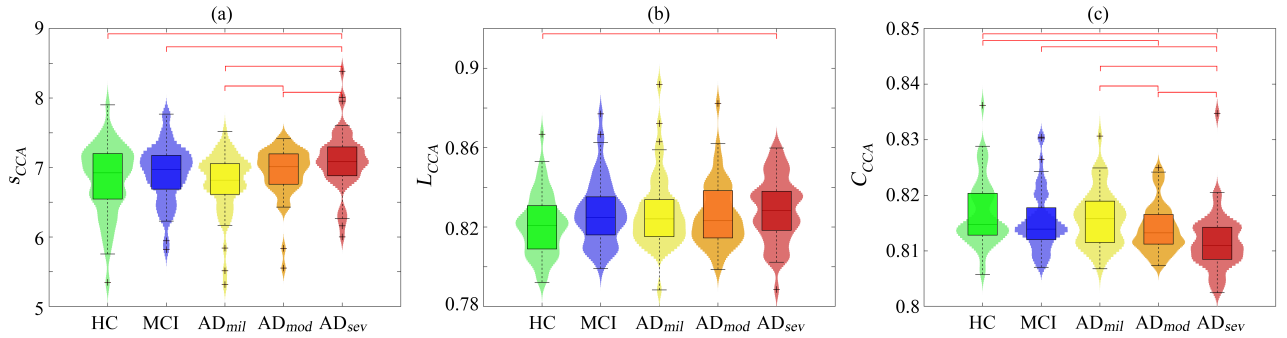

Figure 3. s_{CCA}, L_{CCA}, and C_{CCA} value distributions obtained using real EEG recordings from subjects at different stages of AD continuum: HC subjects, MCI patients, AD_{mil} patients, AD_{mod} patients, and AD_{sev} patients. Statistically significant pairwise differences are marked with red brackets ($p < 0.05$, Mann-Whitney U-test).

The mean clustering coefficient C is a simple measure of segregation, computed as the average fraction of triangles around an individual node [41, 44]. Therefore, lower connected networks result in higher segregated network, as nodes with a low strength may not have clustered connectivity around them [41]. When our proposed CCA-based multiplex approach is applied, C_{CCA} results followed the same trend, showing higher segregation for more connected multiplex networks.

Therefore, considering these results, we can conclude that the interpretation of the CCA-based multiplex parameters can be analogous to their frequency-specific counterparts, as they revealed the same behavior using synthetic signals. Additionally, our proposed method is able to capture and summarize all the relevant information of the whole brain network, while a single-layer approach only provides a limited perspective of the neural network properties.

5.2. Abnormal multiplex network patterns in AD progression

Network organization has conventionally been analyzed in diverse frequency bands, as they are considered to be associated with different cognitive processes. Previous EEG studies have shown that functional connectivity patters in AD networks are frequency-dependent [24, 45]. Specifically, in our database, network topologies in low-frequency bands tend to be less integrated and more segregated as the disease severity increases, reducing their global performance and enhancing their local performance [27]. Here, we integrate all frequency-specific networks in a multiplex framework using CCA to study the brain changes during AD continuum. Our results indicated that multiplex networks are characterized by higher average connectivity as AD severity increases, revealed by an increasing trend in s_{CCA}. Furthermore, these networks tended to be less integrated (revealed by a increase in L_{CCA}) and less segregated (decrease in C_{CCA}) as the disease progresses.

Firstly, a consensus with decreased functional connectivity in alpha band has been reached by previous studies [11, 33, 46, 47]. However, inconsistent results have been found for other frequency bands [33, 48]. These frequency-dependent connectivity variations lead to contradictory results in terms of global strength, which depends on the frequency under analysis. Our previous work using the same database supports these frequency-dependent results, as significant increases of s in the theta band and significant decreases of s in the alpha band were reported as the disease severity increases [27]. To the best of our knowledge, no previous study has analyzed the mean average connectivity in multiplex networks during AD continuum. This could be due to this parameter alone does not give further information about network organization, but it can be useful combined with the integration and segregation information given by L_{CCA} and C_{CCA} results, respectively.

Secondly, our L_{CCA} results are in line with previous studies that reported lower integration in AD networks compared to HC subjects, revealed by an increase in the characteristic path length values [26], a decrease in the network global efficiency [13], and a decrease in the participation coefficient [49]. From the clinical point of view, the global propensity to facilitate information exchange is progressively reduced as a consequence of neurodegeneration as dementia severity increases [50].

Finally, many previous studies showed decreases in global segregation for the AD group using both multiplex and frequency-specific networks [12, 51, 52], in line with our C_{CCA} results. Clinically, this could reflect that local information exchange between frequency bands is progressively disrupted as a consequence of AD [52]. However, increases in clustering coefficient are also reported as a possible

compensatory mechanism that is triggered by the dysfunctional integration in the AD brain networks [25, 53]. This compensatory effect could explain the C_{CCA} values obtained for the AD_{mil} group, as it did not follow the declining trend of C_{CCA} as AD severity increases.

Putting these results together, our findings indicated that dementia due to AD is characterized by a loss of 'small-world' networks as the disease progresses (revealed by decreases in integration and segregation) [16, 54]. This progressive loss of small-world network properties would increase multiplex networks vulnerability [55].

5.3. Limitations and future research lines

Despite the fact that we have reduced the volume-conduction bias (by using *PLI* to build the connectivity network), reduced the frequency-specific results bias (by adopting a multiplex approach), and facilitated the comparison with other EEG-channel configurations and neuroimaging methods (by using source-level results obtained with a standard atlas), the present study has some limitations that should be considered.

Firstly, multiplex network parameters were averaged for all ROIs. Future studies should be carried out taking into account the specific value for each ROI. These spatial analyses could lead us to accurately identify brain areas with abnormal neural patters to gain further insights on AD continuum neurodegeneration processes.

Secondly, even if we had a large database composed of 253 recordings to characterize the brain abnormalities during the AD progression, subjects were divided into five groups depending on their stage. Thus, each group was formed by only 50 or 51 subjects; this sample size is not big enough to perform a classification analysis and prove the usefulness of our proposed methodology as a generalizable diagnostic tool.

Finally, despite we had the AD continuum stratified into five severity groups, it would be useful to perform a longitudinal study focused only on MCI subjects. This group is the most clinically interesting group, since two distinct subgroups of patients with MCI can be distinguished: those that remain stable at MCI condition, and those that progress to AD. This research could provide deeper understanding on the complex neural changes during MCI and may allow to predict AD progression at an early stage.

6. Conclusions

In the present study, an original methodology to build multiplex networks is proposed, evaluated with synthetic signals, and applied to real EEG recordings from patients in different stages of AD continuum. We proved that the proposed CCA-based multiplex parameters can be interpreted in an analogous way compared to their frequency-specific counterparts, as they revealed the same behavior against coupling variations using synthetic signals. Furthermore, our results using real EEG recordings revealed that dementia due to AD altered progressively the small-world network properties, increasing multiplex networks vulnerability as a consequence of the neurodegenerative processes.

Acknowledgments

This research was supported by 'Ministerio de Ciencia e Innovación - Agencia Estatal de Investigación' and 'European Regional Development Fund' (FEDER) under project PGC2018-098214-A-I00, by 'European Commission' and 'FEDER' under projects 'Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer' and 'Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres' ('Cooperation Programme Interreg V-A Spain-Portugal POCTEP 2014–2020'), and by 'CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' through 'Instituto de Salud Carlos III' co-funded with FEDER funds. Saúl J. Ruiz-Gómez, Eduardo Santamaría-Vázquez, and Aarón Maturana-Candelas were in receipt of a predoctoral scholarship from the 'Junta de Castilla y León' and the 'European Social Fund'. Víctor Rodríguez-González was in receipt of a PIF-UVa grant from the 'University of Valladolid'.

References

[1] Mijalkov M, Kakaei E, Pereira J B, Westman E and Volpe G 2017 BRAPH: A graph theory software for the analysis of brain connectivity *PLoS ONE* **12** 1–23

[2] D'Amelio M and Rossini P M 2012 Brain excitability and connectivity of neuronal assemblies in Alzheimer's disease: From animal models to human findings *Progress in Neurobiology* **99** 42–60

[3] Alzheimer's Association 2019 2019 Alzheimer's disease facts and figures *Alzheimer's & Dementia*

[4] McKhann G M, Knopman D S, Chertkow H, Hyman B T, Jack C R, Kawas C H, Klunk W E, Koroshetz W J, Manly J J, Mayeux R, Mohs R C, Morris J C, Rossor M N, Scheltens P, Carrillo M C, Thies B, Weintraub S and Phelps C H 2011 The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease *Alzheimer's and Dementia* **7** 263–269

[5] Aisen P S, Cummings J, Jack C R, Morris J C, Sperling R, Frölich L, Jones R W, Dowsett S A, Matthews B R, Raskin J, Scheltens P and Dubois B 2017 On the path to

2025: Understanding the Alzheimer's disease continuum *Alzheimer's Research and Therapy* **9** 1–10

[6] Petersen R C 2016 Mild cognitive impairment *Continuum* **22** 404–418

[7] McBride J C, Zhao X, Munro N B, Smith C D, Jicha G A, Hively L, Broster L S, Schmitt F A, Kryscio R J and Jiang Y 2014 Spectral and complexity analysis of scalp EEG characteristics for mild cognitive impairment and early Alzheimer's disease *Computer Methods and Programs in Biomedicine* **114** 153–163

[8] Buscema M, Vernieri F, Massini G, Scrascia F, Breda M, Rossini P M and Grossi E 2015 An improved I-FAST system for the diagnosis of Alzheimer's disease from unprocessed electroencephalograms by using robust invariant features *Artificial Intelligence in Medicine* **64** 59–74

[9] Poza J, Gómez C, García M, Tola-Arribas M A, Carreres A, Cano M and Hornero R 2017 Spatio-Temporal Fluctuations of Neural Dynamics in Mild Cognitive Impairment and Alzheimer's Disease *Current Alzheimer research* **14** 924–936

[10] Ruiz-Gómez S J, Gómez C, Poza J, Gutiérrez-Tobal G C, Tola-Arribas M A, Cano M and Hornero R 2018 Automated multiclass classification of spontaneous EEG activity in Alzheimer's disease and mild cognitive impairment *Entropy* **20** 1–15

[11] Dauwels J, Vialatte F and Cichocki A 2010 Diagnosis of Alzheimer's disease from EEG signals: where are we standing? *Current Alzheimer Research* **7** 487–505

[12] Stam C J, De Haan W, Daffertshofer A, Jones B F, Manshanden I, Van Cappellen Van Walsum A M, Montez T, Verbunt J P, De Munck J C, Van Dijk B W, Berendse H W and Scheltens P 2009 Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease *Brain* **132** 213–224

[13] Afshari S and Jalili M 2017 Directed Functional Networks in Alzheimer's Disease: Disruption of Global and Local Connectivity Measures *IEEE Journal of Biomedical and Health Informatics* **21** 949–955

[14] Stam C J, Jones B F, Nolte G, Breakspear M and Scheltens P 2007 Small-world networks and functional connectivity in Alzheimer's disease *Cerebral Cortex* **17** 92–99

[15] Lo C Y, Wang P N, Chou K H, Wang J, He Y and Lin C P 2010 Diffusion tensor tractography reveals abnormal topological organization in structural cortical networks in Alzheimer's disease *Journal of Neuroscience* **30** 16876–16885

[16] Sanz-Arigita E J, Schoonheim M M, Damoiseaux J S, Rombouts S A, Maris E, Barkhof F, Scheltens P and Stam C J 2010 Loss of 'Small-World' Networks in Alzheimer's Disease: Graph Analysis of fMRI Resting-State Functional Connectivity *PLoS ONE* **5** 1–14

[17] Tijms B M, Möller C, Vrenken H, Wink A M, de Haan W, van der Flier W M, Stam C J, Scheltens P and Barkhof F 2013 Single-Subject Grey Matter Graphs in Alzheimer's Disease *PLoS ONE* **8** 1–9

[18] He Y, Chen Z and Evans A 2008 Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease *Journal of Neuroscience* **28** 4756–4766

[19] Yao Z, Zhang Y, Lin L, Zhou Y, Xu C and Jiang T 2010 Abnormal cortical networks in mild cognitive impairment and alzheimer's disease *PLoS Computational Biology* **6** e1001006

[20] Zhao X, Liu Y, Wang X, Liu B, Xi Q, Guo Q, Jiang H, Jiang T and Wang P 2012 Disrupted small-world brain networks in moderate Alzheimer's disease: A resting-state fMRI study *PLoS ONE* **7** e33540

[21] Battiston F, Nicosia V and Latora V 2014 Structural measures for multiplex networks *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **89** 032804

[22] De Domenico M, Sasai S and Arenas A 2016 Mapping multiplex hubs in human functional brain networks *Frontiers in Neuroscience* **10** 1–14

[23] Boccaletti S, Bianconi G, Criado R, del Genio C I, Gómez-Gardeñes J, Romance M, Sendiña-Nadal I, Wang Z and Zanin M 2014 The structure and dynamics of multilayer networks *Physics Reports* **544** 1–122

[24] Yu M, Engels M M, Hillebrand A, Van Straaten E C, Gouw A A, Teunissen C, Van Der Flier W M, Scheltens P and Stam C J 2017 Selective impairment of hippocampus and posterior hub areas in Alzheimer's disease: An MEG-based multiplex network study *Brain* **140** 1466–1485

[25] Kabbara A, Eid H, El Falou W, Khalil M, Wendling F and Hassan M 2018 Reduced integration and improved segregation of functional brain networks in Alzheimer's disease *Journal of Neural Engineering* **15** 026023

[26] Stam C J 2014 Modern network science of neurological disorders *Nature Reviews Neuroscience* **15** 683–695

[27] Ruiz-Gómez S J, Gómez C, Poza J, Marcos R V, Gutiérrez-de Pablo V, Rodríguez-González V and Maturana-Candelas A 2020 Volume Conduction Effects on Connectivity Metrics: Application of Network Parameters to Characterize Alzheimer's Disease Continuum *Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Montreal (Canada)* 30–33

[28] Buldú J M and Porter M A 2018 Frequency-based brain networks: From a multiplex framework to a full multilayer description *Network Neuroscience* **2** 418–441 (*Preprint* `1703.06091`)

[29] Zhang Y, Yin E, Li F, Zhang Y, Tanaka T, Zhao Q, Cui Y, Xu P, Yao D and Guo D 2018 Two-Stage Frequency Recognition Method Based on Correlated Component Analysis for SSVEP-Based BCI *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **26** 1314–1323

[30] Santamaría-Vázquez E, Martínez-Cagigal V, Gomez-Pilar J and Hornero R 2019 Asynchronous Control of ERP-Based BCI Spellers Using Steady-State Visual Evoked Potentials Elicited by Peripheral Stimuli *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **27** 1883–1892

[31] Karageorgiou E, Lewis S M, McCarten J R, Leuthold A C, Hemmy L S, McPherson S E, Rottunda S J, Rubins D M and Georgopoulos A P 2012 Canonical correlation analysis of synchronous neural interactions and cognitive deficits in Alzheimer's dementia *Journal of Neural Engineering* **9** 056003

[32] Cui D, Qi S, Gu G, Li X, Li Z, Wang L and Yin S 2018 Magnitude Squared Coherence Method based on Weighted Canonical Correlation Analysis for EEG Synchronization Analysis in Amnesic Mild Cognitive Impairment of Diabetes Mellitus *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **26** 1908–1917

[33] Ruiz-Gómez S J, Hornero R, Poza J, Maturana-Candelas A, Pinto N and Gómez C 2019 Computational modeling of the effects of EEG volume conduction on functional connectivity metrics. Application to Alzheimer's disease continuum *Journal of Neural Engineering* **16** 066019

[34] US National Library of Medicine The Visible Human Project URL `https://www.nlm.nih.gov/research/visible/visible_human.html`

[35] Yushkevich P A, Piven J, Hazlett H C, Smith R G, Ho S, Gee J C and Gerig G 2006 User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability *NeuroImage* **31** 1116–1128

[36] Albert M S, DeKosky S T, Dickson D, Dubois B, Feldman H H, Fox N C, Gamst A, Holtzman D M, Jagust W J, Petersen R C, Snyder P J, Carrillo M C, Thies B and Phelps C H 2011 The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease *Alzheimer's and Dementia* **7** 270–279

[37] Ruiz-Gómez S J, Gómez C, Poza J, Martínez-Zarzuela M, Tola-Arribas M A, Cano M and Hornero R 2018 Measuring alterations of spontaneous EEG neural coupling in alzheimer's disease and mild cognitive impairment by means of cross-entropy metrics *Frontiers in Neuroinformatics* **12** 1–11

[38] Tadel F, Baillet S, Mosher J C, Pantazis D and Leahy R M 2011 Brainstorm: A user-friendly application for MEG/EEG analysis *Computational Intelligence and Neuroscience* **879716** 1–13

[39] Pascual-Marqui R D 2002 Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details *Methods and Findings in Experimental and Clinical Pharmacology* **24 Suppl D** 5–12

[40] Stam C J, Nolte G and Daffertshofer A 2007 Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources *Human Brain Mapping* **28** 1178–1193

[41] Rubinov M and Sporns O 2010 Complex network measures of brain connectivity: Uses and interpretations *NeuroImage* **52** 1059–1069

[42] Jobson J D 2012 *Applied Multivariate Data Analysis: Volume II: Categorical and Multivariate Methods: Categorical and Multivariate Methods* (New York, NY, USA: Springer)

[43] Moretti D V, Frisoni G B, Pievani M, Rosini S, Geroldi C, Binetti G and Rossini P M 2008 Cerebrovascular disease and hippocampal atrophy are differently linked to functional coupling of brain areas: An EEG coherence study in MCI subjects *Journal of Alzheimer's Disease* **14** 285–299

[44] Watts D and Strogatz S 1998 Collective dynamics of 'small world' networks *Nature* **393** 440–442

[45] Hillebrand A, Tewarie P, Van Dellen E, Yu M, Carbo E W, Douw L, Gouw A A, Van Straaten E C and Stam C J 2016 Direction of information flow in large-scale resting-state networks is frequency-dependent *Proceedings of the National Academy of Sciences of the United States of America* **113** 3867–3872

[46] Wang R, Wang J, Yu H, Wei X, Yang C and Deng B 2014 Decreased coherence and functional connectivity of electroencephalograph in Alzheimer's disease *Chaos* **24** 033136

[47] Babiloni C, Del Percio C, Lizio R, Noce G, Lopez S, Soricelli A, Ferri R, Nobili F, Arnaldi D, Famà F, Aarsland D, Orzi F, Buttinelli C, Giubilei F, Onofrj M, Stocchi F, Stirpe P, Fuhr P, Gschwandtner U, Ransmayr G, Garn H, Fraioli L, Pievani M, Frisoni G B, D'Antonio F, De Lena C, Güntekin B, Hanoğlu L, Başar E, Yener G, Emek-Savaş D D, Triggiani A I, Franciotti R, Taylor J P, Vacca L, De Pandis M F and Bonanni L 2018 Abnormalities of resting-state functional cortical connectivity in patients with dementia due to Alzheimer's and Lewy body diseases: an EEG study *Neurobiology of Aging* **65** 18–40

[48] Hata M, Kazui H, Tanaka T, Ishii R, Canuet L, Pascual-Marqui R D, Aoki Y, Ikeda S, Kanemoto H, Yoshiyama K, Iwase M and Takeda M 2016 Functional connectivity assessed by resting state EEG correlates with cognitive decline of Alzheimer's disease - An eLORETA study *Clinical Neurophysiology* **127** 1269–1278

[49] De Haan W, Van der Flier W M, Koene T, Smits L L, Scheltens P and Stam C J 2012 Disrupted modular brain dynamics reflect cognitive dysfunction in Alzheimer's disease *NeuroImage* **59** 3085–3093

[50] Guillon J, Attal Y, Colliot O, La Corte V, Dubois B, Schwartz D, Chavez M and De Vico Fallani F 2017 Loss of brain inter-frequency hubs in Alzheimer's disease *Scientific Reports* **7** 1–13

[51] Supekar K, Menon V, Rubin D, Musen M and Greicius M D 2008 Network analysis of intrinsic functional brain connectivity in Alzheimer's disease *PLoS Computational Biology* **4** 1–11

[52] Cai L, Wei X, Liu J, Zhu L, Wang J, Deng B, Yu H and Wang R 2020 Functional Integration and Segregation in Multiplex Brain Networks for Alzheimer's Disease *Frontiers in Neuroscience* **14** 1–14

[53] Palesi F, Castellazzi G, Casiraghi L, Sinforiani E, Vitali P, Gandini Wheeler-Kingshott C A and D'Angelo E 2016 Exploring patterns of Alteration in Alzheimer's disease brain networks: A combined structural and functional connectomics analysis *Frontiers in Neuroscience* **10** 1–16

[54] Bullmore E and Sporns O 2009 Complex brain networks: graph theoretical analysis of structural and functional systems *Nature Reviews Neuroscience* **10** 186—-198

[55] Vecchio F, Miraglia F, Piludu F, Granata G, Romanello R, Caulo M, Onofrj V, Bramanti P, Colosimo C and Rossini P M 2017 'Small World' architecture in brain connectivity and hippocampal volume in Alzheimer's disease: a study via graph theory from EEG data *Brain Imaging and Behavior* **11** 473–485