Slowing and loss of complexity in Alzheimer's EEG: Two sides of the same coin?

Justin Dauwels, Srinivasan K, Ramasubba Reddy M, Toshimitsu Musha, François Vialatte, Charles Latchoumane, Jaeseung Jeong, and Andrzej Cichocki

***Abstract*—Medical studies have shown that EEG of Alzheimer's disease (AD) patients is "slower" (i.e., contains more low-frequency power) and is less complex compared to age-matched healthy subjects. The relation between those two phenomena has not yet been studied, and they are often silently assumed to be independent. In this paper, it is shown that both phenomena are strongly related. Strong correlation between slowing and loss of complexity is observed in two independent EEG data sets: (1) EEG of pre-dementia patients (a.k.a. Mild Cognitive Impairment; MCI) and control subjects; (2) EEG of mild AD patients and control subjects. The two data sets are from different patients, different hospitals, and obtained through different recording systems.**

The paper also investigates the potential of EEG slowing and loss of EEG complexity as indicators of AD onset. In particular, relative power and complexity measures are used as features to classify the MCI and MiAD patients vs. age-matched control subjects; linear and quadratic discriminant analysis and support vector machines are applied as classifiers. When combined with two synchrony measures (Granger causality and stochastic event synchrony), classification rates of 83% (MCI) and 98% (MiAD) are obtained. By including the compression ratios as features, slightly better classification rates are obtained than with relative power and synchrony measures alone.

***Index Terms*—Alzheimer's disease (AD), mild cognitive impairment (MCI), electroencephalogram (EEG), compression ratio, relative power, Granger causality, stochastic event synchrony**

I. Introduction

Alzheimer's disease (AD) is the most common form of dementia; it is the sixth leading cause of death in the United States. More than 10% of Americans over age 65 suffer from AD, and it is predicted that the prevalence of AD will triple within next 50 years [1]–[3]. Currently, no known medicine exists for curing AD, but a number of medications are believed to delay the symptoms and the causes of the disease.

Corresponding author: Justin Dauwels, e-mail: jdauwels@ntu.edu.sg

J. Dauwels is with the School of Electrical & Electronic Engineering (EEE), Nanyang Technological university (NTU), 50 Nanyang Avenue, Singapore 639798.

K. Srinivasan is with Department of Applied Mechanics, Indian Institute of Technology Madras, Chennai-600 036, India. He is also affiliated with the School of Electrical & Electronic Engineering (EEE), Nanyang Technological university (NTU), 50 Nanyang Avenue, Singapore 639798. e-mail: srinivasan.sivam@gmail.com

M. Ramasubba Reddy is with Department of Applied Mechanics, Indian Institute of Technology Madras, Chennai-600 036, India.

Toshimitsu Musha is with Brain Functions Laboratory, Inc., Yokohama, Japan.

François Vialatte is with ESPCI ParisTech, Laboratoire SIGMA, France.

Charles Latchoumane is with Korea Institute of Science and Technology (KIST), 39-1 Hawolgok-dong, Seongbuk-gu, Seoul 136-791, South Korea.

Jaeseung Jeong is with KAIST, Department of Bio and Brain Engineering, Daejeon, 305-701, South Korea.

Andrzej Cichocki is with the RIKEN Brain Science Institute, Wako-shi, Saitama, 351-0106, Japan.

The progression of AD can be categorized into three different stages: mild, moderate, and severe AD; there is also a stage known as Mild Cognitive Impairment (MCI) or predementia, that characterizes a population of elderly subjects who are not compromised in their daily living, but have a subclinical and isolated cognitive deficit and are potentially at risk of developing Alzheimer's disease [4, 5]. Around 6% to 25% of people affected by MCI progress to AD. MCI may develop into Mild AD and next Moderate AD; in those stages, cognitive deficits become more severe, and the patients become more dependent on caregivers. In the final stage known as severe AD, the personality of patients may change dramatically, and patients are entirely dependent on caregivers [6].

Diagnosing MCI and Mild Alzheimer's disease is hard, because most symptoms are often dismissed as normal consequences of aging. To diagnose MCI or Mild AD, extensive testing is required, to eliminate all possible alternative causes. Tests include psychological evaluations such as Mini Mental State Examination (MMSE), blood tests, spinal fluid, neurological examination and imaging techniques [7, 8].

Several research groups have investigated the potential of electroencephalograms (EEG) for diagnosing AD in recent years. Since EEG recording systems are nowadays relatively inexpensive and mobile, EEG may potentially be used in the future as a tool to screen a large population for the risk of AD, before proceeding to any expensive imaging or invasive procedures. To date, however, EEG does not have sufficiently high specificity and sensitivity to assume the role of reliable and reproducible method of screening AD.

In recent years, several studies have shown that AD has at least three major effects on EEG (see [9, 10] for an in-depth review): slowing, reduced complexity, and loss of synchrony. However, these effects tend to vary across patients, which makes diagnosis of AD a difficult task. Many recent studies are devoted to improving the sensitivity of EEG for diagnosing AD. We refer to [10] for a detailed review on various EEG statistics that have been used in this context.

In this paper, we investigate the relation between slowing and reduced complexity in AD EEG. Those two phenomena are often silently assumed to be independent. However, since low-frequency signals are more regular than signals with high-frequency components, one would expect that slowing and reduced complexity in AD EEG are strongly related to each other. Nevertheless no study so far has analyzed the relation between both phenomena on a statistical basis though.

In order to investigate the slowing effect in AD EEG, we compute relative power in the standard EEG frequency bands (see Table I). When relative power is larger than usual in

low-frequency bands (delta and/or theta), it is said that the EEG is "slower", and that "EEG slowing" occurs. We quantify the irregularity of EEG by a standard measure, i.e., Lempel-Ziv complexity (see Table I). We also apply several lossless-compression algorithms to the EEG, and we use the resulting compression ratios (reduction in data size after compression) as regularity measures (see Table I). Regular signals are more compressible than irregular signals, and therefore, they result in larger compression ratios; as a consequence, compression ratios are a measure of the regularity of signals.

We consider two EEG data sets: (1) EEG of pre-dementia patients (a.k.a. Mild Cognitive Impairment; MCI) and control subjects; (2) EEG of mild AD patients and control subjects. The two data sets are from different patients, different hospitals, and obtained through different recording systems.

We will show that the theta-band (θ) relative power is significantly larger in both groups of patients compared to age-matched control subjects, and that the lossless compression ratios are significantly larger in MiAD patients than in the age-matched control subjects; however, no significant perturbation of Lempel-Ziv complexity and the lossless-compression ratios is observed for the MCI patients. Interestingly, our numerical analysis will reveal strong correlation between theta relative power on the one hand and Lempel-Ziv complexity and the lossless-compression ratios on the other hand; in other words, the effects of slowing and loss of complexity in AD EEG seem to be significantly coupled, at least in the two EEG data sets at hand.

The paper is structured as follows: In Section II we explain how relative power of EEG may be computed. In Section III, we describe the Lempel-Ziv complexity measure and the lossless-compression schemes used in this study. In Section IV we discuss the two EEG datasets, and in Section V we present our results. We provide concluding remarks and topics of future research in Section VI.

Readers who are not interested in the technical and mathematical details of our data analysis may skip Sections II and III, and may directly proceed to Section IV.

II. Relative Power of EEG

The spectrum of EEG is helpful in describing and understanding brain activity. The EEG spectrum is commonly divided in specific frequency bands: 0.5–4Hz (delta), 4–8Hz (theta), 8–10Hz (alpha 1), 10–12Hz (alpha 2), 12–30Hz (beta), and 30–100Hz (gamma) [11]. Neurological diseases, including MCI and AD, often affect the EEG spectrum. Many studies have shown that MCI and AD cause EEG signals to "slow down" (see [10] and references there in), corresponding to an increase of power in low-frequency bands (delta and theta band, 0.5–8Hz) and a decrease of power in higher-frequency bands (alpha and beta, 8–30Hz).

The EEG spectrum can be computed by means of the Discrete Fourier Transform (DFT) of the EEG [10]. The DFT $X(f_n)$ of the sequence x is usually computed at multiples f_n of $f_T = 1/T$, where T refers to the length of the signal. For computational convenience, the length of the sequence x is often extended to the nearest power of two by zero-padding. As in [10], let us consider an example with T = 20s and the sampling frequency 200Hz, then DFT is computed at 0Hz, 0.05Hz, 0.1Hz, ..., 200Hz. The Nyquist theorem states that only one half the spectrum is of interest, while the other half is the mirror image of the first half; hence for the above example, it is enough to retain the DFT values at 0Hz, 0.05Hz, 0.1Hz, ..., 100Hz. The DFT values $X(f_n)$ are complex , and we are mostly interested in its absolute magnitude $|X(f)|$. The relative power of a frequency band is computed by summing $|X(f_n)|$ over the frequencies f_n in that band, and next by dividing the resulting intra-band sum by the sum of $|X(f)|$ over all DFT frequencies f_n.

III. Complexity Measures

A variety of complexity measures has been used to quantify EEG complexity, stemming from several areas ranging from statistical physics to information theory. We refer to [10] for more information. Earlier studies have reported that the EEG of MCI and AD patients seems to be more regular (i.e., less complex) than in age-matched control subjects. It is conjectured that due to MCI/AD induced loss of neurons and perturbed anatomical and/or functional coupling, fewer neurons interact with each other, and the neural activity patterns and dynamics become simpler and more predictable.

As mentioned earlier, we quantify EEG complexity by a standard measure, i.e., Lempel-Ziv complexity. In addition, we use lossless-compression ratios as regularity measures. In the following, we describe Lempel-Ziv complexity, next we elaborate on lossless compression and its use as measure for regularity.

A. Lempel-Ziv (LZ) complexity

The Lempel-Ziv complexity measure (LZ complexity) computes the number of different patterns present in a sequence of symbols [12]; if the number of different patterns is large, the sequence is complex and hence difficult to compress. LZ complexity is obtained by dividing the number of different patterns by the maximum complexity of a sequence of length N. For more details we refer to [13].

To compute LZ complexity, the time series is first reduced to a symbol list. For the sake of simplicity, we convert the EEG signals into binary sequences $s = s(1), s(2), \ldots, s(N)$, where $s(i) = 0$ if $x(i) < T_d$ and $s(i) = 1$ otherwise; that approach was also followed in [13]. The threshold T_d is chosen as the median of x, since the latter is robust to outliers.

B. Lossless-compression algorithms

In this section, we briefly explain the lossless-compression algorithms applied in this study (see Fig. 2); we will consider three different algorithms, which were all proposed in [14, 15]. The aim of compression is to reduce the size of a given data source (e.g., EEG data). In lossless compression (e.g., ZIP compression algorithm), no information in the original data source is lost after compression, in contrast to lossy compression, where the original can only approximately be constructed after compression (e.g., JPEG compression algorithm for images).

Measure	Description	References
Relative power	Power within specific EEG frequency band normalized by total power Frequency bands: 0.5–4Hz (delta), 4–8Hz (theta), 8–10Hz (alpha 1), 10–12Hz (alpha 2), 12–30Hz (beta)	[11]
Lempel-Ziv complexity	Number of different patterns present in a sequence of symbols (complexity measure)	[12]
Lossless-compression ratio	Reduction of the size of EEG data after lossless compression (regularity measure) Compression algorithms considered here: 1-D SPIHT, 2-D SPIHT, and 2-D SPIHT followed by arithmetic coding	[14, 15]

TABLE I
OVERVIEW OF STATISTICAL MEASURES: RELATIVE POWER, LEMPEL-ZIV COMPLEXITY, AND LOSSLESS-COMPRESSION RATIO.

Biomedical signals such as EEG often have a *decaying* spectrum: The energy is mostly concentrated at low frequencies, and it decays with increasing frequency. Therefore, the spectral components are close to zero at high frequencies; the same holds for coefficients in the time-frequency representation corresponding to high frequencies. To exploit this phenomenon, compression algorithms often subject the given data source to a transform (e.g., time-frequency transform), which results in an alternative representation of the data. The three algorithms used in this study all map the signals into an other domain, i.e., time-frequency domain; the sparseness of the time-frequency representation is then exploited to form a compact code. We now briefly outline the compression process (see Fig. 2). First the EEG signal is *preprocessed*, i.e., the DC component (average value of EEG signal) is removed by applying backward difference; the resulting zero-mean signal is then arranged as a 1D vector (see Fig. 2(a); Algorithm A) or 2D matrix (see Fig. 2(b) and 2(c); Algorithms B and C). The resulting structure is then decomposed into different frequency bands via *integer lifting wavelet transform*, which maps the signals to integers on several time scales; at last, a *set partitioning* coding scheme converts the (integer) wavelet coefficients into a compact representation. In the following sections we describe those different steps in more detail, and then we elaborate on the differences between the three algorithms (Algorithms A, B, and C).

1) Backward difference: First the EEG signal x is *preprocessed*, i.e., the DC component (average value of EEG signal) is removed; this is performed via *backward difference* operation:

$$\tilde{x}(n) = x(n) - x(n-1), \tag{1}$$

where $\tilde{x}(n)$ is the signal obtained by applying the backward difference. Next the EEG is arranged as a vector of size N (1D compression) or as a matrix of size $N \times N$ (2D compression); the latter matrix is filled starting at the top left-hand side, from left to right on the odd rows, and from right to left on the even rows. In matrices, each entry has 8 nearest neighbors (except for entries in the first/last row/column), compared to two nearest neighbors in vectors (except for first and last entries). In the present application, neighboring entries are adjacent EEG samples, which are highly correlated [14]. By leveraging on the additional nearest neighbors (8 instead of 2), 2D compression often yields better compression ratios than 1D compression [14].

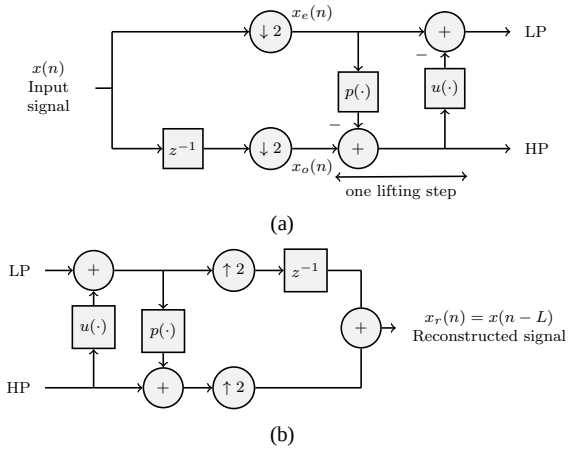


Fig. 1. Wavelet transform realization via lifting scheme (a) Forward transformation, (b) Inverse transformation. The boxes labeled by z^{-1} stand for delays (over one sample). The boxes $\downarrow 2$ and $\uparrow 2$ represent downsampling and upsampling by a factor of two respectively; in the latter a zero is inserted after every sample, whereas in the former, every second sample is removed. The lifting scheme repeats two primitive steps: prediction p and update u.

2) Lifting Wavelet Transform: A wavelet transform decomposes a given signal into different frequency bands; it allows to represent the signal in multiple resolutions (coarse to fine) [16]. Wavelets are usually realized by a set of filters, operating in parallel ("filter banks"). An alternative method of realizing wavelets is a *lifting scheme* [17], which consists of a cascade of simple filters; it may be viewed as the factorization of a filter bank into elementary filters. One such simple filter is depicted in Fig. 1(a)) and Fig. 1(b)). The former shows the forward lifting transformation; the signal x is first split into *odd* and *even* phases x_o and x_e respectively, containing the odd and even samples respectively of input signal x. The odd and even phases contain adjacent samples; in natural signals such as EEG, adjacent samples are highly correlated. Therefore, the odd phase may be predicted from the even phase (and vice versa). By subtracting the prediction $\hat{x}_o = p(x_e)$ from the odd phase, we are left with a high-frequency residue signal (HF) of the odd phase. The latter is used in another lifting step, to predict the even phase x_e ("update" u); the resulting prediction is subtracted from the even phase x_e, which leaves the low-frequency component (LF) of the even phase x_e; this also ensures the complete frequency separation

between a LF and HF component. The forward transform of Fig. 1(a) is easily invertible by reversing the steps and flipping the signs (see Fig. 1(b)). We implement the prediction p and update u by means of the widely used bi-orthogonal 5/3 filter [18], as we did in our previous study on EEG compression [15].

In a lifting scheme, the pair of lifting steps, i.e., prediction p and update u, is repeated several times, leading to multi-scale representation of the input signal x ("wavelet"); the nature and number of lifting steps p and u depends on the type of wavelets [17]. Integer wavelet transforms can easily be realized by systematic rounding and truncation of the intermediate results, i.e., output of p and u [19].

The lifting wavelet transform provides a sparse, multi-resolution representation, that is well suited for effective compression (for example, by means of SPIHT, to be explained in next section); *integer* lifting in particular enables convenient and simple implementations of lossless compression.

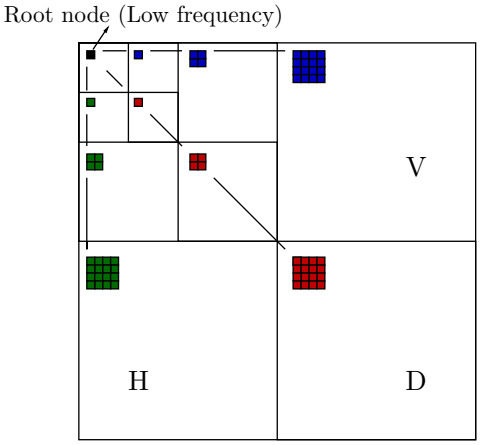


Fig. 3. Wavelet decomposition of the 2-D matrix and associated tree-based set originating from the low frequency band. The root node (black) branches towards horizontal, vertical and diagonal higher-frequency bands (H,V,D).

3) Set partitioning in hierarchical trees algorithm (SPIHT): As the last step in the process, the wavelet transformed signals are compressed. We use a widely known wavelet-based compression scheme, i.e., *Set Partitioning in Hierarchical Trees* (SPIHT) [20]. The underlying idea is *set partitioning*: Sets of samples are recursively split, guided by a series of threshold tests. This approach is particularly well-suited for wavelet transformed data, as wavelet coefficients are naturally clustered. In SPIHT the sample sets are non-overlapping, and they are organized by means of a tree: Each set is rooted in a subset of low-frequency coefficients, and branches successively to subsets of high-frequency coefficients in the same orientation (see Fig. 3). The search for coefficients associated with a particular threshold usually starts at the root node and proceeds successively towards the leaves of the tree, until all significant coefficients are listed. Such tree-based search, starting at coarse resolution at the root and ending with the finest resolution at the leaves, results in output signals of increasing quality and resolution.

The integer wavelet transform, in conjunction with SPIHT, yields a quality and resolution scalable bitstream: The quality and resolution of the signal improves as bitstream progresses. This is a very desirable property for real-time applications. Moreover, the output bitstream is embedded: The bitstream can be truncated at any point to approximately reconstruct the signal. When the bitstream is fully decoded, we obtain a lossless representation.

Though this coding scheme is specifically developed for images, it can be applied to all data sources with decaying spectrum [21].

4) Three SPIHT compression algorithms: The three compression algorithms are depicted in Fig. 2: (1) 1-D SPIHT compression, where the EEG is arranged as a vector (Fig. 2(a)), (2) 2-D SPIHT compression, where the EEG is arranged as a matrix (Fig. 2(b)), and (3) 2-D SPIHT compression (at optimal rate R_o), followed by arithmetic coding for the residuals (Fig. 2(c)). In the 1-D SPIHT compression scheme, backward differentiated EEG is subjected to integer wavelet transformation followed by SPIHT coding. The 2-D SPIHT compression scheme arranges the EEG as a matrix instead of a vector. In the two-stage 2-D SPIHT compression scheme, arithmetic coding is applied to the residuals of 2-D SPIHT compression: First SPIHT encodes the wavelet coefficients till the source loses its memory and behave as independent and identically distributed (corresponding to the optimal bit-rate R_o); next the residuals are encoded by means of single-context arithmetic coding.

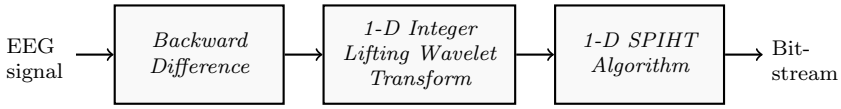
IV. EEG Datasets

A. *Dataset 1: MCI* vs. *Control*

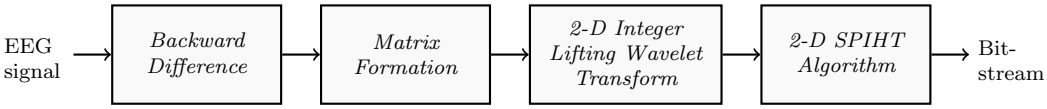
The first EEG data set has been analyzed in previous studies concerning early diagnosis of AD [22]–[26].

Ag/AgCl electrodes (disks of diameter 8mm) were placed on 21 sites according to 10–20 international system, with the reference electrode on the right ear-lobe. EEG was recorded with Biotop 6R12 (NEC San-ei, Tokyo, Japan) at a sampling rate of 200Hz, with analog bandpass filtering in the frequency range 0.5-250Hz and online digital bandpass filtering between 4 and 30Hz, using a third-order Butterworth filter. We used a common reference for the data analysis (right ear-lobe), and did not consider other reference schemes (e.g., average or bipolar references).

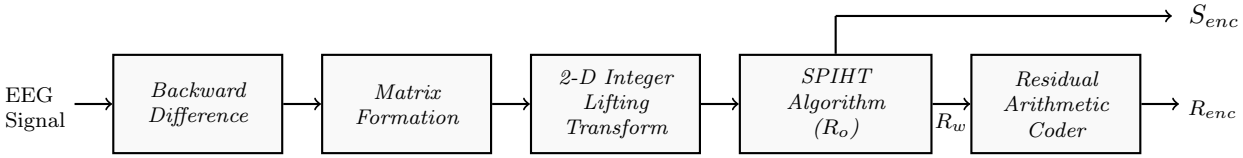
The subjects comprise two study groups. The first consists of 25 patients who had complained of memory problems. These subjects were diagnosed as suffering from mild cognitive impairment (MCI) when the EEG recordings were carried out. Later on, they all developed mild AD, which was verified through autopsy. The criteria for inclusion into the MCI group were a mini mental state exam (MMSE) score = 24, though the average score in the MCI group was 26 (SD of 1.8). The other group is a control set consisting of 56 age-matched, healthy subjects who had no memory or other cognitive impairments. The average MMSE of this control group is 28.5 (SD of 1.6). The ages of the two groups are 71.9 ± 10.2 and 71.7 ± 8.3, respectively. Finally, it should be noted that the MMSE scores of the MCI subjects studied here are quite high compared to a number of other studies. For example, in [27] the inclusion criterion was MMSE = 20, with a mean value of 23.7, while



(a) Algorithm A: EEG compression using 1-D SPIHT.



(b) Algorithm B: EEG compression using 2-D SPIHT.



(c) Algorithm C: EEG compression using 2-D SPIHT (at optimal rate R_o), followed by arithmetic coding for the residuals.

Fig. 2. Lossless EEG compression algorithms apply wavelet transforms followed by Set Partitioning in Hierarchical Trees (SPIHT).

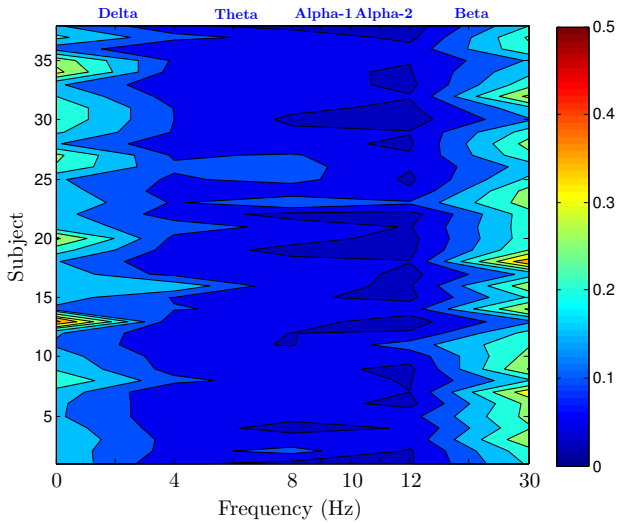
in [28], the criterion was MMSE = 22 (the mean value was not provided); thus, the disparity in cognitive ability between the MCI and control subjects is comparatively small, making the classification task relatively difficult.

B. Dataset 2: Mild AD vs. *Control*

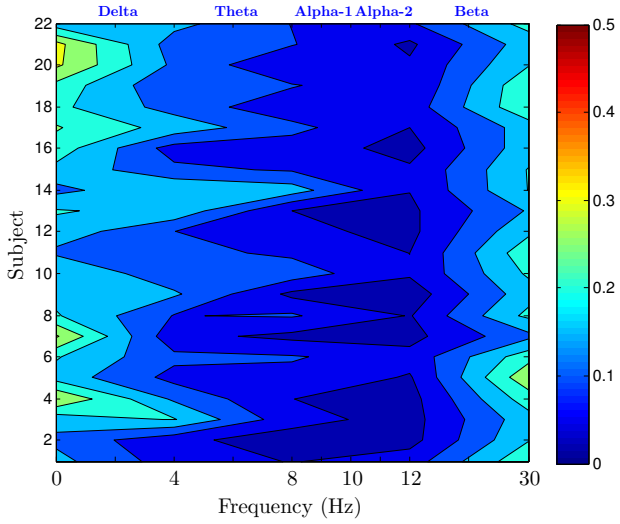
The second EEG data set also has been analyzed in previous studies [29, 30]; these data were obtained using a strict protocol from Derriford Hospital, Plymouth, U.K., and had been collected using normal hospital practices [30]. EEGs were recorded during a resting period with various states: awake, drowsy, alert and resting states with eyes closed and open. All recording sessions and experiments proceeded after obtaining the informed consent of the subjects or the caregivers and were approved by local institutional ethics committees. EEG dataset is composed of 24 healthy control subjects (age: 69.4±11.5 years old; 10 males) and 17 patients with mild AD (age: 77.6±10.0 years old; 9 males). The patient group underwent full battery of cognitive tests (Mini Mental State Examination, Rey Auditory Verbal Learning Test, Benton Visual Retention Test, and memory recall tests). The EEG time series were recorded using 19 electrodes positioned according to Maudsley system, similar to the 10-20 international system, at a sampling frequency of 128 Hz. EEGs were band-pass filtered with digital third-order Butterworth filter (forward and reverse filtering) between 0.5 and 30 Hz.

C. Recording Conditions Common to Both Datasets

In both data sets, all recording sessions were conducted with the subjects in an awake but resting state with eyes closed, and the length of the EEG recording was about 5 minutes, for each subject. The EEG technicians prevented the subjects from falling asleep (vigilance control). After recording, the EEG data has been carefully inspected. Indeed, EEG recordings are prone to a variety of artifacts, for example due to electronic smog, head movements, and muscular activity. For each patient, an EEG expert selected by visual inspection one segment of 20s artifact free EEG, blinded from the results



(a)



(b)

Fig. 4. Relative power distribution in various frequency bands for all the datasets, (a) Control group, (b) Mild cognitive impaired subjects

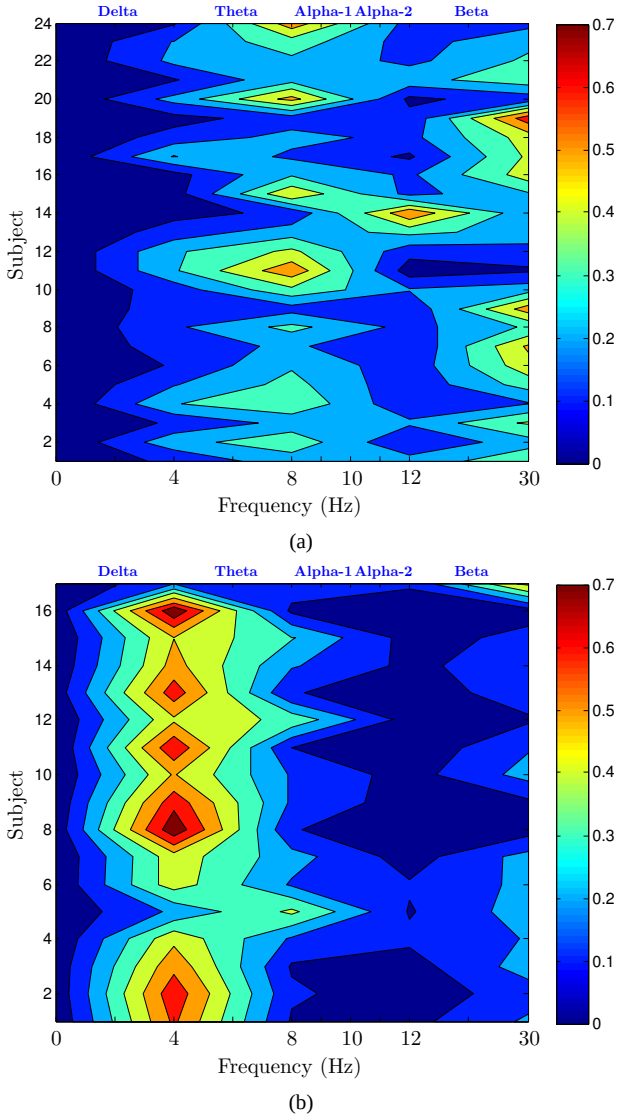


Fig. 5. Relative power distribution in various frequency bands for all the datasets, (a) Control group, (b) Mild Alzheimer's disease subjects

of the present study. Only those subjects were retained in the analysis whose EEG recordings contained at least 20s of artifact-free data. Based on this requirement, the number of subjects of EEG Dataset 1 was further reduced to 22 MCI patients and 38 control subjects; in EEG Dataset 2 no such reduction was required. From each subject in the two data sets, one artifact-free EEG segment of 20s was analyzed.

V. Results and Discussion

We compute relative power, compression ratios and LZ complexity for the EEG signals of all subjects. More specifically, we calculate those measures for all individual EEG channels, and then the measures are averaged over all channels; this results in average measures for all subjects. Our results are summarized in Table II and III and Figs. 6 and 7. In the analysis we also include two measures of EEG synchrony: stochastic event synchrony (ρ) [31, 32] and a Granger causality measure, i.e., full frequency directed transfer function (ffDTF) [33]; in an earlier study we observed that those two measures indicated statistically significant differences between MCI/MiAD and age-matched control subjects, for the data sets described in Section IV [25, 26]. It is noteworthy that, since the two data sets (MCI and MiAD) were obtained through different recording systems and at different hospitals, a direct comparison of the results obtained from MCI with those from mild AD is not straightforward.

In Table II we list statistics of the average measures, including the average computed across the entire subject groups and the standard deviation. We apply the Mann-Whitney test for the average measures between MCI and the reference subjects (Dataset 1) and MiAD and reference subjects (Dataset 2). The Mann-Whitney test allows us to investigate whether the statistics at hand (EEG measures) take different values between two subject populations. Low p-values indicate large difference in the medians of the two populations. The resulting p-values are listed in Table II. Since we conduct multiple statistical tests simultaneously, we need to apply statistical post-correction. We adopt Bonferroni post-correction [34], and multiply the p-values by the number of tests (11). In Table II we indicate which EEG measures remain statistically significant after post-correction.

Theta relative power is significantly larger in MCI patients compared to reference subjects, whereas beta power is significantly larger. In the MiAD patients the perturbations on EEG relative power are stronger: Delta and theta relative power is significantly larger than in the reference subjects, whereas alpha and beta power is significantly smaller. In other words, slowing occurs in both the MCI and MiAD patients, which is in agreement with earlier studies (see [10] for a review). The slowing effect can also readily be seen from the (normalized) EEG spectra, shown in Fig. 4 and 5 for dataset 1 and 2 respectively. The effect of slowing in the MiAD subjects is very clear from Fig. 5: Power is obviously more concentrated in theta-band in MiAD patients than in the age-matched control subjects. For the MCI patients (see Fig. 4), no such clear effect can be observed from the spectra; this is no surprise, since MCI is a less severe disease state than MiAD. However, one may notice a slight increase (decrease) in theta (beta) relative power in MCI patients. In both the MCI patients and control subjects, power is concentrated in low-frequency bands (delta and theta band) and in high-frequency band (beta band); high-frequency power (beta band) is much smaller in the MiAD patients. In summary, as in earlier studies (see [10] for a review), we observe slowing in MCI and MiAD EEG.

No significant effect on the complexity and regularity measures can be observed in MCI patients. On the other hand, the regularity measures and complexity measures are significantly larger and smaller respectively for MiAD patients than for control subject; in other words, the EEG signals of MiAD patients are significantly less complex than in healthy subjects. This observation is in agreement with several earlier studies (see [10] for a review).

We also try to classify patients vs. control subjects by means of the most discriminative EEG measures ($p < 0.05$). We test those measures individually and jointly for their discriminative ability. Table III shows the resulting classification performance

TABLE II
MEAN AND STANDARD DEVIATION VALUES OF COMPRESSION RATIO, LZ COMPLEXITY, RELATIVE POWER AND SYNCHRONY MEASURES. SENSITIVITY OF THE MEASURES IN DISCRIMINATING BETWEEN MCI AND MILD AD IS GIVEN IN LAST COLUMN. UNCORRECTED P-VALUES FROM MANN-WHITNEY TEST, WHERE * AND ** INDICATE $p < 0.05$ AND $p < 0.005$ RESPECTIVELY; † INDICATES P-VALUES THAT REMAIN SIGNIFICANT AFTER POST-CORRECTION (BONFERRONI, $p < 0.05$).

MCI *vs.* Control			
Measure	Control	MCI	p-value
1-D SPIHT CR	1.34±0.04	1.35±0.03	0.3077
2-D SPIHT CR	1.36±0.04	1.37±0.03	0.3778
2-D SPIHT+AC	1.36±0.04	1.37±0.03	0.4477
LZ complexity	0.65±0.07	0.62±0.09	0.0830
ρ	0.25±0.07	0.36±0.10	**0.00044****†
ffDTF	0.05±0.003	0.051±0.003	**0.0012****†
delta	0.20±0.06	0.21±0.06	0.2934
theta	0.08±0.03	0.12±0.04	**0.0001****†
alpha-1	0.07±0.03	0.08±0.03	0.1698
alpha-2	0.05±0.02	0.05±0.02	0.9939
beta	0.24±0.05	0.21±0.03	0.0116 *
Mild AD *vs.* Control			
Measure	Control	Mild AD	p-value
1-D SPIHT CR	1.09±0.01	1.12±0.04	**3.45** $\times 10^{-5}$**†
2-D SPIHT CR	1.11±0.02	1.15±0.04	**6.09** $\times 10^{-5}$**†
2-D SPIHT+AC	1.07±0.02	1.11±0.04	**4.86** $\times 10^{-5}$**†
LZ complexity	0.63±0.06	0.55±0.08	**0.0024****†
ρ	0.46±0.04	0.49±0.03	**0.0024****†
ffDTF	0.04±0.004	0.037±0.009	**0.0001****†
delta	0.001±0.004	0.017±0.01	**0.0029****†
theta	0.17±0.08	0.54±0.16	**8** $\times 10^{-7}$**†
alpha-1	0.32±0.12	0.18±0.10	**0.0009****†
alpha-2	0.17±0.11	0.06±0.02	**3.41** $\times 10^{-6}$**†
beta	0.33±0.14	0.18±0.11	**0.0006****†

TABLE III
CLASSIFICATION RATES FOR DISCRIMINANT ANALYSIS (DA) OF THE LOSSLESS COMPRESSION RATIOS, LZ COMPLEXITY AND RELATIVE POWER IN THETA BAND

MCI *vs.* Control			
Measure	**Linear DA**	**Quadratic DA**	**SVM**
theta	76.67%	76.67%	76.67%
ffDTF	63.33%	71.67%	78.33%
ρ	75%	75%	76.67%
ffDTF+ρ	76.67%	**83.33%**	80.00%
theta+ρ	78.33%	**83.33%**	80.00%
Mild AD *vs.* Control			
Measure	**Linear DA**	**Quadratic DA**	**SVM**
1-D SPIHT CR	80.49%	80.49%	80.49%
2-D SPIHT CR	82.93%	82.93%	**85.37%**
2-D SPIHT+AC CR	75.61%	80.49%	82.93%
LZ complexity	68.29%	68.29%	68.29%
theta	**95.12%**	**95.12%**	**95.12%**
ffDTF	58.54%	78.05%	82.93%
ρ	56.10%	63.41%	63.41%
ffDTF + ρ	65.85%	70.73%	78.05%
theta + ffDTF	**95.12%**	**92.68%**	**95.12%**
theta + ffDTF + 1-D SPIHT CR	**95.12%**	**92.68%**	**97.56%**

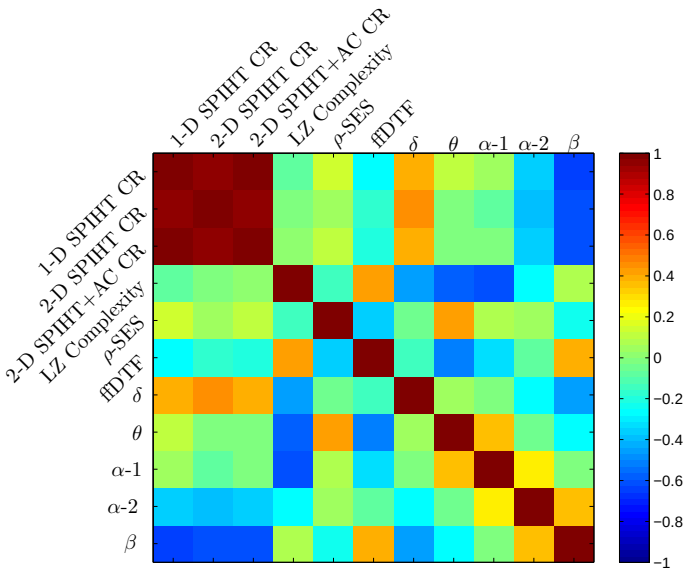
with linear and quadratic discriminant analysis, and support vector machine, determined through leaving-one-out crossvalidation [35]. Only the best performing combinations of EEG measures are listed. From Table III we can see that theta-band relative power yields good performance when used separately, and results in even better performance when combined with the most discriminative lossless-compression ratio and synchrony measure. The other relative power measures are less discriminative, for both datasets (not shown here); this observation is in agreement with the p-values listed in Table II. The compression ratios and LZ complexity fail to discriminate MCI patients from control subjects (not shown here). However, those measures yield good classification performance for the MiAD patients. Interestingly, the lossless-compression ratios result in better classification rates than LZ complexity; this may be explained as follows: LZ complexity is based on binary approximations of the continuous EEG signals, whereas the former are derived from accurate representations of the EEG, associated with lossless compression.

In order to gain more insight in the relationship between the different measures, we calculate the correlation between those measures (see Fig. 6). The correlation coefficient among each pair of measures is calculated as follows:

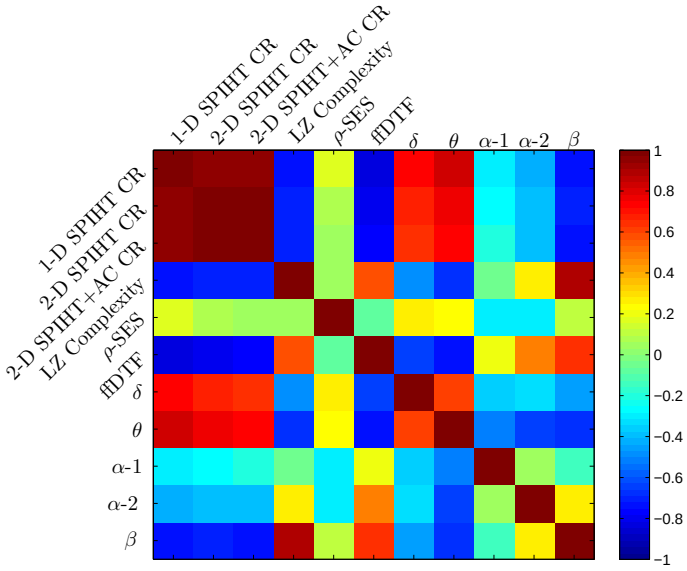
$$r_{ij} = \frac{1}{N_{subject}} \sum_{k=1}^{N_{subject}} \frac{m_i(k) - \bar{m}_i}{\sigma_i} \frac{m_j(k) - \bar{m}_j}{\sigma_j}, \quad (2)$$

where $m_i(k)$ and $m_j(k)$ is the average value of EEG measure i and j respectively for subject k, the sum is computed over all subjects, and $\bar{m}_i$, $\bar{m}_j$, σ_i and σ_j are the mean and standard deviation of m_i and m_j respectively. The resulting correlation coefficients are displayed in Fig. 6, for Dataset 1 and Dataset 2 separately. We also conduct the Pearson correlation test, to verify whether the correlations or anti-correlations are statistically significant. The resulting p-values are shown in Fig. 7 (logarithmic scale). Since we have multiple simultaneous tests, statistical post-correction is required. Again we adopt Bonferroni post-correction [34], and multiply the p-values by the number of tests (55).

As expected, the compression measures are significantly mutually correlated as all the schemes are based on the same principle; they are also significantly anti-correlated with LZ complexity in the MiAD dataset (Dataset 2).

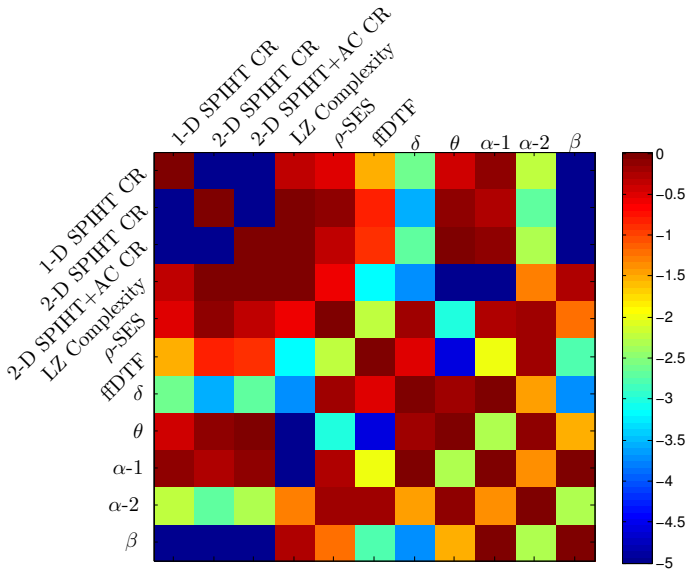


(a) MCI (Dataset 1)

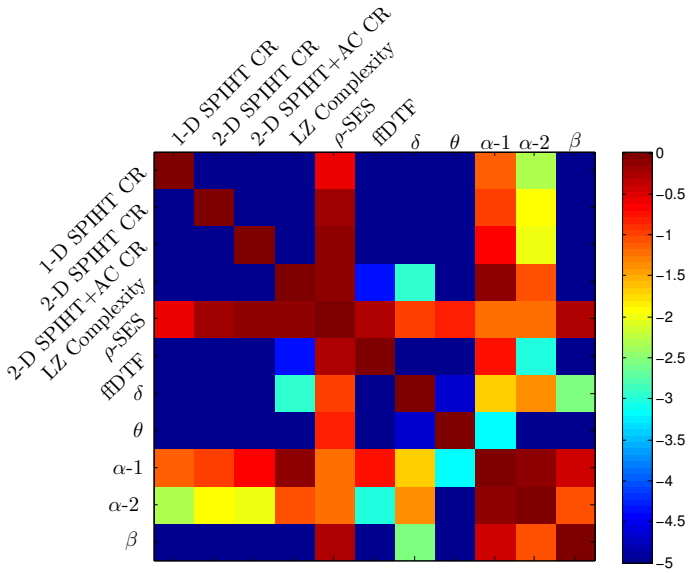


(b) Mild AD (Dataset 2)

Fig. 6. Correlation between the lossless compression ratios, LZ complexity, relative power in different bands, Granger causality (ffDTF), and stochastic event synchrony (ρ); red and blue indicate strong correlation and anti-correlation respectively.



(a) MCI (Dataset 1)



(b) Mild AD (Dataset 2)

Fig. 7. Pearson correlation test between the lossless compression ratios, LZ complexity, relative power in different bands, Granger causality (ffDTF), and stochastic event synchrony (ρ). The (uncorrected) p-values are shown on a logarithmic scale.

Interestingly, the compression ratios are significantly correlated with low-frequency relative power (delta and theta; MiAD) and anti-correlated with high-frequency relative power (beta; both data sets). Likewise LZ complexity is strongly anti-correlated with low-frequency relative power (delta and theta; both data sets) and correlated with high-frequency relative power (beta; MiAD). Taken together, this observation strongly suggests that slowing and loss of complexity in AD EEG are *not independent* phenomena but are strongly related; to the best of our knowledge, this observation has not been reported before in the literature.

Perhaps surprisingly, Granger causality (ffDTF) [33] is significantly correlated with LZ complexity and high-frequency relative power (MiAD), and significantly anti-correlated with lossless compression ratios (MiAD) and low-frequency relative power (both datasets). We believe that this observation has not been documented yet. We conjecture that the observed statistical (anti-)correlation between ffDTF and the other measures is an artefact of the multivariate model underlying Granger causality (and ffDTF in particular). More specifically, Granger causality is derived from a multivariate autoregressive model (MVAR). The order of the latter needs to be kept low, since the coefficients of the MVAR need to be inferred from a short EEG segment; high-order MVARs contain many coefficient, which cannot be reliably inferred from the limited amount of data. Low-order MVARs have short memory, and cannot capture low-frequency components in the EEG. Consequently Granger causality may underestimate the correlation among

brain signals when the EEG contains strong low-frequency components.

Stochastic event synchrony (ρ) [31, 32] seems to be uncorrelated with the other measures (both datasets), and therefore, it may provide complementary information.

VI. Conclusion

In this study, we investigated the use of relative power, LZ complexity, and lossless compression ratio as EEG markers for MCI and Mild AD. Lossless compression ratio is shown to be discriminative for Mild AD, whereas it is not discriminative for MCI. On the other hand, theta-band relative power was observed to be statistically larger in MCI and Mild AD patients than in control subjects. Maximum discrimination is obtained by combining the compression ratio, relative power and synchrony measures (Granger causality and/or stochastic event synchrony).

We would like to reiterate, however, that the two data sets analyzed (MCI and MiAD) were obtained through different recording systems and at different hospitals; a direct comparison of the results obtained from MCI with those from mild AD is therefore difficult. On the other hand, since the data sets are independent, our observations are probably not due to particularities of the recording systems and/or procedures at the hospitals.

Interestingly, compression ratios were found to be significantly correlated to delta and theta band relative power, showing their strong correlation with relative power at low frequencies; also strong anti-correlation between compression ratios and beta relative power was observed. Therefore, slowing and loss of complexity in the EEG of MCI and MiAD patients may be strongly related phenomena.

More generally, this study also underlines the importance of analyzing MCI and AD EEG by means of a variety of statistical measures (relative power, complexity/regularity measures, synchrony measures), in order to detect potential correlations between various observed phenomena associated with MCI and AD.

References

[1] M. P. Mattson, "Pathways towards and away from alzheimer's disease," *Nature*, vol. 430, pp. 631 – 639, 2004.

[2] P. D. Meek, K. McKeithan, and G. T. Schumock, "Economics considerations of alzheimer's disease," *Pharmacotherapy*, vol. 18 (2 Pt 2), pp. 68–73, Mar-Apr 1998.

[3] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, H.M. Arrighi, "Forecasting the global burden of Alzheimer's disease," *Alzheimer's and Dementia* 3(3):186-91, 2007.

[4] A.R. Frank and R.C. Petersen, "Mild cognitive impairment", *Handb Clin Neurology*, vol. 89:217–21, 2008.

[5] R.C. Petersen, "Early diagnosis of Alzheimer's disease: is MCI too late?", *Curr Alzheimer Res.*, vol. 6(4):324–30, 2009.

[6] A. Shimokawa, N. Yatomi, S. Anamizu, S. Torii, H. Isono, Y. Sugai, and M. Kohno, "Influence of deteriorating ability of emotional comprehension on interpersonal behavior in alzheimer-type dementia," *Brain and Cognition*, vol. 47, no. 3, pp. 423 – 433, 2001.

[7] K. Palmer, A. K. Berger, R. Monastero, B. Winblad, L. Bäckman, and L. Fratiglioni, "Predictors of progression from mild cognitive impairment to Alzheimer disease," *Neurology*, vol. 68, no. 19, pp. 1596–1602, 2007.

[8] K.A. Wollen, "Alzheimer's disease: the pros and cons of pharmaceutical, nutritional, botanical, and stimulatory therapies, with a discussion of treatment strategies from the perspective of patients and practitioners," *Altern Med Rev.*, 15(3):223–44, 2010.

[9] J. Jeong, "EEG dynamics in patients with alzheimers disease," *Clinical Neurophysiology*, vol. 115, pp. 1490 – 1505, 2004.

[10] J. Dauwels, F. Vialatte, and A. Cichocki, "Diagnosis of alzheimers disease from EEG signals: Where are we standing?" *Current Alzheimer Research*, vol. 7, pp. 487–505, 2010.

[11] P. L. Nunez and R. Srinivasan, *Electric fields of the brain*. Oxford university press, 2006.

[12] A. Lempel and J. Ziv, "On the complexity of finite sequences," *IEEE Transctions on Information Theory*, vol. IT-22, pp. 75–81, 1976.

[13] R. Hornero, D. Abásolo, J. Escuredo, and C. Gómez, "Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with alzheimer's disease," *Philosophical Transactions of The Royal Society A*, vol. 367, pp. 317–336, 2008.

[14] K. Srinivasan and M. R. Reddy, "Efficient preprocessing technique for lossless real-time EEG compression," *Electronics Letters*, vol. 46, no. 1, pp. 26–27, Jan. 2010.

[15] K. Srinivasan, J. Dauwels, and M. R. Reddy, "A two-dimensional approach for lossless EEG compression," *Biomedical signal processing and control*, in press, 2011.

[16] S. G. Mallat, "A theory for multiresolution signal decomposition:the wavelet representation," *IEEE Transactions Pattern analysis and machine intelligence*, vol. 2, no. 7, pp. 674–693, July 1989.

[17] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.*, vol. 4, no. 3, pp. 245–267, 1998.

[18] M. D. Adams and F. Kossentini, "Reversible Integer-to-Integer wavelet transforms for image compression: Performance evaluation and analysis," *IEEE Transactions on Image Processing*, vol. 9, no. 6, pp. 1010–1024, 2000.

[19] R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers," *Appl. Comput. Harmon. Anal.*, vol. 5, no. 3, pp. 332–369, 1998.

[20] A. Said and W. Pearlman, "A new, fast and efficient image codec basec on set partitioning in hierarchial trees," *IEEE transactions on circuits and systems for video Technology*, vol. 6, no. 3, pp. 243–250, June 1996.

[21] Z. Lu, D. Y. Kim, and W. A. Pearlman, "Wavelet compression of ECG signals by the set partitioning in hierarchical trees algorithm," *IEEE Transactions on Bio-Medical Engineering*, vol. 47, no. 7, pp. 849 – 856, July 2000.

[22] A. Cichocki, S. L. Shishkin, T. Musha, Z. Leonowicz, T. Asada, and T. Kurachi, "EEG filtering based on blind source separation (BSS) for early detection of alzheimer's disease," *Clinical Neurophysiology*, vol. 116, no. 3, pp. 729–737, 2005.

[23] T. Musha, T. Asada, F. Yamashita, T. Kinoshita, Z. Chen, H. Matsuda, M. Uno, and W. R. Shankle, "A new EEG method for estimating cortical neuronal impairment that is sensitive to early stage alzheimer's disease," *Clinical Neurophysiology*, vol. 113, no. 7, pp. 1052–1058, 2002.

[24] F. Vialatte, A. Cichocki, G. Dreyfus, T. Musha, T. Rutkowski, and R. Gervais, "Blind source separation and sparse bump modelling of time frequency representation of eeg signals: New tools for early detection of alzheimer's disease," in *IEEE Workshop on Machine Learning for Signal Processing*, 2005, pp. 27 –32.

[25] J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki, "A comparative study of synchrony measures for the early diagnosis of alzheimer's disease based on EEG," *NeuroImage*, vol. 49, pp. 668–693, 2010.

[26] J. Dauwels, F. Vialatte, C. Latchoumane, J. Jeong, and A. Cichocki, "EEG synchrony analysis for early diagnosis of alzheimer's disease: A study with several synchrony measures and EEG data sets," in *31st Annual International Conference of the IEEE EMBS, Minneapolis, Minnesota, USA*, september 2009, pp. 2224 – 2227.

[27] M. J. Hogan, G. R. J. Swanwick, J. Kaiser, M. Rowan, and B. Lawlor, "Memory-related EEG power and coherence reductions in mild alzheimer's disease," *International Journal of Psychophysiology*, vol. 49, no. 2, pp. 147 – 163, 2003.

[28] R. M. Chapman, G. H. Nowlis, J. W. McCrary, J. A. Chapman, T. C. Sandoval, M. D. Guillily, M. N. Gardner, and L. A. Reilly, "Brain event-related potentials: Diagnosing early-stage alzheimer's disease," *Neurobiology of aging*, vol. 28, no. 2, pp. 194–201, 2007.

[29] C. Goh, E. Ifeachor, G. Henderson, C. Latchoumane, J. Jeong, C. Bigan, M. Besleaga, N. Hudson, P. Capotosto, and S. Wimalaratna, "Characterisation of EEG at different stages of Alzheimer's disease (AD)," *Clinical Neurophysiology*, vol. 117, pp. 138–139, 2006.

[30] G. Henderson, E. Ifeachor, N. Hudson, C. Goh, N. Outram, S. Wimalaratna, C. Del Percio, and F. Vecchio, "Development and assessment of methods for detecting dementia using the human electroencephalogram," *IEEE Transaction on Biomedical Engineering*, vol. 53, pp. 1557–1568, 2006.

[31] J. Dauwels, F. Vialatte, T. Weber, and A. Cichocki, "Quantifying statistical interdependence by message passing on graphs, Part I: One-dimensional point processes," *Neural Computation* 21(8):2203–2268, 2009.

[32] J. Dauwels, F. Vialatte, T. Weber, T. Musha, and A. Cichocki, "Quantifying statistical interdependence by message passing on graphs, Part II: Multi-dimensional point processes," *Neural Computation* 21(8):2152–2202, 2009.

[33] Kamiński M. and Blinowska K. J., 1991. A new method of the description of the information flow in the brain structures. Biol. Cybern. 65, 203–210.

[34] Bonferroni C. E., 1936. Teoria statistica delle classi e calcolo delle probabilitá. Pubblicazioni del R. Instituto Superiore di Scienze Economiche e Commerciali di Firenze 8, 3–62.

[35] R. Duda, P. Hart, and D. Stork, *Pattern classification*. Wiley-Interscience, 2000.