Permutation Jaccard Distance-Based Hierarchical Clustering to Estimate EEG Network Density Modifications in MCI Subjects

Permutation Jaccard Distance-based Hierarchical Clustering to estimate EEG network density modifications in MCI subjects

Nadia Mammone, *Member IEEE*, Cosimo Ieracitano, Hojjat Adeli, *Fellow IEEE*, Alessia Bramanti and Francesco C. Morabito, *Senior Member IEEE*

***Abstract*—In this paper, a novel EEG-based method is introduced for the quantification of brain-electrical connectivity changes over a longitudinal evaluation of Mild Cognitive Impaired (MCI) subjects. In the proposed method, a dissimilarity matrix is costructed by estimating the coupling strength between every pair of EEG signals, Hierarchical Clustering (HC) is then applied to group the related electrodes according to the dissimilarity estimated on pairs of EEG recordings. Subsequently, the connectivity density of the electrodes network is calculated. The technique was tested over two different coupling strength descriptors: Wavelet Coherence (WC) and Permutation Jaccard Distance (PJD), a novel metric of coupling strength between time series introduced in the present work. Twentyfive MCI patients were enrolled within a follow-up program that consisted of two successive evaluations, at time T0 and at time T1, three months later. At T1, four subjects were diagnosed to have converted to Alzheimers Disease (AD). When applying the PJD-based method, the converted patients exhibited a significantly increased PJD ($p < 0.05$), i.e., a reduced overall coupling strength, specifically in delta and theta bands and in the overall range (0.5-32Hz). In addition, in contrast to stable MCI patients, converted patients exhibited a network density reduction in every sub-band (delta, theta, alpha, beta). When WC was used as coupling strength descriptor, the method resulted in a less sensitive and specific outcome. The proposed method, mixing nonlinear analysis to a machine learning approach, appears to provide an objective evaluation of the connectivity density modifications associated to the MCI-AD conversion, just processing non-invasive EEG signals.**

***Index Terms*—EEG, Mild Cognitive Impairment, Alzheimer's Disease, Permutation Entropy, Permutation Jaccard Distance, Brain Connectivity, Hierarchical Clustering, Network density.**

I. Introduction

Today, 47 million people live with dementia worldwide. This number is expected to increase to more than 131 million by 2050, together with the average age of the population. Alzheimer's Disease (AD) represents nearly 60% of the total dementia cases [1]. AD is typically diagnosed after the age of 65. AD patients survive on average only 4 to 8 years after diagnosis, because AD is nowadays still intractable [2], [3], [4], [5], [6], [7]. The National Institute on Aging and the Alzheimer's Association (NIA/AA) workgroup of experts postulated that what is commonly considered Alzheimer's should rather be considered a stage of a long and complex degenerating process [8]. They have hypothesized three stages of AD progression:

1) Preclinical-AD: when the disease has already triggered the brain degeneration but no clinical symptoms are visible yet; 2) Mild Cognitive Impairment (MCI) due to AD (MCI-AD or Prodromal-AD): an intermediate stage where symptoms related to the thinking ability may start to be noticeable, but they do not affect the daily life of the subject; 3) Dementia due to AD (Dementia-AD): In the last stage, impairments in memory, thinking and behaviour undermine a person's ability to live and act independently.

Therefore, longitudinal (follow-up) studies on MCI patients and normal aging subjects can be of great help in the development of biomarkers for monitoring the progression of dementias. Unfortunately, keeping the patients and their caregivers loyal to a followup program is challenging, resulting in a lack of longitudinal studies in the literature. MCI subjects can be classified as: amnestic MCI (aMCI) and non-amnestic MCI [9], aMCI subjects are more likely to convert to AD [10]. The aim of the present paper is to propose a method, based on hierarchical clustering, for studying longitudinal changes in the brain-electrical connectivity density. The method will be specifically tested over longitudinal changes in the brain-electrical connectivity density associated with the conversion from MCI to AD. The method is based on the analysis of Electroencephalographic (EEG) recordings because EEG is a quick, cheap, non-invasive and well tolerated neurophysiological measurement which has provided encouraging results in MCI and AD diagnosis [11] [12] [13] and could be the basis of future systems for the early-diagnosis of AD. EEG is the most popular tool for investigating the cerebral electrical activity [14]. Abnormal patterns in the electrical potentials detected on the scalp may indeed reflect abnormalities in the communication between neurons and can be used as diagnostic and prognostic markers [14], [15], [16].

Ahmadlou et al. [17] present a chaos-wavelet methodology for EEG-based diagnosis of AD employing the concept of Visibility Graph (VG) from the graph theory. They report a high diagnostic accuracy of 97.7% using a two-stage classifier consisting of Radial Basis Function Neural Network

N. Mammone is with the IRCCS Centro Neurolesi Bonino-Pulejo, Via Palermo c/da Casazza, SS. 113, 98124 Messina, Italy, e-mail: nadia.mammone@irccsme.it, nadiamammone@tiscali.it.

C. Ieracitano and F. C. Morabito are with the Department DICEAM of the *Mediterrean* University of Reggio Calabria, Via Graziella Feo di Vito, 89060 Reggio Calabria, Italy.

H. Adeli is with the Ohio State University, College of Engineering, 470 Hitchcock Hall, 2070 Neil Avenue, Columbus, Ohio.

A. Bramanti is with the Institute of Applied Sciences and Intelligent Systems *Eduardo Caianiello* (ISASI), National Research Council (CNR), Messina (Italy)

and Principal Component Analysis. Ahmadlou et al. [18] use two different fractal dimensions (FD), Katzs FD (KFD) and Higuchis FD (HFD), for evaluation of the dynamical changes in the AD brain. They report a high accuracy of 99.3% for diagnosis of the AD based on the global KFD in the beta band of the eyes-closed condition. Moretti et al. observed a decay of the temporoparietal cortex associated with MCI prodromal to AD, additionally, shrinkage of the temporoparietal cerebral area was correlated with an increase in alpha3/alpha2 EEG power ratio [19]. Frantzidis et al. investigated functional disorganization of small-world brain networks in mild AD and aMCI applying Relative Wavelet Entropy (RWE) to the EEG and then a complex network approach to model brain networks functioning. They found a declined functional network organization even during the prodromal stage [20]. Ahmadlou et al. [21] investigate the complexity of functional connectivity networks in MCI patients during a working memory task using magnetoencephalography (MEG) signals. They introduce two measures for brain networks complexities: Graph Index Complexity and Efficiency Complexity. Houmani et al. [22] use epoch-based entropy for early diagnosis of the AD. Ortiz-Garcia et al. [23] use ensembles of deep neural network learning architectures for the early diagnosis of the AD. F. J. Martinez-Murcia et al. [24] study the structural parametrization of the brain using hidden Markov models-based paths in AD.

Recently, Amezquita-Sanchez et al. [25] presented a novel 4-step methodology for identifying MCI patients during a working memory task using MEG signals. The enhanced probabilistic neural network classifier of Ahmadou and Adeli [26] is employed to distinguish MCI patients from healthy subjects. They report a high accuracy of 98.4% to distinguish MCI subjects. A state-of-the art review of imaging and machine learning techniques for diagnosis of AD is presented in a recent article by Mirzaei et al. [27]. Unfortunately, the above mentioned studies are cross-sectional.

There are just a few longitudinal studies on MCI [28], resulting in a lack of diagnostic tools to allows the neurologist to objectively monitor the progression of the disease. Morabito et al. [29] present a longitudinal EEG study of AD progression based on a complex network approach. Romero-Garcia et al. [30] note different scales of cortical organization are selectively targeted during the progression to AD. Buscema et al. [28] proposed a system which exploits special types of artificial neural networks (ANNs) assembled in a novel methodology named IFAST (implicit function as squashing time). They reported that IFAST method was able to predict the conversion from amnestic MCI to AD with high accuracy (85.98%) in a 1-year follow-up study. This methodology was later improved but it has been validated only in a MCI vs AD vs Controls cross sectional study so far [31].

Motivated by the preliminary results achieved with the joint application of standard coherence and hierarchical clustering on AD patients [32], the goal of the present study was to evaluate the evolution of the cortical connectivity in MCI patients, indirectly, through the EEG, within a follow-up program. Let us recall that AD (and therefore MCI, when prodromal to AD) is considered a disconnection disorder because it weakens the connectivity between the different areas of the brain [20], [14], [12], [33]. Changes in the connectivity between cortical areas are expected to induce changes in the coupling strength between the corresponding EEG signals, therefore, a measure of coupling strength between EEG signals should be able to indirectly quantify the changes in the brain-electrical connectivity due the disease's progression [21].

The present work resulted from a translational research whose goal was to compare the electrical connectivity density at time T0 and at time T1, in MCI subjects, in order to assess if it could reflect the brain connectivity reduction that is expected to be induced by the disease's progression towards AD.

To this purpose, a novel method to evaluate changes in the brain-electrical connectivity was introduced: the method is based on the construction of a dissimilarity matrix based on the coupling strength between the EEG electrodes, the application of Hierarchical Clustering to the dissimilarity matrix and the subsequent estimation of the network density as a function of the fusion level. The study involved 25 amnestic MCI patients, who were longitudinally evaluated at time T0 and, 3 months later, at time T1, within a follow-up program.

In this research, the authors first used spectral coherence, as an estimator of the coupling strength, because it is the most common descriptor in the analysis of EEG recorded from MCI/AD patients, coherence-based analysis is indeed routinely used in most MCI/AD EEGs clinical systems [31]. In particular, Wavelet Coherence (WC) definition is used in this research, because it was shown to outperform the standard coherence in the analysis of AD EEGs [34],[35]. Applying the method with WC as coupling strength estimator, it was not possible to detect a common behaviour for the EEGs of the converted patients.

To cope with this problem, the Permutation Jaccard Distance (PJD) is introduced here as a novel symbolic measure of coupling strength between nonlinear time series. Applied to EEG signals, PJD can be interpreted as an indirect measure of the coupling strength between two or more cortical areas, which is estimated by measuring the "distance" between the corresponding projected EEG time series. PJD was hereby defined and implemented on simulated unidirectionally coupled Henon maps, in order to test its ability to measure the coupling strength between interacting dynamic systems. Since PJD successfully reflected the coupling strength variations in the simulated systems, it was applied as coupling strength estimator instead of WC.

When the method was applied using PJD as coupling strength descriptor, the EEGs of the converted patients exhibited a peculiar behaviour: network density reduction in every sub-band from T0 to T1. In contrast, stable MCI subjects did not exhibit such a behaviour. The paper is organized as follows: Section II describes how the patients were recruited and how the EEGs were recorded and preprocessed. Section III introduces the concept of PJD and describes how it was tested both on simulated dynamical interacting systems and on the experimental EEG data. Section IV reports the achieved results, Section V discusses them and Section VI addresses the conclusions.

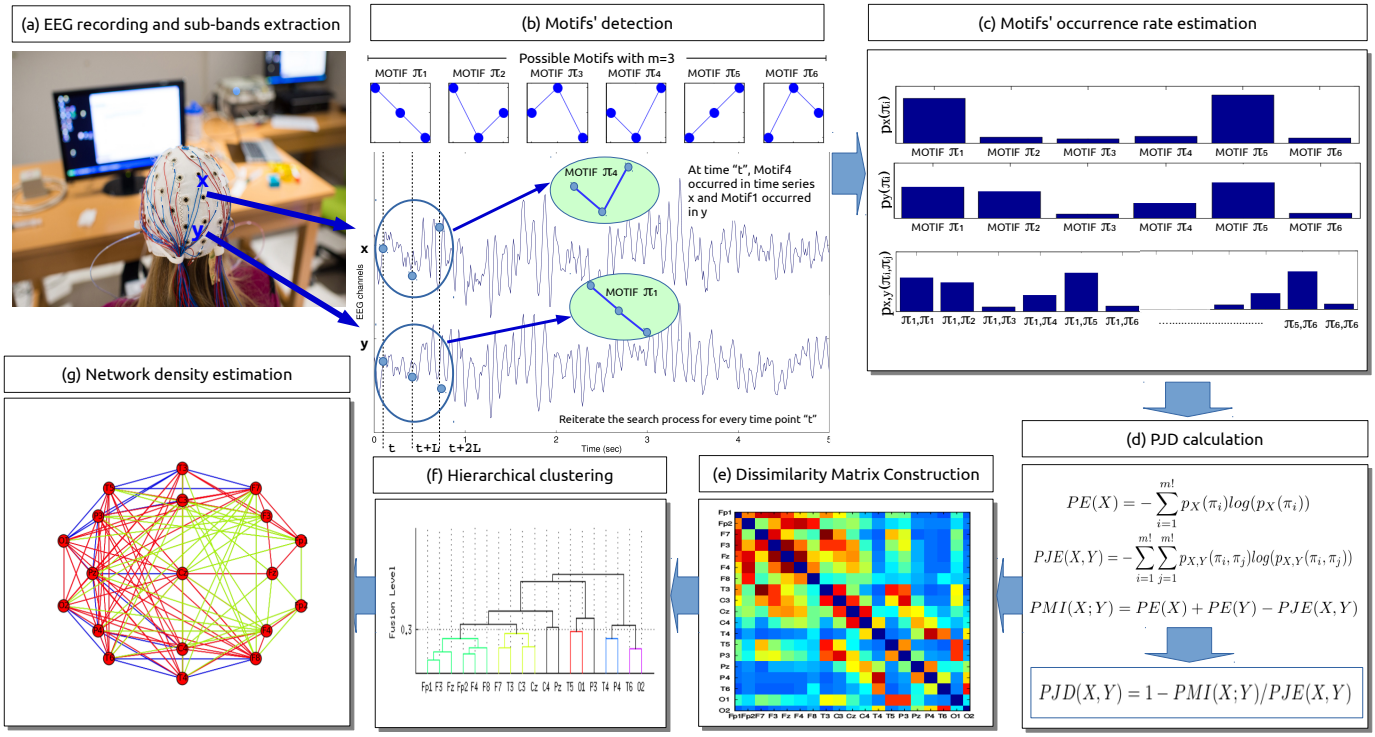

Fig. 1. Given two time series x and y (i. e. two EEG signals), a time point t, an embedding dimension m and a lag L, the two time series can be projected into the vectors X_t and Y_t. When m=3, X_t and Y_t are vectors with 3 elements and 6 possible ordinal patterns (Motifs) can occur. Given a time point t, the algorithm checks which motifs occurred in x and which one in y. In the example illustrated in Figure, at time t, motif π_4 occurred in x and motif π_1 occurred in y. The procedure is reiterated for every time point t so that we can end up with a final occurrence rate of every motif π_i in x ($p_X(\pi_i)$) and in y ($p_Y(\pi_i)$). We also end up with a joint occurrence rate $p_{X,Y}(\pi_i, \pi_j)$ for every possible couple of motifs π_i and π_j, which accounts for the probability that π_i occurs in x and π_j occurs in y.

II. EEG DATA RECORDING AND PREPROCESSING

A. Study population

A study group of 25 amnestic MCI subjects was recruited at the IRCCS Centro Neurolesi *Bonino-Pulejo* of Messina (Italy). Every patient was enrolled according to a cooperation agreement that included a clinical protocol approved by the local Ethical Committee. The patients signed an informed consent form. The clinical diagnosis was formulated according to the guidelines of the Diagnostic and Statistical Manual of Mental Disorders (fifth edition, DSM-5) [36], by a multidisciplinary team of neurologists, psychologists, psychiatrists and EEG experts. The same examiners carried out all the cognitive and clinical evaluations through a complete medical assessment.

Every patient underwent a neuroradiological examination to rule out other pathological conditions (tumors, strokes, damage from severe head trauma or buildup of fluid in the brain, etc). The patients were not undergoing any medical treatment. They were evaluated at a baseline time, T0, and 3 months later (time T1). At the time of the second evaluation, T1, 4 out of 25 subjects resulted converted to AD (Pt 03, Pt 32, Pt 51, Pt 71).

B. EEG recording

The EEG was recorded according to the 10-20 International System (19 channels: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz and Pz), with linked earlobe (A1-A2) reference, during a comfortable resting state.

The sampling rate was set at 1024 Hz and a notch filter at 50Hz was used. Before recording the EEG, every patient and her/his caregivers were questioned about the quality and duration of the last night sleep and about the last meal timing and content. The EEGs were recorded in the morning.

During the acquisition, the patients kept their eyes closed but remained awake. The recorded EEG were later manually reviewed in order to assess that no sleep pattern was visible and to label the segments affected by artifacts.

C. EEG preprocessing

Four major EEG sub-bands are commonly considered when processing the EEGs of MCI/AD patients: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), and beta (12-32 Hz). Each of the four sub-bands refers to different functional and physiological parts of the brain.

The EEG signals were band-pass filtered at 0.5-32 Hz. Every EEG signal was then split into four sub-signals, each one associated with a specific sub-band. The components of the signals were extracted through a set of band-pass filters implemented by the toolbox EEGLab (https://sccn.ucsd.edu/eeglab/) [37]. In particular, the function *eegfiltfft*, based on Fast Fourier Transform (FFT) and inverse FFT, was used to reconstruct the signal in the desired frequency range.

In this way, a n-channels EEG recording is basically split into 4 different n-channels EEG recordings corresponding to the four sub-bands: EEG_δ, EEG_θ, EEG_α, EEG_β. Every

sub-band-EEG was downsampled to 256Hz and it was then partitioned into 5 s non-overlapping windows (epochs).

All the algorithms were implemented in Matlab 2016a (The MathWorks, Inc., Natick, MA, USA).

Before processing the filtered EEG through the proposed algorithm, the artifactual segments, labelled by EEG experts, were discarded. The average time length of the EEG recordings, after artifact cancellation, was 5.9 min.

III. A NEW EEG-BASED METHOD TO INDIRECTLY ESTIMATE THE BRAIN CONNECTIVITY IN MCI PATIENTS

To study the connectivity of the brain network represented by the scalp electrodes (nodes), it is necessary to define a suitable measure of dissimilarity, between every pair of electrodes, *i* and *j*. This metric should be able to interpret the coupling strength between the brain areas *i* and *j* covered by the two electrodes. The inter-electrode dissimilarity is inversely quantified by estimating the coupling strength between the corresponding EEG signals. In this way, one can associate a weighted graph to the EEG recorded at T0 and T1 and the two corresponding symmetric matrices of dissimilarities between the electrodes. In the proposed method, the matrices of dissimilarities between electrodes are the input for the hierarchical clustering (HC) step, which allows to group the electrodes depending on the relative coupling strength between the corresponding EEG signals. The output of HC ia a *dendrogram* representation of the relationship between the nodes. The node clusters will depend on the selected fusion level threshold. Fixed a fusion level threshold, a set of clusters is determined; it is then possible to give an estimate of the network density (i.e., the ratio between the actual number of connections and the total number of possible connections). The connectivity changes of the brain can be indirectly quantified by comparing the connectivity density of the two graphs, at T0 and T1. A schematic view of the proposed methodology is illustrated in Figure 1 and will be discussed in detail, step by step, following the labelling used in Figure 1. The proposed method includes the following steps: a) Recording the EEG and extracting the sub-band signals (delta, theta, alpha, beta) as described in Section II-C; b) For every EEG epoch, detecting which motifs occur within that window; c) Estimating the occurrence rate of the motifs; d) Calculating the PJD between every pair of EEG signals (i.e. every pair of electrodes); e) Constructing a network where the nodes are the electrodes and the dissimilarity between pair of nodes is represented by the PJD between the two corresponding EEG signals (in other words, a dissimilarity matrix is constructed, where the element (i, j) is the PJD between electrodes *i* and *j*; f) Feeding the dissimilarities into hierarchical clustering; g) Estimating the network density, for different fusion level thresholds, according to the resulting clusters (two electrodes are considered "connected" if they belong to the same cluster). Steps 2 to 6 are applied to every EEG sub-band, independently. Given a patient, the methodology is applied either to EEG-T0 and EEG-T1, in order to estimate the connectivity densities at T0 and T1, to compare them and to assess the brain-electrical connectivity changes over the follow-up.

A. The Permutation Jaccard Distance (PJD) as a novel measure of coupling strength between time series

Permutation Entropy (PE) was introduced by Bandt and Pompe [38] as a symbolic descriptor of dynamic complexity changes in time series. PE projects the time trace into symbols (motifs), thus estimating the randomness of the time series regardless of its amplitude, which plays a key role when analysing EEG. In fact, the amplitude of the EEG recorded through a given electrode, depends on the distance from the reference one. When processing the EEG recordings through amplitude dependent techniques, each EEG signal should be first normalized to cancel the effect of closeness to the reference electrode. In contrast, normalization is not necessary when using a symbolic procedure like PE. However, PE is a univariate descriptor which can only describe the randomness of a single time series (i.e. an EEG signal) and cannot quantify the coupling strength between two or more time series (i.e. between two or more EEG signals). The proposed descriptor, PJD, is based on the same projection into symbols adopted by PE, but it is a bivariate descriptor that can quantify the coupling strength between two time series. In what follows, before introducing PJD, some preparatory concepts of Information Theory are briefly discussed.

1) Preparatory Information Theory Concepts: Given a time series *x*, with *N* samples, and its probability density function $p(x)$, the Shannon Entropy of *x* is defined as:

$$H(X) = -\sum_{i=1}^{N} p_X(x_i) log(p_X(x_i)) \tag{1}$$

Given two time series *x* and *y*, with *N* samples, and their joint probability density function $p_{X,Y}(x, y)$, their Joint Entropy is defined as:

$$H(X,Y) = -\sum_{i=1}^{N}\sum_{j=1}^{N} p_{X,Y}(x_i, y_j) log(p_{X,Y}(x_i, y_j)) \tag{2}$$

The Mutual Information between *x* and *y* is defined as:

$$MI(X;Y) = H(X) + H(Y) - H(X,Y) \tag{3}$$

The Variation of Information is defined as:

$$VI(X,Y) = H(X,Y) - MI(X;Y) \tag{4}$$

When normalized by $H(X,Y)$, $VI(X,Y)$ becomes a measure of distance between time series *x* and *y* [39]:

$$JD(X,Y) = 1 - MI(X;Y)/H(X,Y) \tag{5}$$

which is a metric because it satisfies the properties of symmetry, positivity, boundedness ($0 \leq JD(X,Y) \leq 1$) and triangle inequality [39]. In particular, the joint entropy of X and Y, $H(X,Y)$, represents the union of the subsets $H(X)$ and $H(Y)$. Considering Eq. 3, $MI(X;Y)$ represents the intersection between $H(X)$ and $H(Y)$ [40]. This means that $MI(X;Y)/H(X,Y)$ represents the ratio between the intersection and the union of $H(X)$ and $H(Y)$, which recalls the concept of "Jaccard similarity coefficient" (JS) [41], also known as "Intersection over Union", traditionally used for comparing the similarity and diversity between sample sets.

The Jaccard distance (JD) is defined as $JD = 1 - JS$ [41]. This is the reason why the term "Jaccard Distance" for Eq. 5 was adopted in the present work. The concept of Permutation Jaccard Distance (PJD) will be introduced by taking advantage of the properties of JD as well as from the advantages of projecting the time series into symbols.

2) Projecting the time series into symbols: Given two time series *x* and *y* with *N* samples, they can be mapped into a *m*-dimensional space, where *m* is the embedding dimension. Given the EEG window under analysis, starting from two given samples $x(t)$ and $y(t)$ and given a time-lag *L*, two *m*-dimensional vectors, X_t and Y_t can be constructed as follows:

$$X_t = [x(t), x(t+L), ..., x(t+(m-1)L)]^T \quad (6)$$

and

$$Y_t = [y(t), y(t+L), ..., y(t+(m-1)L)]^T \quad (7)$$

In order to make the explanation easily understandable, Figure 1 shows an example of how the method works with $m = 3$. In this case, X_t and Y_t are both vectors with three elements. The algorithm disregards the absolute values of X_t and Y_t as it only takes into account the relative amplitude of their elements: *low*, *medium*, *high*, which are indeed related to no specific value as they represent the possible levels when an embedding dimension $m = 3$ is selected. At each iteration, three samples ($m = 3$) of the time series are selected and compared to each other in order to assess which sample was relatively "low", "medium" or "high". If $m = 3$, 3 possible levels are considered and 6 possible ordinal sequences (*Motifs*) can be identified (Figure 1). Motifs are denoted with π_i, where $i = 1, ..., 6$. The algorithm checks which motif occurs in X_t (motif π_4, in the example shown in Figure 1) and which one occurs in Y_t (motif π_1, in the example shown in Figure 1). According to the example shown in Figure 1, in the first iteration, the algorithm increments the number of occurrences of motif π_4 in time series *x*, $\eta_X(\pi_4)$, and the number of occurrences of motif π_1 in time series *y*, $\eta_Y(\pi_1)$. The algorithm also increments the number of joint occurrence of the two motifs π_4 and π_1, $\eta_{X,Y}(\pi_4, \pi_1)$. Then the algorithm moves to the next sample $x(t+1)$, constructs two new vectors X_{t+1} and Y_{t+1} and reiterates the procedure. Eventually, the algorithm estimates the overall probability that a given motif π_i (with $i = 1, ..., 6$) occurs in *x* (Eq. 8) and in *y* (Eq. 9), by normalizing the number of occurrences η by the total number of iterations:

$$p_X(\pi_i) = \eta_X(\pi_i)/(N - (m-1)L) \quad (8)$$

$$p_Y(\pi_i) = \eta_Y(\pi_i)/(N - (m-1)L) \quad (9)$$

The probability that a couple of motifs occurs jointly is:

$$p_{X,Y}(\pi_i, \pi_j) = \eta_{X,Y}(\pi_i, \pi_j)/(N - (m-1)L) \quad (10)$$

As explained early in this Section, by discarding the absolute amplitude of the elements of vectors X_t and Y_t, by matching them with the predetermined Motifs, the procedure becomes amplitude independent. This characteristic is very useful when dealing with EEG, because a signal recorded through an electrode close to the reference one, will inherently have a lower amplitude, compared to an electrode located farther away.

3) Permutation Jaccard Distance Definition: Given a time series *x*, with *N* samples and embedding dimension *m*, the Permutation Entropy of *x* is defined as [38]:

$$PE(X) = -\sum_{i=1}^{m!} p_X(\pi_i) log(p_X(\pi_i)) \quad (11)$$

Given two time series *x* and *y*, with *N* samples and embedding dimension *m*, their Permutation Joint Entropy (PJE) can defined as:

$$PJE(X,Y) = -\sum_{i=1}^{m!}\sum_{j=1}^{m!} p_{X,Y}(\pi_i, \pi_j) log(p_{X,Y}(\pi_i, \pi_j)) \quad (12)$$

Their Permutation Mutual Information (PMI) is defined as:

$$PMI(X;Y) = PE(X) + PE(Y) - PJE(X,Y) \quad (13)$$

Their Permutation Variation of Information (PVI) is defined as:

$$PVI(X,Y) = PJE(X,Y) - PMI(X;Y) \quad (14)$$

Therefore, the Permutation Jaccard Distance between time series *x* and *y* can be defined as:

$$PJD(X,Y) = 1 - PMI(X;Y)/PJE(X,Y) \quad (15)$$

As the coupling strength between *x* and *y* increases, PJD is expected to decrease, because the two time series become more synchronized. In fact, as the coupling strength increases, PMI increases and the joint randomness (therefore PJE) decreases. As a consequence of JD definition (Section III-A1) PJD satisfies the properties of a metric and is bounded between 0 and 1. The advantages of using PJD are manifold: since PJD is a symbolic methodology, it is less sensitive to artifacts because it projects the EEG time series into a set of symbols (motifs). This is because the amplitude alteration in the EEG signals due to artifacts, would not alter the amplitude of the symbols. The robustness to noise of motifs-based descriptors, is discussed in detail in [42] and the independence from the amplitude of motifs-based descriptors is shown in [43]. Furthermore, PJD is a nonlinear metric that is able to capture the nonlinear dynamics in the EEG.

4) Wavelet Coherence as a descriptor to be compared to PJD: Conventional coherence measures the synchronization between two signals, observing only the spectral components and losing time information. Wavelet Coherence (WC) takes into account the temporal and frequency information, thus, it measures the coherence between two signals in the time-frequency space. As noted, WC outperforms conventional coherence in the analysis of AD EEGs [34],[35]. WC is based on the definition of Continuous Wavelet Transform (CWT), used in time-frequency analysis. The CWT of a time series *x* is defined as:

$$CWT(a,b) = \int_{-\infty}^{+\infty} x(t)\Psi^*_{a,b}(t)dt \quad (16)$$

where Ψ is the mother wavelet, *a* the scaling parameter and *b* the shifting parameter. In this study, the Morlet wavelet was

chosen as mother wavelet because it is reasonably localized in both time and frequency [44]. As each scale is inversely related to a specific frequency, the CWT is a time-frequency representation. The wavelet coherence WC_{xy} between two time series is defined as:

$$WC_{xy}(a,b) = \frac{|S(C_x^*(a,b)C_y(a,b))|^2}{S(|C_x(a,b)|^2)S(|C_y(a,b)|^2)} \quad (17)$$

where $C_x(a,b)$ and $C_y(a,b)$ are the wavelet functions of x and y at scales a and shift b; the asterisk sign * indicates the complex conjugate operator. S denotes a smoothing operator in time and scale. For the Morlet wavelet, a proper smoothing function S was provided by Torrence and Webster [45]. The numerator is the wavelet cross spectrum and it displays the areas with high common power between two signals. The wavelet coefficients of each $WC_{xy}(t,f)$ are then averaged, over the frequency, in order to come up with a single average value of wavelet coherence $WC_{xy}(t)$, for a given EEG epoch $EEG(t)$ (where $t = T0$ or $t = T1$), in every specific sub-band under consideration $f_1 - f_2$ (delta, theta, alpha or beta):

$$WC_{xy}(t) = \frac{1}{f_2 - f_1}\int_{f_1}^{f_2} WC_{xy}(t,f)df \quad (18)$$

where f_2 and f_1 are the upper and lower frequency bounds of the sub-band.

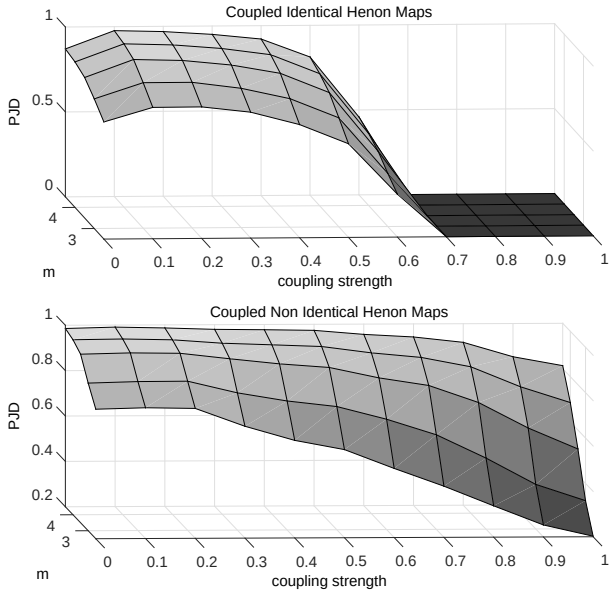

Fig. 2. PJD as a function of coupling strength c and embedding dimension m for the identical Henon Maps (top) and the non-identical Henon maps (bottom).

B. Estimating PJD from the EEG signals

The PJD between every possible pair of electrodes is computed, for every patient, in every sub-band, at T0 and T1. Given a sub-band sb (delta, theta, alpha or beta) and the corresponding EEG signals EEG_{sb} (extracted as described in Section II-C), given a generic window w under analysis, the $PJD^w_{x,y}(sb)$ is calculated between every pair of electrodes

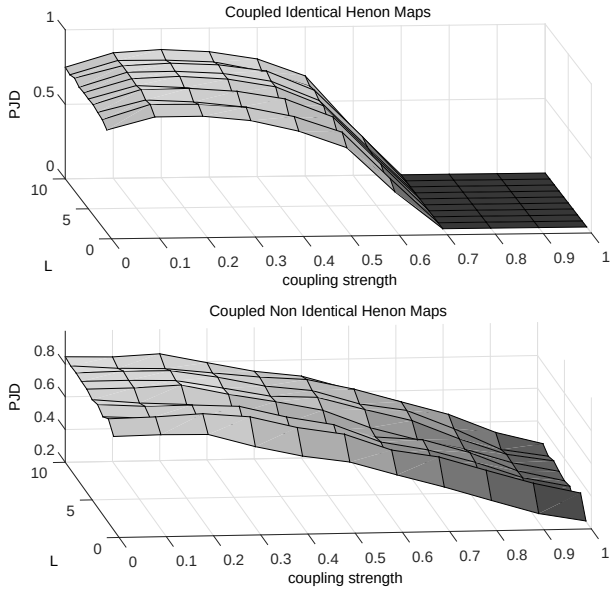

Fig. 3. PJD as a function of coupling strength c and lag L for the identical Henon Maps (top) and the non-identical Henon maps (bottom).

x and y. Such values are then averaged over the time (i.e. over the windows) in order to come up with an average PJD value $\overline{PJD}_{x,y}(sb)$ for every couple of electrodes x and y, in every sub-band sb. In the end of the analysis, $\overline{PJD}^{T0}_{x,y}(sb)$ and $\overline{PJD}^{T1}_{x,y}(sb)$ are computed for every pair of electrodes. The same procedure is carried out with WC and $\overline{WC}^{T0}_{x,y}(sb)$ and $\overline{WC}^{T1}_{x,y}(sb)$ are estimated.

C. PJD-based hierarchical clustering

The goal of the present paper is to use HC to cluster the electrodes according to their mutual $PJDs$ and to subsequently estimate the connectivity density of the network of electrodes. HC partitions a dataset of N samples into nC clusters. There are two kind of HC: agglomerative and divisive. The divisive HC assigns the entire dataset to a cluster and then iteratively splits it into groups until all groups are single clusters. This algorithm is computationally expensive and is not frequently used [46]. In this work, the agglomerative approach is adopted. It assigns an individual cluster to each data point and then, iteratively, merges the two most similar clusters. The procedure is repeated until all subsets belong to a single cluster [47]. Different agglomerative HC algorithms have been proposed in literature [48], in this paper, a complete linkage algorithm is employed. This method, called also *furthest neighbour*, defines the dissimilarity $D(A,B)$ between two clusters A and B through the maximum distance between two objects a and b where $a \in A$ and $b \in B$. The choice of the "distance" between objects is therefore crucial. In this paper, a distance metric sensitive to longitudinal changes in EEG coupling strength was necessary, this is the reason why PJD was adopted. Using PJD, the dissimilarity between objects (i. e. electrodes) x and y is defined as: $D^{T0}_{x,y}(sb) = \overline{PJD}^{T0}_{x,y}(sb)$,

Table I: Statistics (p value) of Descriptors' variation, from time T0 to time T1, in every sub-band

	WC					PJD				
Patient	**Overall**	**Delta**	**Theta**	**Alpha**	**Beta**	**Overall**	**Delta**	**Theta**	**Alpha**	**Beta**
Pt 03 (prodromal AD)	**1,98E-06**	**3,71E-12**	5,02E-01	7,66E-02	NS	**3,92E-03**	**5,95E-18**	**1,43E-05**	8,00E-01	NS
Pt 23	NS	NS	7,28E-02	NS	**1,00E-03**	4,48E-01	1,47E-01	**8,97E-03**	NS	2,67E-01
Pt 30	**3,40E-09**	**1,18E-08**	NS	8,46E-01	**3,58E-02**	NS	NS	2,49E-01	NS	6,51E-01
Pt 32 (prodromal AD)	**1,94E-02**	**5,88E-05**	**8,69E-05**	1,03E-01	**2,44E-03**	**4,61E-17**	**1,81E-10**	**3,10E-09**	**1,69E-09**	**1,90E-20**
Pt 41	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Pt 51 (prodromal AD)	NS	NS	**2,13E-13**	**1,62E-11**	7,79E-01	**1,01E-22**	**7,81E-16**	**6,50E-23**	**8,20E-20**	**1,16E-21**
Pt 57	1,36E-01	NS	NS	NS	NS	**3,69E-05**	6,16E-01	NS	**3,87E-06**	**4,82E-04**
Pt 71 (prodromal AD)	**6,86E-22**	**2,21E-34**	**8,59E-32**	**5,91E-04**	**2,11E-06**	**1,50E-32**	**3,46E-50**	**2,52E-34**	**1,82E-23**	**5,26E-35**
Pt 72	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Pt 127	NS	NS	**2,85E-02**	8,36E-01	NS	3,79E-01	NS	9,15E-01	4,88E-01	6,59E-01
Pt 128	NS	NS	NS	NS	NS	NS	NS	NS	6,58E-01	NS
Pt 132	NS	NS	**3,87E-02**	NS	NS	NS	NS	NS	NS	8,20E-01
Pt 137	NS	NS	NS	3,25E-01	NS	3,18E-01	1,20E-01	2,35E-01	NS	2,80E-01
Pt 143	NS	**1,94E-02**	NS	NS	NS	9,03E-01	5,97E-01	NS	1,73E-01	NS
Pt 144	1,81E-01	**2,84E-03**	**4,15E-08**	**3,31E-14**	1,43E-01	6,71E-01	1,24E-01	7,97E-01	7,08E-02	NS
Pt 146	NS	NS	NS	NS	NS	NS	NS	NS	**5,18E-02**	NS
Pt 147	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Pt 150	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Pt 152	**2,51E-03**	**2,54E-02**	NS	NS	NS	1,44E-01	NS	NS	**2,39E-02**	NS
Pt 153	NS	NS	NS	NS	**1,83E-03**	NS	NS	NS	NS	NS
Pt 155	1,02E-01	**2,65E-05**	NS	NS	2,57E-01	NS	NS	NS	NS	NS
Pt 156	NS	NS	**5,34E-04**	NS	NS	NS	NS	NS	NS	NS
Pt 157	8,86E-01	8,85E-01	NS	6,78E-01	6,35E-03	NS	NS	NS	NS	6,03E-01
Pt 162	NS	NS	NS	NS	NS	9,01E-01	7,46E-01	NS	NS	**2,55E-03**
Pt 173	NS	NS	NS	NS	NS	NS	NS	NS	NS	8,70E-01

when processing EEG-T0, and on $D^{T1}_{x,y}(sb) = \overline{PJD}^{T1}_{x,y}(sb)$ when processing EEG-T1. In this way, two dissimilarity matrices $\mathbf{D}^{T0}(sb)$ and $\mathbf{D}^{T1}(sb)$, are created, for every patient, in every sub-band. For comparison purpose, the same procedure is applied using WC instead of PJD. With regards to WC, the dissimilarity between electrodes x and y is defined as the complementary of wavelet coherence (that is bounded between 0 and 1): $D^{T0}_{x,y}(sb) = 1 - \overline{WC}^{T0}_{x,y}(sb)$, when analysing EEG-T0, and on $D^{T1}_{x,y}(sb) = 1 - \overline{WC}^{T1}_{x,y}(sb)$ when analysing EEG-T1. Therefore, two WC-based dissimilarity matrices are created, for every patient, in every sub-band.

D. Network density estimation

Given a patient, a time Ti of patient monitoring in the follow-up program (T0 or T1), a sub-band sb under analysis (delta, theta, alpha or beta) and a selected coupling strength descriptor (PJD or WC), a dissimilarity matrix $\mathbf{D}^{Ti}(sb)$ is constructed as described in Section III-C. HC is applied to the matrix $\mathbf{D}^{Ti}(sb)$. The output of *HC* can be represented through the *dendrogram*, which gives a view of the links among electrodes as a function of the *fusion level*. Given a fusion level *FL*, the electrodes linked at a level lower than *FL* will belong to the same cluster and will be considered *connected*. Therefore, for every dissimilarity matrix $\mathbf{D}^{Ti}(sb)$ and for every fusion level threshold *FL*, a set of clusters is determined. Subsequently, the number of active connections $AC^{Ti}_{FL}(sb)$ is calculated. Noting that the total number of possible connections between *n* nodes of a network is $n*(n-1)/2$, the network density ND can be defined as:

$$ND^{Ti}_{FL}(sb) = \frac{AC^{Ti}_{FL}(sb)}{n*(n-1)/2} \tag{19}$$

ND represents the ratio between the number of active connections and the number potential connections of a network. $ND = 0$ represents a totally disconnected network whereas $ND = 1$ represents a fully connected one. Appendix provides a pseudo-code of the whole procedure.

IV. Results

A. Sensitivity of PJD to changes in the coupling strength between simulated interacting dynamic systems

The aim of the present Section is to test PJD as a descriptor of the coupling strength between simulated time series, in order to assess if it is sensitive to changes in the coupling strength between interacting dynamic systems. PJD was tested on two unidirectionally coupled Henon Maps *X* and *Y*. Henon Maps have been extensively used in the literature to validate measures of coupling strength [49] [50]. The unidirectionally coupled Henon Maps are defined by the following equations:

$$\begin{cases} X : x_{n+1} = 1.4 - x_n^2 + b_x x_{n-1} \\ Y : y_{n+1} = 1.4 - [cx_n + (1-c)y_n]y_n + b_y y_{n-1} \end{cases} \tag{20}$$

where *n* is the step of iteration. System *X* drives system *Y* nonlinearly, according to the coupling strength c, which ranges from 0 to 1, where *c*=0 represents no coupling and *c*=1 represents complete coupling. In the analysis of identical systems, *bx*=*by*=0.3. In the analysis of nonidentical systems, *bx*=0.3 and *by*=0.1. Identical and nonidentical systems are discussed in detail in [51]. *X* and *Y* are initialized randomly. After initialization, *X* and *Y* are iteratively computed according to Eq. 20. Every 1000 steps, c is updated by $c = c + 0.01$. In this way, two time series with 110000 samples are generated. The two simulated time series are then processed within non-overlapping windows of 1000 samples. The PJD between *X*

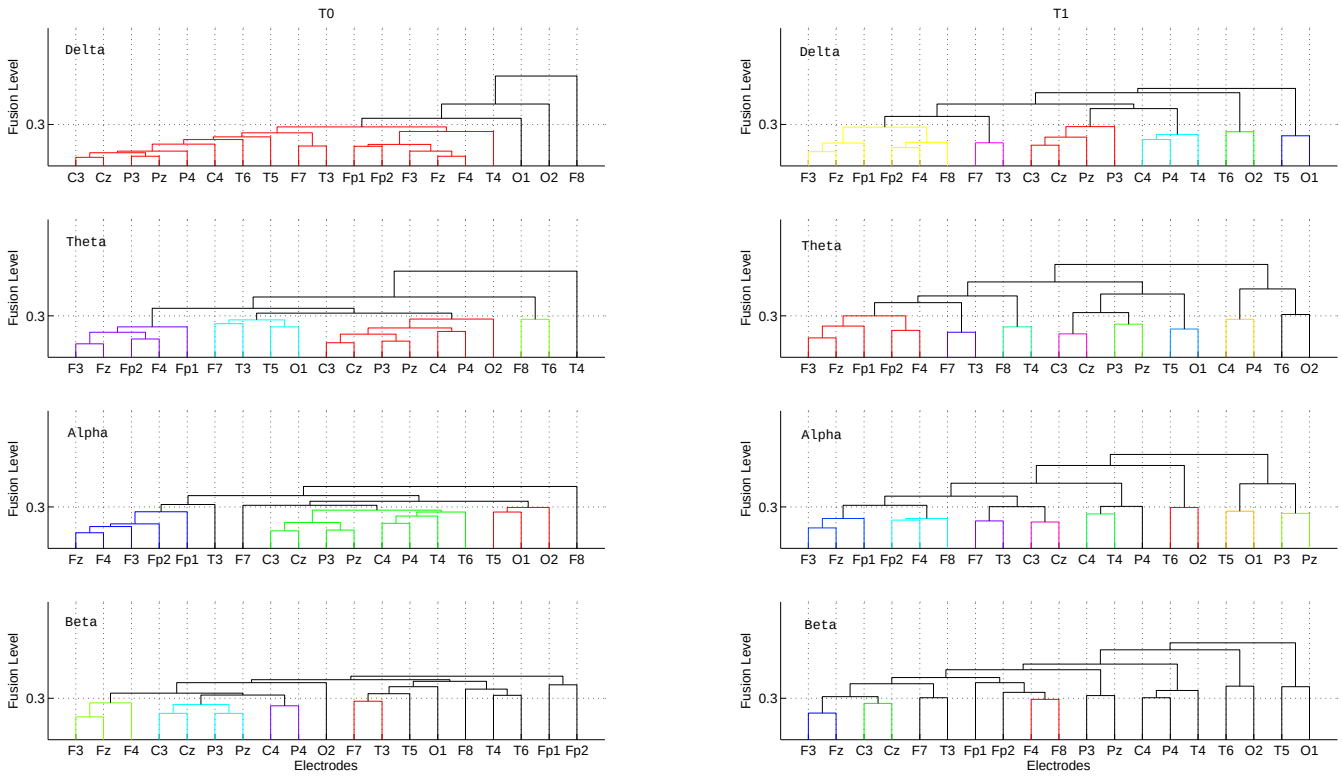

Fig. 4. Dendrograms of patient 51 (who was diagnosed MCI at T0 and AD at T1), in every EEG sub-band, at T0 and T1. The vertical axis of the dendrogram represents the distance or dissimilarity between clusters (fusion level). The horizontal axis represents the electrodes. Each joining (fusion) of two clusters is represented by a vertical line splitting into two vertical lines. The horizontal position of the split, shown by the short vertical bar, provides the distance (dissimilarity) between the two clusters. An example of how the electrodes are clustered when the fusion level thresold is set at 0.3 is shown.

and *Y* is calculated within each window, therefore, a sample of PJD is estimated every 1000 simulated samples of *X* and *Y*. In the end of the analysis, 110 PJD values are calculated, each one associated to a specific c value. In order to come up with a smooth PJD estimation, given a c_0 value and the corresponding $c_0 \div c_0 + 0.1$ range, the 10 PJD values estimated within that range are averaged. Figure 2 shows PJD as a function of both the coupling strength c and embedding dimension m, for identical (top sub-plot) and non-identical Henon maps (bottom sub-plot). PJD decreases as the coupling strength increases. This suggests that a decreasing PJD reflects an increasing coupling strength, as expected. With regards to identical systems, the behaviour looks similar for different m values. For $0 < c < 0.7$, PJD decreases monotonically as c increases, thus showing that PJD is sensitive even to small coupling variations. The critical threshold $c = 0.7$ corresponds to the condition of identical synchronization between the two systems. For $c > 0.7$, PJD reveals a strong synchronization between the two coupled systems. With regard to non-identical systems, as m increases, PJD becomes less sensitive to coupling strength variations. In fact, increasing m inherently reduces the joint probability of every pair of motifs and, therefore, inherently reduces the estimated synchronization between time series x and y. Finally, Figure 3 shows PJD as a function of the coupling strength c and the lag L. Once again, PJD decreases as the coupling strength increases. The behaviour looks the same for different L values. The issue of optimal parameter setting in the projection of EEG time series into symbols, specifically for AD/MCI patients, is discussed in detail in [42]. That paper reported that the setting $m = 3$ and $L = 1$ allowed for better capturing the *slowing* effect typical of MCI/AD EEGs. In conclusion, since $m = 3$ and $L = 1$ was shown to work fine on the EEGs of AD patients [42] and since that setting worked fine also in the theoretical experiments on Henon Maps shown in the present work, $m = 3$ and $L = 1$ was adopted herein.

B. Sensitivity of PJD to changes in the EEG coupling strength associated to the transition from MCI to AD

After having evaluated the ability of PJD to follow changes in the coupling strength between simulated coupled time series, the significance of the variation of PJD levels from T0 to T1 was assessed, for every patient, in every sub-band. The same analysis was carried out by using WC instead of PJD. PJD and WC were estimated at time T0 and T1, according to the procedure described in Section III-B. Once $\overline{PJD}^{Ti}_{x,y}(sb)$ is estimated, for every pair of electrodes, in every sub-band, we can define the two vectors $\overline{\mathbf{PJD}}^{T0}(sb)$ and $\overline{\mathbf{PJD}}^{T1}(sb)$,

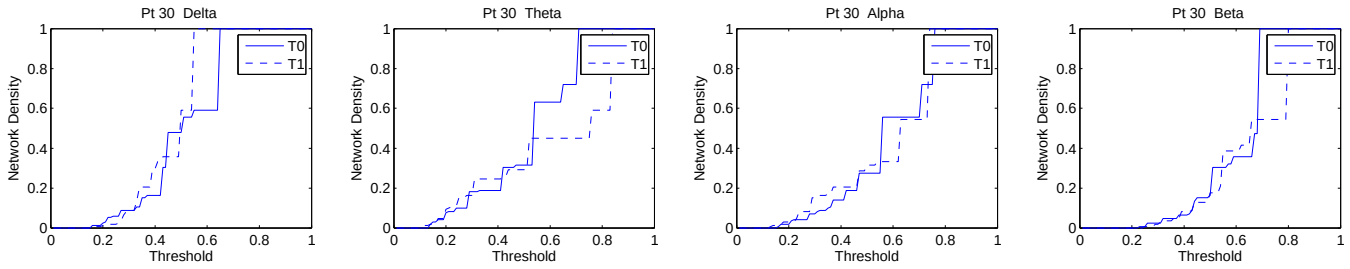

Fig. 5. ND evolution in subject Pt 30 (stable MCI). ND is depicted as a function of the fusion level threshold, in every sub-band, at T0 (continuous line) and at T1 (dashed line).

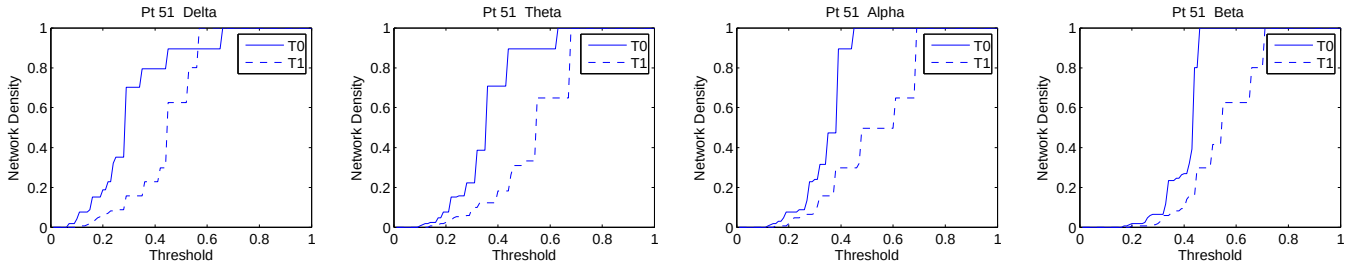

Fig. 6. ND evolution in subject Pt 51 (converted to AD at T1). ND is depicted as a function of the fusion level threshold, in every sub-band, at T0 (continuous line) and at T1 (dashed line).

where every element of the vector is associated with a pair of electrodes x and y. In order to evaluate the significance of the overall PJD variation in the transition from T0 to T1, the two populations (i. e. the two vectors $\overline{\mathbf{PJD}}^{T0}(sb)$ and $\overline{\mathbf{PJD}}^{T1}(sb)$) were statistically compared. The intra-subject statistical analysis was carried out by the Wilcoxon rank sum test [52], using the Matlab Statistics Toolbox (Mathworks). Given a patient and a sub-band, the null hypothesis that the two vectors under consideration, $\overline{\mathbf{PJD}}^{T0}(sb)$ and $\overline{\mathbf{PJD}}^{T1}(sb)$, are independent samples from identical continuous distributions with equal medians, was tested. The statistical significance was set at $p < 0.05$.

Table I reports the results. Table I reports the p values associated with decreased coupling strength (increased PJD or decreased Coherence). The significant values ($p < 0.05$) are highlighted in bold. The p values associated with increased coupling strength are denoted with "Not Significant" (NS). With regard to PJD analysis, the intra-subject variability study shows that PJD increased significantly ($p < 0.05$), between T0 and T1, in the overall range (0.5-32Hz), in delta and theta sub-bands specifically in the prodromal AD patients (Pt 03, 32, 51 and 71). In patients 32, 51 and 71, PJD significantly increased also in alpha and beta bands. The results show that PJD appears to be sensitive to the connectivity reduction that is expected to be induced by MCI progression towards AD, specifically in delta and theta bands, as well as in the overall range $0.5-32Hz$. As regards WC, three out of four converted patients (Pt 03, 32 and 71) exhibited a significant decrease of WC ($p < 0.05$), between T0 and T1, in the overall range ($0.5 - 32Hz$) and in delta band, however, such a behaviour was also observed in the stable patient Pt 152.

In conclusion, a peculiar behaviour of converted patients could be detected only using PJD: PJD significantly increased, between T0 and T1, in the overall range (0.5-32Hz) as well in delta and theta sub-bands. With regard to the intra-subject statistical analysis of PJD and WC variations from T0 to T1, PJD seems to outperform WC as a coupling strength descriptor, however, although PJD seems to outperform WC, the proposed approach (coupling strength descriptor + HC + network density estimation) either to PJD and WC in order to investigate in detail how they influenced the performance of the whole methodology. The results will be shown in the following sections.

C. Estimating changes in connectivity density through PJD-based HC

The present Section illustrates the results of the proposed method, coupling strength (PJD or WC) + HC + network density estimation. As described in those sections, considering a patient, a time Ti in the follow-up program (T0 or T1), a sub-band sb under analysis (delta, theta, alpha or beta) and a selected coupling strength descriptor (PJD or WC), the dissimilarity matrix $\mathbf{D}^{Ti}(sb)$ was estimated and HC was applied to it. The output of *HC* was represented as a dendrogram, in order to plot the connection among the electrodes as a function of the fusion level. Given a fusion level *FL*, the electrodes linked at a level lower than *FL* were considered connected. The explanatory dendrogram of a converted patient, Pt 51 is shown in Figure 4. In order to provide a view at-a-glance of how the connectivity changed from T0 to T1, an arbitrary fusion level was selected (0.3), for either time T0 and time T1, and the corresponding clusters are depicted in Figure 4. It is worth noting that the clusters significantly changed for Pt 51. For example, in delta band, at time T0, three single elements were observed and a large cluster with 16 elements

Table II: Variation of the Network Density (%), from time T0 to time T1, in every sub-band

	WC					PJD				
Patient	Overall	Delta	Theta	Alpha	Beta	Overall	Delta	Theta	Alpha	Beta
Pt 03 (prodromal AD)	-10,5427	-14,0529	-9,4731	-7,2903	16,3243	-11,3890	-16,8506	-17,7979	-2,8107	-1,0863
Pt 23	-9,8430	-6,1193	-7,0933	-8,2425	-15,4060	-6,7426	-3,2468	-8,3310	1,6503	-8,1050
Pt 30	-12,8046	-12,2395	-0,8787	0,8127	3,5099	-8,3897	8,8624	-17,8502	-0,1248	-11,7887
Pt 32 (prodromal AD)	-9,9223	-11,8551	-7,7799	-22,0394	-4,3528	-11,7276	-8,6111	-24,7501	-8,9659	-18,1170
Pt 41	20,0593	13,8176	4,8418	11,8544	3,8310	16,2234	20,7818	19,3970	22,9322	18,7450
Pt 51 (prodromal AD)	5,6631	8,6061	-15,4737	-9,9883	-2,8248	-25,5922	-18,8540	-26,7961	-26,9115	-23,7278
Pt 57	-6,1756	11,3768	-0,2753	10,0621	5,3042	-6,0188	4,2953	2,1169	-6,3780	-4,0384
Pt 71 (prodromal AD)	-26,9451	-28,6622	-30,5229	-11,4021	-12,2785	-37,9529	-25,7139	-35,0678	-37,6276	-38,5537
Pt 72	29,4773	1,8311	55,3970	45,4063	48,2014	20,9794	43,6527	47,4403	18,3776	16,6373
Pt 127	-1,9072	5,9160	-7,4871	0,5350	7,1542	1,8087	11,1583	7,1589	-2,9027	6,0705
Pt 128	-2,9874	14,3250	0,9943	-13,8910	-1,3079	4,0780	5,8894	-2,9138	-2,0185	9,6852
Pt 132	20,5724	18,7410	-5,4458	2,3999	7,1640	13,8482	23,2922	5,1045	13,3417	-1,7921
Pt 137	-9,5128	-8,6157	2,5696	-6,1000	0,0765	1,9772	0,3092	1,2644	2,3925	1,1458
Pt 143	-4,5972	-6,6329	-1,1076	-2,9539	3,7436	3,0132	1,6417	12,1239	-13,0592	5,0949
Pt 144	-4,5397	-8,6550	-12,2995	-17,4007	-8,6074	2,0564	4,6984	1,1681	-1,6820	6,4471
Pt 146	16,3487	8,9824	15,8411	9,1919	2,4468	-1,3928	7,3463	10,4766	-1,6517	3,9755
Pt 147	9,1365	-0,9617	10,6373	12,4769	-2,0669	4,0782	9,8692	3,6231	3,1022	3,8321
Pt 150	-1,5728	1,4722	7,6679	8,8323	5,5795	12,6309	29,0709	24,7136	11,9896	6,2864
Pt 152	-9,5168	-6,1373	1,7331	-0,4500	1,6214	-9,5383	2,5866	-7,8974	-9,9029	-0,2606
Pt 153	46,5915	48,4782	12,0332	-4,2626	4,1338	23,9232	29,8034	12,6058	20,7825	25,5532
Pt 155	-2,9383	-3,2941	3,3666	9,0457	11,0433	8,6336	2,4756	0,5857	5,8362	7,6411
Pt 156	4,5310	9,9463	-6,4259	0,5514	3,1690	2,4276	1,6760	-0,5822	7,5739	2,8292
Pt 157	-3,6335	0,8262	1,3870	0,8256	-5,7942	15,1749	6,0364	20,9418	10,6449	13,4512
Pt 162	3,1793	6,3553	3,2088	10,6541	0,2047	-3,2225	8,3359	14,0567	-1,1975	-3,8045
Pt 173	15,3454	4,2756	14,8657	19,9628	17,9548	7,9661	4,6500	25,8033	20,3664	-0,1755

was obtained, whereas, at time T1, 6 clusters were obtained, with a size ranging from 2 to 6 elements.

In order to quantify the results of such a visual evaluation, the network density ND was estimated, according to the method described in Section III-D. ND measures how many connections are active, as a function of the selected fusion level threshold. ND was estimated, according to Eq. 19, for different fusion level thresholds, ranging from 0 to 1, with a step of 0.01. Figures 5 and 6 show two explanatory representations of ND evolution, as a function of the fusion level threshold, in every sub-band, for patients Pt 30 (stable MCI) and Pt 51 (converted to AD), respectively. In Figures 5 and 6, ND was estimated using the PJD-based method. With regard to Pt 30 (Figure 5), it is observed that the two trends ($ND(T0)$ and $ND(T1)$) essentially overlap for $FL < 0.55$ in every sub-band, whereas they differ significantly in Pt 51, even for $FL < 0.55$.

This result indicates that ND significantly changed in patient Pt 51, in the transition from T0 to T1, whereas it remained stable in patient Pt 30. In order to quantify the results shown in Figures 5 and 6, the percent variation of the area below the two curves was calculated, $ND^{T0}(sb)$ and $ND^{T1}(sb)$, for every patient, in every sub-band sb:

$$\triangle ND(sb)\% = \frac{(ND^{T1}(sb) - ND^{T0}(sb))}{ND^{T0}(sb)} * 100 \qquad (21)$$

The results are reported in Table II, for both the PJD-based and the WC-based methods. With regard to the PJD-based method, patients converted to AD (prodromal AD) experienced a reduction of the ND area in every sub-band and have been highlighted in grey in Table II. With regard to WC, three out of four converted patients exhibited such a reduction, together with two non converted patients, thus making the WC-based method less sensitive and specific. Limited to the analysed dataset, by using the PJD-based method, the ND reduction in every sub-band, as well as in the overall range $0.5 - 32Hz$, seems to be associated with the conversion from MCI to AD. A behaviour peculiar of converted patient could not be detected by the WC-based method.

V. Discussion

This paper proposes a method for studying the connectivity of complex systems based on hierarchical clustering and on a novel metric called Permutation Jaccard Distance (PJD). The method is novel from either the theoretical and the application point of view. PJD is a novel measure of coupling strength between time series hereby introduced as the basic metric of a hierarchical clustering-based method for connectivity analysis. The choice of the metric is indeed decisive in hierarchical clustering. From the application point of view, innovation lies in the fact that the proposed method has been applied to the longitudinal study of EEG connectivity in MCI patients, which has never been studied in literature. The brain is indeed a complex system and the neurophysiological signals, like EEG, are the output of such a system. Learning from EEG data can help to understand the behaviour of the system that generated them, this is the reason why neural networks and learning systems have been extensively used in EEG analysis, for example in disease diagnosis [11], [23], [28], [34], [53], [54], [55], Brain Computer Interface applications [56], EEG-based studies of the brain condition [57], cognitive [58] or task analyses [59].

The present paper addressed the importance of developing novel EEG-based biomarkers, able to objectively quantify the changes caused by the progression of Alzheimer's disease, from the stage of MCI to the final dementia stage.

Since it is commonly believed that AD weakens the connectivity between the different cortical areas in the brain, the authors developed an EEG-based method for the indirect estimation of brain-electrical connectivity changes within a patient follow-up program.

The present work resulted from a translational research conducted at the IRCCS Neurolesi Center Bonino-Pulejo of Messina (Italy). Twentyfive MCI subjects were recruited and involved in a follow-up program that consisted in reevaluating the patients at time T1, three months after the first evaluation (time T0). The goal of the present study was to develop novel biomarkers for the objective quantification of the effects of disease's progression on the brain-electrical connectivity.

The proposed method considers the EEG electrodes as nodes of a network. The dissimilarity between the nodes is estimated as the complement of the coupling strength between the corresponding EEG signals. Such dissimilarities are then used as input to Hierarchical Clustering (HC) to cluster the electrodes and subsequently estimate the connectivity density of the network (ND). To our best knowledge, such a method has not been proposed in the literature yet. Therefore, the novelty of the paper does not rely in replacing WC with PJD but in introducing a novel method (EEG signals Coupling Strength Descriptor + Hirerachical Clustering + Network Density estimation) to indirectly estimate cortical connectivity changes induced by MCI degeneration towards AD.

In the early experiments, the coupling strength was estimated by Wavelet Coherence, due to the promising results in the literature on MCI/AD EEG analysis. Since the proposed method, when based on WC, did not exhibit a specific behaviour for converted patients, the authors decided to propose a new method for estimating the coupling strength between EEG signals, taking advantage of the projection of time series into symbols, according to Bandt and Pompe theory [38]. In the present research, the Permutation Jaccard Distance (PJD) was introduced and its ability to estimate the coupling strength between simulated dynamic interacting systems was tested. By using the PJD-based method, a ND reduction could be detected in every sub-band, as well as in the overall range $0.5 - 32Hz$, only in the converted patients.

The WC-based method did not show any peculiar behaviour for converted patients. According to the achieved results, the PJD-based network density estimation method seems to show potential as an EEG-marker in the follow-up analysis of MCI subjects. Both WC- and PJD-based methods will be tested on the upcoming extended dataset currently under construction at the IRCCS Neurolesi Center. Our future goal is to build a multistep longitudinal database where the patient is evaluated every 3 (up to 4) months.

The results presented in this study can be considered encouraging for the following reasons: 1) PJD is a new method for estimating the coupling strength between time series, particularly useful when it is not desirable to pre-normalize the available recordings; 2) This study could stimulate the clinical centers to perform further follow-up studies on MCI and test the proposed methodology on additional data; and 3) The proposed method can be in principle used on other neurological disorders affecting the cortical connectivity such as Parkinsons disease [60] [53], schizophrenia [54], [61], epilepsy [62] [63], ADHD [64] and autism [55].

Hopefully, the present work will encourage longitudinal studies on MCI patients, which are long overdue and strictly necessary to develop biomarkers that are able to objectively quantify the diseases progression towards AD. It could also encourage the clinical centers to cooperate and share their follow-up data, which would dramatically increase the size of the available dataset and speed-up the development of the aforementioned biomarkers.

VI. Conclusions

In the present paper, the importance of developing novel EEG-based biomarkers for the objective quantification of the effects of brain degeneration from the stage of Mild Cognitive Impairment (MCI) to the possible dementia-due-to-AD stage, was addressed. AD is defined in the literature as a disconnection disorder, because it weakens the connectivity between different cortical areas.

In this paper, the authors introduced a novel, EEG-based, method for the indirect estimation of brain connectivity in MCI patients and for the objective quantification of connectivity changes over the follow-up. The proposed method considers the EEG electrodes as nodes of a network. The dissimilarity between the nodes is estimated through PJD. Hierarchical Clustering (HC) is then applied in order to cluster the electrodes according to their mutual dissimilarity. The connectivity density ND of the network is eventually calculated for different fusion levels. The method was tested either on PJD and WC. First, the ability of PJD to estimate the coupling strength between simulated dynamic interacting systems was tested. Then, the method was applied for estimating the coupling strength between the EEG signals of the MCI patients involved in the study. The intra-subject variability analysis showed that the four patients who converted to AD exhibited a significantly ($p < 0.05$) increased PJD (reduced overall coupling strength), particularly in delta and theta bands. According to the proposed method, only the four converted patients showed a connectivity density reduction, in every sub-band under consideration (delta, theta, alpha, beta), as well as in the overall $0.5 - 32Hz$ range. The remaining 21 stable MCI patients did not exhibit such a behaviour. The reduced connectivity density reflects the expected reduced coupling strength between the brain areas that comes with MCI progression towards AD.

In the future, the database will be extended to include more patients and more steps in the follow-up program. The retrospective study of the resulting longitudinal database might help to develop EEG-based predictive biomarkers that can estimate the probability of conversion to AD.

Acknowledgment

This work was funded by the Italian Ministry of Health, project code: GR-2011-02351397.

Appendix

Pseudo-Code of the procedure.

PSEUDO-CODE

// EEG pre-processing: *s= #EEG signals analyzed*

1. **for** each EEG{s}, s=1,..,25 **do**
2. *eegfiltfft* ranged [0.5-32] Hz
3. *extraction* of $EEG_\delta, EEG_\theta, EEG_\alpha, EEG_\beta$, band
4. *downsampling* from 1024 Hz to 256 Hz
5. *partitioning* into *n* non-overlapping epochs of 5sec
6. **output**: EEG{s} filtered and downsampled

// Motifs' detection: *m = embedding dimension; L= time-lag; n= #epochs*

7. **for** each EEG epoch{w}, w=1,..n
8. **for** every pair of channels x and y
9. *mapping* each sample x(t) and y(t) into two *m* dimensional vectors, t=1,..., N

$$X_t = [x(t), x(t+L), ..., x(t+(m-1)L)]^T$$

$$Y_t = [y(t), y(t+L), ..., y(t+(m-1)L)]^T$$

10. **define** $\boldsymbol{\pi}_i$ motifs, i=1,...,m!

// Motifs' occurance rate estimation: *N=#samples; i=1,...,m!; j=1,...,m!*

11. **if** $\boldsymbol{\pi}_i$ occurs in Xt or Yt
12. **count** the number of occurrence $\boldsymbol{\eta_X}(\boldsymbol{\pi}_i)$ or $\boldsymbol{\eta_Y}(\boldsymbol{\pi}_i)$, respectivily
13. **if** $\boldsymbol{\pi}_i$ occurs in Xt and $\boldsymbol{\pi}_j$ occurs in Yt
14. **count** the number of joint occurrence $\boldsymbol{\eta_{X,Y}}(\boldsymbol{\pi}_i, \boldsymbol{\pi}_j)$
15. **end**
16. **normalize** $\boldsymbol{\eta_X}(\boldsymbol{\pi}_i)$, $\boldsymbol{\eta_Y}(\boldsymbol{\pi}_i)$, $\boldsymbol{\eta_{X,Y}}(\boldsymbol{\pi}_i, \boldsymbol{\pi}_j)$ by the number of iterations:

$$p_X(\pi_i) = \eta_X(\pi_i)/(N-(m-1)L)$$

$$p_Y(\pi_i) = \eta_Y(\pi_i)/(N-(m-1)L)$$

$$p_{X,Y}(\pi_i, \pi_j) = \eta_{X,Y}(\pi_i, \pi_j)/(N-(m-1)L)$$

17. **output:** $p_X(\pi_i)$, $p_Y(\pi_i)$, $p_{X,Y}(\pi_i)$

// PJD and WC calculation

18. **calculate** the Permutation Entropy PE(X) and PE(Y):

$$PE(X) = -\sum_{i=1}^{m!} p_X(\pi_i)\log\left(p_X(\pi_i)\right)$$

$$PE(Y) = -\sum_{i=1}^{m!} p_Y(\pi_i)\log\left(p_Y(\pi_i)\right)$$

19. **calculate** the Permutation Joint Entropy PJE(X,Y):

$$PJE(X,Y) = -\sum_{i=1}^{m!}\sum_{j=1}^{m!} p_{X,Y}(\pi_i,\pi_j)log(p_{X,Y}(\pi_i,\pi_j))$$

20. **calculate** the Permutation Mutual Information PMI (X,Y):

$$PMI(X;Y) = PE(X) + PE(Y) - PJE(X,Y)$$

21. **calculate** the Permutation Jaccard Distance PJD (X,Y):

$$PJD(X,Y) = 1 - PMI(X;Y)/PJE(X,Y)$$

22. **calculate** the Wavelet Coherence

$$WC_{xy}(a,b) = \frac{|S(C_x^*(a,b)C_y(a,b))|^2}{S(|C_x(a,b)|^2)S(|C_y(a,b)|^2)}$$

23. w+1; **end**
24. **calculate** the average $\overline{\boldsymbol{PJD}}$ and $\overline{\boldsymbol{WC}}$ over *n* epochs
25. **output:** $\overline{\boldsymbol{PJD}}$, $\overline{\boldsymbol{WC}}$

// Hierarchical Clustering: *T_i= T_0T_1 ; sb=sub-band*

26. **calculate** the dissimilarity matrix: $D_{x,y}^{T_i}(sb)$

26.1 PJD-based hierarchical clustering:

$$D_{x,y}^{T_i}(sb) = \overline{PJD}_{x,y}^{T_i}(sb)$$

26.2 WC-based hierarchical clustering:

$$D_{x,y}^{T_i}(sb) = 1 - \overline{WC}_{x,y}^{T_i}(sb)$$

27. **output:** dendrogram at T_0 T_1 for each sub-band

// Network density estimation: *AC=active connections; FL=fusion level; n=nodes; T_i= T_0T_1*

28. **while** fusion level FL < 1

$$ND_{FL}^{T_i}(sb) = \frac{AC_{FL}^{T_i}(sb)}{n*(n-1)/2}$$

29. **output:** ND at T_0 T_1 for each sub-band
30. s+1; **end**

REFERENCES

[1] M. Prince, A. Comas-Herrera, M. Knapp, M. Guerchet, and M. Karagiannidou. *World Alzheimer Report 2016*. Alzheimer's Disease International (ADI), London, 2016.

[2] M. Ganguli, H. H. Dodge, C. Shen, R. S. Pandav, and S. T. DeKosky. Alzheimer disease and mortality: A 15-year epidemiological study. *Arch Neurol*, 62(5):779–84, 2005.

[3] S. C. Waring, R. S. Doody, V. N. Pavlik, P. J. Massman, and W. Chan. Survival among patients with dementia from a large multi-ethnic population. *Alzheimer Dis Assoc Disord*, 19(4):178–8, 2005.

[4] R. Brookmeyer, M. M. Corrada, F. C. Curriero, and C. Kawas. Survival following a diagnosis of Alzheimer disease. *Arch Neurol*, 59(11):1764–7, 2002.

[5] E. B. Larson, M. F. Shadlen, L. Wang, W. C. McCormick, J. D. Bowen, L. Teri, and et al. Survival after initial diagnosis of Alzheimer disease. *Ann Intern Med*, 140(7):501–509, 2004.

[6] E. P. Helzner, N. Scarmeas, S. Cosentino, M. X. Tang, N. Schupf, and Stern. Survival in Alzheimer disease: A multiethnic, population-based study of incident cases. *Neurology*, 71(19):1489–1495, 2008.

[7] J. Xie, C. Brayne, and F. E. Matthews. Survival times in people with dementia: Analysis from a population based cohort study with 14-year follow-up. *BMJ*, 336(7638):259–262, 2008.

[8] R. A. Sperling, P. S. Aisen, L. A. Beckett, D. A. Bennett, S. Craft, A. M. Fagan, T. Iwatsubo, C. R. Jr Jack, J. Kaye, T. J. Montine, D. C. Park, E. M. Reiman, C. C. Rowe, E. Siemers, Y. Stern, K. Yaffe, M. C. Carrillo, B. Thies, M. Morrison-Bogorad, M. V. Wagster, and C. H. Phelps. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.*, 7(3):280–292, 2011.

[9] T. Varghese, R. Sheelakumari, J. S. James, and P. Mathuranath. A review of neuroimaging biomarkers of Alzheimer's disease. *Neurol Asia*, 18(3):239–248, 2013.

[10] L. K. McEvoy, C. Fennema-Notestine, J. C. Roddey, D. J. Hagler, D. Holland, D. S. Karow, and et al. Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology*, 251(1):195–205, 2009.

[11] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr. Alzheimer's disease and models of computation: Imaging, classification, and neural models. *Journal of Alzheimer's Disease*, 7(3):187–199, 2005.

[12] A. Alberdi, A. Aztiria, and A. Basarab. On the early diagnosis of Alzheimer's disease from multimodal signals: A survey. *Artif Intell Med.*, 71:1–29, 2016.

[13] G. Fiscon, E. Weitschek, G. Felici, P. Bertolazzi, S. De Salvo, P. Bramanti, and M. C. De Cola. Alzheimer's disease patients classification through EEG signals processing. In *Computational Intelligence and Data Mining (CIDM), 2014 IEEE Symposium on*, pages 105–112. IEEE, 2015.

[14] J. Dauwels, F. Vialatte, and A. Cichocki. Diagnosis of Alzheimer's disease from EEG signals: where are we standing? *Curr Alzheimer Res*, 7(6):487–505, 2010.

[15] S. De Salvo, A. Naro, L. Bonanno, M. Russo, N. Muscarà, P. Bramanti, and S. Marino. Assessment of nociceptive system in vegetative and minimally conscious state by using laser evoked potentials. *Brain injury*, 29(12):1467–1474, 2015.

[16] S. De Salvo, F. Caminiti, L. Bonanno, M. C. De Cola, F. Corallo, A. Caizzone, C. Rifici, P. Bramanti, and S. Marino. Neurophysiological assessment for evaluating residual cognition in vegetative and minimally conscious state patients: a pilot study. *Functional neurology*, 30(4):237, 2015.

[17] M. Ahmadlou, H. Adeli, and A. Adeli. New diagnostic EEG markers of the Alzheimer's disease using visibility graph. *Journal of neural transmission*, 117(9):1099–1109, 2010.

[18] M. Ahmadlou, H. Adeli, and A. Adeli. Fractality and a wavelet-chaos-methodology for EEG-based diagnosis of Alzheimer disease. *Alzheimer Disease & Associated Disorders*, 25(1):85–92, 2011.

[19] D. V. Moretti. Conversion of Mild Cognitive Impairment patients in Alzheimer's disease: prognostic value of alpha3/alpha2 electroencephalographic rhythms power ratio. *Alzheimers Res Ther.*, 7(80):14pp, 2015.

[20] C. A. Frantzidis, A. B. Vivas, A. Tsolaki, M. A. Klados, M. Tsolaki, and P. D. Bamidis. Functional disorganization of small-world brain networks in mild Alzheimer's Disease and amnestic Mild Cognitive Impairment: an EEG study using Relative Wavelet Entropy (RWE). *Front Aging Neurosci.*, 6(224).

[21] M. Ahmadlou, A. Adeli, R. Bajo, and H. Adeli. Complexity of functional connectivity networks in Mild Cognitive Impairment subjects during a working memory task. *Clinical Neurophysiology*, 125(4):694–702, 2014.

[22] N. Houmani, G. Dreyfus, and F. B. Vialatte. Epoch-based entropy for early screening of Alzheimer's Disease. *International journal of neural systems*, 25(08):1550032, 2015.

[23] A. Ortiz, J. Munilla, J. M. Górriz, and J. Ramirez. Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's Disease. *International Journal of Neural Systems*, 26(07):1650025, 2016.

[24] F. J. Martinez-Murcia, J. M Górriz, J. Ramírez, and A. Ortiz. A structural parametrization of the brain using hidden markov models-based paths in Alzheimer's Disease. *International Journal of Neural Systems*, 26(07):1650024, 2016.

[25] J. P Amezquita-Sanchez, A. Adeli, and H. Adeli. A new methodology for automated diagnosis of mild cognitive impairment (MCI) using magnetoencephalography (MEG). *Behavioural brain research*, 305:174–180, 2016.

[26] M. Ahmadlou and H. Adeli. Wavelet-synchronization methodology: a new approach for EEG-based diagnosis of ADHD. *Clinical EEG and Neuroscience*, 41(1):1–10, 2010.

[27] G. Mirzaei and H. Adeli. Resting state functional magnetic resonance imaging processing techniques in stroke studies. *Reviews in the Neurosciences*, 27(8):871–885, 2016.

[28] M. Buscema, E. Grossi, M. Capriotti, C. Babiloni, and P. Rossini. The IFAST model allows the prediction of conversion to Alzheimer disease in patients with mild cognitive impairment with high degree of accuracy. *Curr Alzheimer Res*, 7(2):173–187, 2010.

[29] F. C. Morabito, M. Campolo, D. Labate, G. Morabito, L. Bonanno, A. Bramanti, S. de Salvo, A. Marra, and P. Bramanti. A longitudinal EEG study of Alzheimer's disease progression based on a complex network approach. *Int J Neural Syst*, 25(2):1550005(1–18), 2015.

[30] R. Romero-Garcia, M. Atienza, and J. L. Cantero. Different scales of cortical organization are selectively targeted in the progression to Alzheimer's Disease. *International journal of neural systems*, 26(02):1650003, 2016.

[31] M. Buscema, F. Vernieri, G. Massini, F. Scrascia, M. Breda, PM. Rossini, and et al. An improved I-FAST system for the diagnosis of Alzheimer's disease from unprocessed electroencephalograms by using robust invariant features. *Artif Intell Med*, 64(1):59–74, 2015.

[32] N. Mammone, L. Bonanno, S. De Salvo, A. Bramanti, P. Bramanti, H. Adeli, C. Ieracitano, M. Campolo, and F. C. Morabito. Hierarchical clustering of the electroencephalogram spectral coherence to study the changes in brain connectivity in Alzheimer's disease. In *Evolutionary Computation (CEC), 2016 IEEE Congress on*, pages 1241–1248. IEEE, 2016.

[33] T. Koenig, L. Prichep, T. Dierks, D. Hubl, L. O. Wahlund, E. R. John, and V. Jelic. Decreased EEG synchronization in Alzheimer's disease and mild cognitive impairment. *Neurobiol Aging*, 26(2):165–171, 2005.

[34] Z. Sankari and H. Adeli. Probabilistic neural networks for diagnosis of Alzheimer's disease using conventional and Wavelet coherence. *J Neurosci Methods*, 197(1):165–170, 2011.

[35] Z. Sankari, H. Adeli, and A. Adeli. Wavelet coherence model for diagnosis of Alzheimer disease. *Clin EEG Neurosci.*, 43(4):268–278, 2012.

[36] American Psychiatric Association, editor. *Diagnostic and statistical manual of mental disorders (5th ed.)*. 2013.

[37] A. Delorme and S. Makeig. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including Independent Component Analysis. *Journal of Neuroscience Methods*, 134:9–21, 2004.

[38] C. Bandt and B. Pompe. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.*, 88 (17), 2002.

[39] A. Kraskov, H. Stogbauer, R. G. Andrzejak, and P. Grassberger. Hierarchical clustering based on mutual information. *arXiv:q-bio/0311039*, (7).

[40] T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[41] S. Mushfeq-Us-Saleheen and F. Raihana. An efficient k-means algorithm integrated with jaccard distance measure for document clustering. In *Internet, 2009. AH-ICI 2009. First Asian Himalayas International Conference on*, pages 1–6. IEEE, 2009.

[42] F. C. Morabito, D. Labate, F. La Foresta, A. Bramanti, G. Morabito, and I. Palamara. Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG. *Entropy*, 14(7):1186–1202, 2012.

[43] P. L. Vuong, A. S. Malik, and J. Bornot. Weighted-permutation entropy as complexity measure for electroencephalographic time series of different physiological states. In *Biomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on*, pages 979–984. IEEE, 2014.

[44] J. C. Moore, S. Jevrejeva, and A. Grinsted. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, (40), 2004.

[45] C. Torrence and G. P. Compo. A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, 79(1):61–78, 1998.

[46] I. Cattinelli, G. Valentini, E. Paulesu, and N. A. Borghese. A novel approach to the problem of non-uniqueness of the solution in hierarchical clustering. *IEEE Transactions on Neural networks and Learning Systems*, 24(7):1166–1173, 2013.

[47] X. Rui and D. Wunsch. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678, 2005.

[48] B. Everitt, S. Landau, and M. Leese. *Cluster Analysis*. 2001.

[49] R. Quian Quiroga, J. Arnhold, and P. Grassberger. Learning driver-response relationships from synchronization patterns. *Physical Review E*, 61(5):5142–5148, 2000.

[50] C. J. Stam and B. W. van Dijk. Synchronization likelihood: an unbiased measure of generalized sysnchronization in multivariate data sets. *Physica D*, 163:236–251, 2002.

[51] S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladares, and C. S. Zhou. *The synchronization of chaotic systems*. Elsevier, 2000.

[52] H. B. Mann and Whitney D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.*, 18(1):50–60, 1947.

[53] T. J Hirschauer, H. Adeli, and J. A. Buford. Computer-aided diagnosis of parkinson's disease using enhanced probabilistic neural network. *Journal of medical systems*, 39(11):179, 2015.

[54] D. Chyzhyk, M. Graña, D. Öngür, and A. K. Shinn. Discrimination of schizophrenia auditory hallucinators by machine learning of resting-state functional MRI. *International journal of neural systems*, 25(03):1550007, 2015.

[55] M. Ahmadlou, H. Adeli, and A. Adeli. Improved visibility graph fractality with application for the diagnosis of autism spectrum disorder. *Physica A: Statistical Mechanics and its Applications*, 391(20):4720–4726, 2012.

[56] H. Zeng and A. Song. Optimizing single-trial EEG classification by stationary matrix logistic regression in brain–computer interface. *IEEE transactions on neural networks and learning systems*, 27(11):2301–2313, 2016.

[57] C.-T. Lin, S.-F. Tsai, and L.-W. Ko. EEG-based learning system for online motion sickness level estimation in a dynamic vehicle environment. *IEEE transactions on neural networks and learning systems*, 24(10):1689–1700, 2013.

[58] R. Khosrowabadi, C. Quek, K. K. Ang, and A. Wahab. Ernn: A biologically inspired feedforward neural network to discriminate emotion from EEG signal. *IEEE transactions on neural networks and learning systems*, 25(3):609–620, 2014.

[59] J. del R Millan, J. Mouriño, M. Franzé, F. Cincotti, M. Varsta, J. Heikkonen, and F. Babiloni. A local neural classifier for the recognition of eeg patterns associated to mental tasks. *IEEE transactions on neural networks*, 13(3):678–686, 2002.

[60] F. Su, J. Wang, B. Deng, X.-L. Wei, Y.-Y. Chen, C. Liu, and H.-Y. Li. Adaptive control of Parkinson's state based on a nonlinear computational model with unknown parameters. *International journal of neural systems*, 25(01):1450030, 2015.

[61] S. A. Akar, S. Kara, F. Latifoğlu, and V. Bilgiç. Analysis of the complexity measures in the EEG of schizophrenia patients. *International journal of neural systems*, 26(02):1650008, 2016.

[62] M. Koppert, S. Kalitzin, D. Velis, F. Lopes Da Silva, and M. A. Viergever. Preventive and abortive strategies for stimulation based control of epilepsy: A computational model study. *International Journal of Neural Systems*, 26(08):1650028, 2016.

[63] J. R. Villar, P. Vergara, M. Menéndez, E. de la Cal, V. M. González, and J. Sedano. Generalized models for the classification of abnormal movements in daily life and its applicability to epilepsy convulsion recognition. *International Journal of Neural Systems*, 26(06):1650037, 2016.

[64] M. Ahmadlou and H. Adeli. Enhanced probabilistic neural network with local decision circles: A robust classifier. *Integrated Computer-Aided Engineering*, 17(3):197–210, 2010.

Responses to Associate Editor and reviewers

Comments to Author(s) by Associate Editor:
Associate Editor

One reviewer pointed out the scope issue, which is very important. The authors should clearly state the significant original contributions to the core of NNLS. The scope of TNNLS refers to the "theory, design, and applications of neural networks and related learning systems." Therefore, each submission should have a clear and strong connection to the core of NN and related learning systems.

++++++++++++++++++++

Individual Reviews:

Reviewer(s)' Comments to Author(s):

Reviewer: 1

Comments to the Author
There are no comments.

Reviewer: 2

Comments to the Author
In this paper, the authors propose a new hierarchical clustering approach based on Permutation Jaccard Distance applied to the detection of change in brain connectivity using EGG signal.

RESPONSE: We appreciate the reviewer has confirmed the newness of the approach.

The paper is clear and well written, with an adequate structure. The proposed approach is correctly explained and the experiments show the quality of the method.

RESPONSE: We appreciate the fact that paper is "well written, with an adequate structure. The proposed approach is correctly explained and the experiments show the quality of the method."

My only concern is that it is a very specific application to EEG data for Alzheimer Disease. It is an important subject, although I am not sure TNNLS is the most suitable Journal for this kind of research with medical application.

RESPONSE: The revised manuscript now clarifies the connection of the research to the scope of IEEE Transactions on Neural Networks and Learning Systems. In particular, the paper proposes a new method for studying the connectivity of complex systems based on hierarchical clustering and on a novel metric called Permutation Jaccard Distance (PJD). The method is novel from both the theoretical and the application point of view. PJD is a novel measure of coupling strength between time series hereby introduced as the basic metric for a hierarchical clustering-based

method of complex systems connectivity analysis. The choice of the metric is indeed decisive in hierarchical clustering. The proposed one, PJD, can be applied with great advantage whenever it is necessary to study the connectivity of a complex system whose elements evolve and interact with each other dynamically and whenever it is necessary to minimize the effects of possible artifacts. As such, the proposed method can find many other applications. We have added the following reference from the journal to establish connection to the content of the journal.

1. ***H. Zeng and A. Song. Optimizing single-trial EEG classification by stationary matrix logistic regression in brain–computer interface. IEEE transactions on neural networks and learning systems, 27(11):2301–2313, 2016.***
2. ***C.-T. Lin, S.-F. Tsai, and L.-W. Ko. EEG-based learning system for online motion sickness level estimation in a dynamic vehicle environment. IEEE transactions on neural networks and learning systems, 24(10):1689–1700, 2013.***
3. ***R. Khosrowabadi, C. Quek, K. K. Ang, and A. Wahab. Ernn: A biologically inspired feedforward neural network to discriminate emotion from EEG signal. IEEE transactions on neural networks and learning systems, 25(3):609–620, 2014.***
4. ***Jose del R Millan, Josep Mouri˜no, Marco Franz´e, Febo Cincotti, Markus Varsta, Jukka Heikkonen, and Fabio Babiloni. A local neural classifier for the recognition of eeg patterns associated to mental tasks. IEEE transactions on neural networks, 13(3):678–686, 2002.***

Also, I can' find the pseudocode in the appendix.

RESPONSE: The pseudo-code is on page 12 (boxed in yellow). Furthermore, the paper was polished (changes are highlighted in yellow)